Biology of Rhodococcus

Bearbeitet von
Héctor M Alvarez

1. Auflage 2010. Buch. xii, 368 S. Hardcover
ISBN 978 3 642 12936 0
Format (B x L): 15,5 x 23,5 cm
Gewicht: 800 g

Weitere Fachgebiete > Technik > Biotechnologie

Zu Inhaltsverzeichnis

schnell und portofrei erhältlich bei

beck-shop.de
DIE FACHBUCHHANDLUNG

Die Online-Fachbuchhandlung beck-shop.de ist spezialisiert auf Fachbücher, insbesondere Recht, Steuern und Wirtschaft. Im Sortiment finden Sie alle Medien (Bücher, Zeitschriften, CDs, eBooks, etc.) aller Verlage. Ergänzt wird das Programm durch Services wie Neuerscheinungsdienst oder Zusammenstellungen von Büchern zu Sonderpreisen. Der Shop führt mehr als 8 Millionen Produkte.
The Rhodococcal Cell Envelope: Composition, Organisation and Biosynthesis

Iain C. Sutcliffe, Alistair K. Brown, and Lynn G. Dover

Contents

1 Introduction .............................................................................................................. 30
2 Cell Envelope Composition in the Genus Rhodococcus:
   Covalently Associated Components .................................................................. 30
   2.1 Mycolic Acids ................................................................................................. 30
   2.2 The Peptidoglycan–Arabinogalactan Complex ........................................... 32
3 Organisation of the Rhodococcal Cell Envelope .................................................. 34
4 Non-Covalently Associated Cell Envelope Components ....................................... 37
   4.1 Channel Forming Porins .................................................................................. 37
   4.2 Lipoglycans ..................................................................................................... 38
   4.3 Cell Envelope Lipids ....................................................................................... 39
   4.4 Capsules and Cell Envelope Polysaccharides ................................................. 40
   4.5 Lipoproteins ................................................................................................... 41
5 Biosynthesis of Key Cell Envelope Components ................................................... 42
   5.1 Mycolic Acid Biosynthesis ............................................................................. 43
   5.2 Arabinogalactan Biosynthesis ........................................................................ 47
   5.3 LAM Biosynthesis .......................................................................................... 56
6 Concluding Comments ............................................................................................ 59
References ................................................................................................................. 60

Abstract The cell envelopes of rhodococci and their closest relatives are dominated by the presence of large branched chain fatty acids, the mycolic acids. Here we review the structural features underlying the incorporation of the mycolic acids into the rhodococcal cell envelope, notably their covalent anchoring to the peptidoglycan–arabinogalactan complex and their organisation into an outer lipid permeability barrier. Rhodococcal cell envelopes also accommodate diverse non-covalently associated components such as channel-forming porin proteins, free lipids,

I.C. Sutcliffe (✉), A.K. Brown, and L.G. Dover
Biomolecular and Biomedical Research Centre, School of Applied Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, UK
e-mail: iain.sutcliffe@northumbria.ac.uk; alistair.brown@northumbria.ac.uk;
lynn.dover@northumbria.ac.uk

lipoglycans, lipoproteins and capsules or cell envelope polysaccharides. Based on the extensive studies of cell envelope biogenesis in corynebacteria and mycobacteria, we have used a comparative genomics approach to examine the pathways for the biosynthesis of the major cell envelope components of *Rhodococcus jostii* RHA1.

1 Introduction

The genus *Rhodococcus* belongs to the suborder Corynebacterineae, a distinctive lineage within the phylum Actinobacteria (Gürtler et al. 2004; Jones and Goodfellow in press; Zhi et al. 2009). The members of this taxon are characterised by distinctive cell envelopes typically dominated by large branched chain lipids, the mycolic acids, and collectively they are often referred to as the mycolata. In addition to the mycolic acids, these bacteria share a number of other common cell envelope features, most notably an arabinogalactan (AG) cell wall polysaccharide that is covalently attached to the cell wall peptidoglycan and in turn provides a scaffold for the covalent anchoring of mycolic acids. Thus, the chemistry and organisation of these components in a distinctive cell envelope architecture represents one of the defining features of the biology of the mycolata. The mycolate cell envelope has received extensive study in the context of understanding the targets of several crucial antibiotics that are used against the pathogenic mycobacteria, most notably *Mycobacterium tuberculosis* (Dover et al. 2008a). These studies have been vital in providing comparative insights into cell envelope biology in the genus *Rhodococcus* (Sutcliffe 1998; Gürtler et al. 2004). Here we review recent progress in understanding of the composition, architecture and biosynthesis of the mycolate cell envelope, with particular reference to the rhodococci. In particular insights into the biosynthesis of these cell envelope components are now possible using a comparative genomics approach, based on the recently published genome sequence of *Rhodococcus jostii* RHA1 (McLeod et al. 2006).

2 Cell Envelope Composition in the Genus *Rhodococcus*: Covalently Associated Components

The mycolyl–arabinogalactan–peptidoglycan complex represents the defining covalently interlinked structure of the cell envelope of the mycolata. These components and their linkages are reviewed in the following sections.

2.1 Mycolic Acids

Mycolic acids are 2-alkyl branched, 3-hydroxy long chain fatty acids, which vary in size and complexity with the different genera of the mycolata (Fig. 1). Those of the
Corynebacteria are typically the smallest (size range 22–38 total carbons) and those of the mycobacteria are the most complex, with a size range of 60–90 carbons and a greater diversity of meromycolate chain functional groups such as cyclopropane, methoxy- and keto- modifications (Dover et al. 2004; Gürtler et al. 2004; Takayama

Fig. 1 Structures of representative mycolic acids from Corynebacterium, Rhodococcus and Mycobacterium species, illustrating their differing complexity. M. tuberculosis methoxymycolate is an example of the most complex mycolic acids. The less complex M. smegmatis mycolates encompass either double or single unsaturations whilst retaining the longer chain length. Rhodococcus sp. mycolates are of an intermediate size. They present an aliphatic 2-alkyl chain varying from 12 to 16 carbons, whilst the 3-hydroxyl meromycolate typically contains 18–40 carbons (i.e. x, y and z total 18–40). The rhodomeromycolates have relatively simple modifications containing up to four unsaturations at presently undetermined positions. Corynebacterium sp. mycolates are the simplest known mycolates. For simplicity, corynemycolates are represented as the typical 32–36 carbon aliphatic mycolic acids but a proportion of the total cell wall mycolates may also contain single or double unsaturations.
et al. 2005). Members of the genus *Rhodococcus* produce mycolic acids of an intermediate size, typically with 28–54 carbons in total (Alshamaony et al. 1976; Klatte et al. 1994; Sutcliffe 1998; Stratton et al. 1999; Nishiuchi et al. 2000). One important feature of mycolic acid structure is the length of the main meromycolate chain compared to the alkyl side branch. In rhodococci, the alkyl side branch is typically a saturated alkyl chain of 10–16 carbons in length, whereas the meromycolate is a longer chain (C20–C42) with up to four carbon–carbon double bonds (Fig. 1). Thus consideration of the detailed profile of the mycolates present in *Rhodococcus rhodochrous* (Stratton et al. 1999) reveals that the lengths of each meromycolate chain will extend beyond that of its alkyl partner (Fig. 2) and similar data are evident for other rhodococci (Nishiuchi et al. 2000). The positions of the unsaturated bonds have not been determined unequivocally but studies of the mycolic acids of *Nocardia asteroides* suggest it is likely that they are localised in the distal regions of the meromycolate chain (Minnikin and O’Donnell 1984; Sutcliffe 1998). This would mean the region of the meromycolate chain proximal to the ester-linked terminus is effectively a saturated chain and would pack closely with the saturated alkyl branch. Thus these features need to be considered when predicting the arrangement of the mycolates esterified to AG within the rhodococcal cell envelope (see below).

### 2.2 The Peptidoglycan–Arabinogalactan Complex

As in other members of the mycolata, the peptidoglycan structure of members of the genus *Rhodococcus* has been determined to be of the A1γ type, i.e. with the diamino acid meso-diaminopimelic acid forming direct cross-linkages between...
the stem peptides (Jones and Goodfellow in press). The muramic acid residues of the glycan strands are N-glycolyated, which is a comparatively unusual peptido-glycan modification (Uchida and Aida 1979; Vollmer 2008; Jones and Goodfellow in press). N-glycolylation is likely carried out during cytoplasmic peptidoglycan precursor biosynthesis, prior to lipid II formation (Raymond et al. 2005; Vollmer 2008; Jones and Goodfellow in press). A clear orthologue of the NamH protein recently identified as the Mycobacterium smegmatis oxygen-dependent hydroxylase responsible for N-glycolylation was identified in the R. jostii RHA1 genome (Raymond et al. 2005; RHA1_ro04045). Although the physiological function of N-glycolylation remains unclear, it is notable that deletion of namH in M. smegmatis increased susceptibility to lysozyme and β-lactam antibiotics (Raymond et al. 2005). As an extra hydroxyl group is introduced to the glycan chain, there may be additional hydrogen binding possibilities within the cell envelope, which could contribute to novel aspects of its supramolecular organisation.

The AG of the cell envelope is phosphodiester linked to the peptidoglycan by a well conserved linker unit (LU) of L-rhamnose-D- N-acetylglucosamine phosphate (Daffe´ et al. 1993). Mycobacterial AG has been extensively structurally characterised as the scaffold for the attachment of the mycolic acids (Besra et al. 1995; Bhamidi et al. 2008) and the target for the action of the anti-tubercular ethambutol (EMB; Takayama and Kilburn 1989). The heteropolymer is divided into distinct homopolymer galactan and arabinan domains. A galactan anchored to the peptidoglycan via the LU will typically carry three arabinan domains, the branched termini of which carry the mycolic acids. Comparatively little is known of the fine structure of AG from most mycolic acid containing bacteria, although an important comparative study revealed that the AG of Rhodococcus equi and R. rhodochrous have a similar domain organisation of a linear homogalactan bearing discrete arabinan domains (Daffe´ et al. 1993). The galactan of R. equi contained both 1→3, 1→5 and 1→6 glycosidic linkages whereas that of R. rhodochrous contained 1→2 and 1→5 linkages. Further galactan diversity was revealed in the galactans of Nocardia spp. The arabinan domains of the AG also exhibit similar variations, that of mycobacteria typically present pentaarabinosyl branched termini, which can carry four mycolates each (Besra et al. 1995; Bhamidi et al. 2008). In contrast, in R. equi, a range of arabinose termini are present including a linear arabinan, a triarabinosyl branched terminus and termini bearing mannose caps (Daffe´ et al. 1993). As minor variations on the AG core structure have been reported in Tsukamurella paurometabolum (Tropis et al. 2005a), it appears that AG is likely to be subject to genus and species-specific variations in fine structure, which may also extend to the presence of substituents such as succinate (Bhamidi et al. 2008). This may have implications regarding the extent to which rhodococcal cells are covered with covalently bound mycolates, the significance of which is discussed later.

In addition to the “secondary” cell wall polymers, Gram-positive bacteria also anchor proteins to their peptidoglycan through the action of sortase transpeptidase enzymes (Marraffini et al. 2006). Sortase substrates have a characteristic LPXTG motif (or variants thereof), of which the threonine residue is targeted for the
transpeptidation reaction. Sortase-mediated anchoring of proteins to the cell wall does not appear to be as prominent in Actinobacteria compared to Firmicute Gram-positive bacteria; important examples include the anchoring of the larger chaplins during production of aerial hyphae by *Streptomyces coelicolor* (Di Berardo et al. 2008) and the polymerisation and anchoring of pili in *Corynebacterium diphtheriae* (Mandlik et al. 2008a, b). Bioinformatic analyses have identified only a single sortase encoded in the *R. jostii* RHA1 genome (see PFAM family PF04203 at http://pfam.sanger.ac.uk//family/pf04203) and no members of the LPXTG family (PF00746) of canonical sortase substrates. In *R. jostii*, the sortase protein RHA1_ro03500 is apparently encoded in an operon with an adjacent coding sequence (RHA1_ro03501), which has the requisite C-terminal features, including an HPETG motif, that suggest RHA1_ro03501 might be the sortase substrate. However, this pairing aside, it is clear that sortase-anchored proteins are not numerically abundant in the predicted proteome of *R. jostii*.

3 Organisation of the Rhodococcal Cell Envelope

Determining how significant quantities of high molecular weight lipids (i.e. the mycolic acids) are organised within the cell envelope presented a significant challenge in earlier studies of the mycolata. However, the landmark studies of Minnikin (1982, 1991) provided a model that, following extensive biochemical, biophysical and structural analyses, has become accepted as the definitive model of the mycobacterial cell envelope (Brennan and Nikaido 1995; Daffé and Draper 1998). Subsequently, this model has been applied to models of the corynebacterial (Puech et al. 2001; Dover et al. 2004) and rhodococcal cell envelopes (Sutcliffe 1997, 1998). Recently, excellent cryo-electron microscopic studies have provided clear visualisations of this structure for both *Corynebacterium* and *Mycobacterium* spp. (Hoffmann et al. 2008; Zuber et al. 2008) and variations on this theme are most likely applicable to all mycolic-acid containing actinomycetes.

The central tenet of the Minnikin model is that the mycolic acids covalently attached to the AG have a perpendicular orientation with respect to the plane of the plasma membrane (Fig. 3). Thus, the mycolates form the basis of a second hydrophobic permeability barrier outside of the plasma membrane. This structure is analogous to the outer membranes of Gram-negative bacteria but is chemically and structurally distinct, most notably in that the defining feature of the permeability barrier is not a bilayer but the monolayer of bound mycolates. Nevertheless, depending on the extent to which bound mycolates are able to provide coverage of the whole bacterial cell surface, additional components may be needed to “plug” potential gaps in the mycolyl layer, a role proposed to be taken most likely by mycolic acid containing-glycolipids (see Sect. 4.3; Sutcliffe 1998; Puech et al. 2001; Zuber et al. 2008). Indeed, trehalose mycolates are likely to act as carriers for incorporation on newly synthesised mycolic acids into the cell envelope (Tropis et al. 2005b).
**Fig. 3** *Model for the organisation of the rhodococcal cell envelope.* This adaptation of the classic Minnikin model (Minnikin 1991) emphasises that vertically orientated mycolic acids form the basis of the outer lipid permeability barrier. Components of the model in panel A are identified in panel B. No specific conformation for the peptidoglycan–AG complex is favoured in this schematic representation. Consideration of the sizes and conformations of rhodococcal mycolates suggest that there may be filler lipids (and possibly lipoglycans and lipoproteins) associated with the outer surface of mycolate layer (*grey box*) but this remains hypothetical. Not shown are the outermost layers composed of the cell envelope polysaccharides and/or capsules that are known to be present in many rhodococci.
A second consideration is the extent to which the longer meromycolate chains of the mycolic acids project beyond the alkyl branches that are packed alongside them (Sect. 2.1) and, with longer chain mycolates, the extent to which folded conformations can be adopted (Villeneuve et al. 2007). In rhodococci, we propose that as the typical length of these meromycolate extensions corresponds most closely with the lengths of typical fatty acyl lipids (Fig. 2), it is likely that the permeability barrier provided by the mycolate layer is bolstered externally by the interaction of smaller amphiphiles (e.g., acylglycerol-based lipids), which may vary on a species and strain-specific basis, as discussed previously (Sutcliffe 1998; Puech et al. 2001). Moreover, it is notable that rhodococcal mycolates lack both the length and chain modifications necessary to achieve the more complex “Z” and “W” type conformations (e.g., see Villeneuve et al. 2007). Genus, species and strain-specific variations in the length of the mycolic acids (Dover et al. 2004; Gurtler et al. 2004; Takayama et al. 2005) dictate the precise configurations of the mycolates and so the extent to which this necessitates their interactions with “filler” lipids is thus likely to vary in a species-specific manner. Moreover, the mycolate layer is highly unlikely to be a static barrier, and the permeability of this barrier is likely to be regulated, consistent with studies showing that mycolic acid composition can vary with growth conditions (Sutcliffe 1998; Sokolovská et al. 2003; Stratton et al. 2003).

The representation in Fig. 3 is in reasonable agreement with the recently proposed “zippered” version of the classic Minnikin model, wherein free lipids are shown intercalated with the mycolates (Zuber et al. 2008). This model is in reasonable agreement with the measured thicknesses of the outer permeability barrier in corynebacteria (4–5 nm) and mycobacteria (7–8 nm) (Hoffmann et al. 2008; Zuber et al. 2008). However, specific studies are needed to further define the details of the organisation of the rhodococcal cell envelope. Such studies will be of considerable interest given that the length of the rhodococcal mycolates represents an intermediate stage between those of the corynebacterial and mycobacterial species whose cell envelopes have been most extensively studied (Hoffmann et al. 2008; Zuber et al. 2008). Thus the extent to which the projecting meromycolate chains interact with covering amphiphiles and other outermost components (notably capsules and other polysaccharides; see Sect. 4.4) is an important question for future study, particularly as these features will profoundly influence the cell surface hydrophobicity and thus the possible biotechnological applications of rhodococcal strains.

The Minnikin model focuses primarily on the organisation of the mycolates within the cell envelope. Thus, the organisation of the peptidoglycan in the mycolata has traditionally been assumed to be comparable to that of the peptidoglycan in other bacteria, that is, a layered structured wherein the peptidoglycan strands are orientated in parallel with the plane of the plasma membrane (Vollmer and Höltje 2004). Alternatively, it has been proposed that helical glycan strands of the peptidoglycan may have a novel vertical orientation (Dmitriev et al. 2005) and that in mycobacteria this allows for helical galactan chains of the AG to be intercalated within a grid of glycan “pillars” (Dmitriev et al. 2000). As yet this novel “scaffold” hypothesis has not yet received extensive support from studies on other organisms (Gan et al. 2008; Hayhurst et al. 2008) and it remains to be verified in the mycolata.
Finally, the presence of an outer lipid permeability barrier suggests that the location between this layer and the plasma membrane should be considered a “pseudo-periplasm” and also dictates that pathways must exist for both solute uptake (see Sect. 4.1) and secretion (notably of proteins but also of DNA). However, although there is a growing understanding of pathways of protein secretion in mycobacteria (DiGiuseppe Champion and Cox 2007), no specific systems for the export of proteins beyond the pseudoperiplasm have yet been identified.

4 Non-Covalently Associated Cell Envelope Components

As illustrated in Fig. 3, the mycolyl–arabinogalactan–peptidoglycan complex provides a scaffold upon which several crucial classes of cell envelope component can be localised. These non-covalently associated cell envelope components are reviewed in the following sections.

4.1 Channel Forming Porins

The organisation of the covalently linked mycolates and other cell envelope lipids into an outer lipid permeability barrier suggests that channel forming proteins (porins) need to be present to allow the accumulation of hydrophilic solutes. This prediction was confirmed first for Mycobacterium chelonae (Trias et al. 1992) and subsequently channel forming proteins have been identified in a considerable range of mycolic acid containing actinomycetes (Nikaido 2003; Ziegler et al. 2008). The channel forming proteins that have been best characterised to date, that is the cation selective MspA channel of M. smegmatis (Faller et al. 2004) and the anion selective channel PorB of Corynebacterium glutamicum (Ziegler et al. 2008), are both relatively small proteins that oligomerise to form their respective channels. MspA forms a novel octameric 16-stranded $\beta$-barrel structure (Faller et al. 2004) whereas PorB forms a putative pentameric structure that is unusual in containing $\alpha$-helices (Ziegler et al. 2008).

Three studies have confirmed the presence of porins in rhodococci. Cation selective channels have been isolated from Rhodococcus (formerly Nocardia), corynebacteroides (Rieß and Benz 2000) and Rhodococcus erythropolis (Lichtinger et al. 2000), whilst complementary anion and cation selective channels were identified in organic solvent extracts from R. equi (Rieß et al. 2003). The ca. 2-nm wide cation selective channels of R. equi and R. erythropolis have similar biophysical properties. However, the N-terminal sequence determined for the R. erythropolis protein (Lichtinger et al. 2000) does not at present generate any significant homology matches to any known proteins. M. smegmatis MspA is the prototype of a porin family (PF09203; http://pfam.sanger.ac.uk//family/PF09203) and three members of this family (RHA1_ro03127; RHA1_ro04074; RHA1_ro08561) with
significant sequence homology to *M. smegmatis* MspA can also be identified in the *R. jostii* genome.

The identity of the anion selective channel of *R. equi* has yet to be determined, and it is notable that the known anion selective channels of corynebacteria appear to be unique to this genus (Ziegler et al. 2008). Thus other novel channel forming proteins are likely to be present in rhodococcal cell envelopes.

In addition to the above porins, Rv1968 from *M. tuberculosis* was recently described as the prototype for a new class of channel forming proteins (Siroy et al. 2008). In vitro, Rv1968 forms channels with a weak selectivity for cations. A single significant homologue of Rv1968 is encoded in the *R. jostii* genome (RHA1_ro00932; 145/308 amino acid sequence identity) suggesting that channels of this type are also likely to be present in rhodococcal cell envelopes.

### 4.2 Lipoglycans

The cell envelopes of most, but not all, Actinobacteria bacteria are characterised by the presence of membrane-anchored polysaccharides, the lipoglycans (Sutcliffe 1994; Rahman et al. 2009). In all mycolic acid-containing Actinobacteria studied to date, the lipoglycans present belong to the lipoarabinomannan (LAM) family (Nigou et al. 2003, 2008; Gilleron et al. 2005). As with AG, this lipoglycan family is characterised by a conserved core structure that then exhibits considerable species and strain-specific variation in fine structure. The core structure is defined by the presence of a phosphatidylinositolmannoside-based lipid anchor, which is extended into a 1→6 linked mannan domain of variable length (Nigou et al. 2003). In addition to mannose side chains, this mannan core will also carry arabinose or arabinan branches, which in turn may carry a variety of substituent motifs, most notable mannose caps in some strains (Nigou et al. 2003).

Three species of *Rhodococcus* have been investigated as to their lipoglycan structure. The structure of the LAM-like lipoglycan of *R. equi* (ReqLAM) was found to be the first known example of a ‘truncated’ LAM wherein the typical phosphatidylinositol-anchored lipomannan core is decorated with 1→2 linked mannose branches, some of which bear a single capping by *t*-arabinofuranose residues (Garton et al. 2002). Thus, the substantial arabinan domains of mycobacterial LAMs are not present in this structure, which has immunomodulatory properties that may be relevant to the pathogenesis of disease in foals (Garton et al. 2002; Nigou et al. 2008). Likewise, the lipoarabinomannan of *Rhodococcus ruber* (RruLAM) is also a truncated LAM structure in which the lipomannan core is directly substituted with *t*-arabinofuranose residues (Gibson et al. 2003b). These truncated LAMs are thus closely structurally related yet distinct from each other and it is apparent that truncated LAMs represent as distinct subfamily within the LAM archetype (Gilleron et al. 2005). A LAM-like lipoglycan has also been identified in *Rhodococcus rhodnii* (Flaherty et al. 1996) but has not yet received full structural
characterisation. However, the arabinose content determined by gas chromatography suggests that this LAM might be more extensively arabinosylated (Flaherty et al. 1996). Notably, in these rhodococci, the LAM-like lipoglycans appear to be the sole membrane-anchored polysaccharides, whereas in mycobacteria, the LAM is accompanied in the membrane by a structurally inter-related lipomannan (Nigou et al. 2003).

The physiological functions of these lipoglycans remain obscure (Sutcliffe 2005) but recent advance in understanding the genetic basis of LAM biosynthesis have led to the generation of mutants abrogated in various stages of the LAM biosynthesis pathway. Mutation of an early stage in LAM biosynthesis (assembly of the mannan core) was achieved in C. glutamicum, although at a low frequency of homologous recombination and the mutant obtained exhibited notably poor in vitro growth (Mishra et al. 2008b). This provides the clearest evidence to date that LAM lipoglycans may not be essential for the growth of mycolic acid-containing Actinobacteria but that they are likely to be necessary for optimal growth.

By analogy with both lipoteichoic acids and other lipoglycans (Sutcliffe 1994; Rahman et al. 2009), it has generally been assumed that LAM family lipoglycans are anchored to the outer leaflet of the plasma membrane with the glycan polymer projecting into the ‘pseudo-periplasm’. However, it remains possible that a subfraction of lipoglycans is surface exposed through trafficking and intercalation of the lipid anchor into the outer mycolate-based lipid layer (Fig. 3), as recently discussed for mycobacterial LAM (Pitarque et al. 2008). These two subfractions can be usefully distinguished as ‘parietal’ LAM (associated with the mycolate layer) and ‘cellular’ LAM (associated with the plasma membrane) (Gilleron et al. 2000; Pitarque et al. 2008).

### 4.3 Cell Envelope Lipids

The cell envelopes of rhodococci are rich sources of structurally diverse lipids, some of which have pronounced surfactant properties that facilitate the growth of the bacteria on hydrophobic substrates and may be of biotechnological significance (Lang and Philp 1998; Kuyukina, this volume; Sutcliffe 1998). These lipids are typically glycolipids including both acyl- and mycolyl-glycolipids (Table 1). There is also a rich diversity of lipopeptides and glycolipopeptides known to be produced by rhodococci (Table 1). The nature of the associations and the specific functions of these lipids within the rhodococcal cell envelope are largely unknown but it is likely that they can interact/intercalate with the covalently bound mycolic acids (see above). Whether the roles of these lipids are simply structural (i.e. as fillers to complete to the outer lipid permeability barrier) or more dynamic (e.g. in modulating surface physicochemical properties) remain to be determined.
4.4 Capsules and Cell Envelope Polysaccharides

The capsular polysaccharides of *Rhodococcus* spp. have received surprisingly little attention. Seven capsular serotypes of *R. equi* were initially defined by Prescott (1981), and the structures of six of these have been extensively characterised by Richards and co-workers (Richards 1994; Severn and Richards 1999). These polysaccharides are structurally diverse acidic heteropolysaccharides, typically characterised by the presence of acetal-linked pyruvate or lactic acid ether substituents. In many cases, the acidic character in part stems from the presence of glucuronic acid in the polymer repeating unit, although the structure of the serotype 4 capsule is notable for containing a 5-amino-3,5-dideoxynonulosonic (rhodaminic) acid (Richards 1994; Severn and Richards 1999). A recent study showed that inactivation of a gene encoding a putative mycolic acid transferase (*fbpA*) resulted in a failure to correctly encapsulate *R. equi* strain 103 (Syror et al. 2008), possibly due to a failure to correctly incorporate capsule polymer into the cell envelope. Intriguingly, although the capsule has long been considered a potential virulence factor of *R. equi*, it was found that the *fbpA* mutant strain was not attenuated in macrophage or mouse infection models (Syror et al. 2008).

A cell envelope polysaccharide of *R. jostii* has recently been characterised as having a tetrasaccharide repeating unit containing D-glucuronic acid, D-glucose, 2-acetyl-D-galactose and L-fucose (Perry et al. 2007). Close association of the polysaccharide with the cell envelope was suggested by the need to use hot (60°C) 50% aqueous phenol to extract significant yields of the polymer, an extraction method similar to that used for lipoglycans (Garton et al. 2002; Gilleron et al.

---

**Table 1** Representative cell envelope lipids of rhodococci

<table>
<thead>
<tr>
<th>Lipid</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycerol monomycolates</td>
<td><em>R. erythropolis</em></td>
<td>Ioneda and Ono (1996)</td>
</tr>
<tr>
<td>Glycosyl monomycolates</td>
<td><em>R. erythropolis</em></td>
<td>Kurane et al. (1995)</td>
</tr>
<tr>
<td></td>
<td><em>R. rhodochrous</em></td>
<td>de Almeida and Ioneda (1989)</td>
</tr>
<tr>
<td></td>
<td><em>R. ruber</em></td>
<td>Matsunaga et al. (1996)</td>
</tr>
<tr>
<td>Trehalose mycolates</td>
<td><em>R. corynebacteroides</em></td>
<td>Powalla et al. (1989)</td>
</tr>
<tr>
<td></td>
<td><em>R. erythropolis</em></td>
<td>Kretschmer et al. (1982); Kurane et al. (1995)</td>
</tr>
<tr>
<td></td>
<td><em>R. opacus</em></td>
<td>Niescher et al. (2006)</td>
</tr>
<tr>
<td></td>
<td><em>R. rhodochrous</em></td>
<td>Asselineau and Asselineau (1978); de Almeida and Ioneda (1989)</td>
</tr>
<tr>
<td>Acylated carotenoid glucosides</td>
<td><em>R. ruber</em></td>
<td>Matsunaga et al. (1996)</td>
</tr>
<tr>
<td>Acyl pentaglucoside</td>
<td><em>R. rhodochrous</em></td>
<td>Takaichi et al. (1997)</td>
</tr>
<tr>
<td>Succinylated acyl trehaloses</td>
<td><em>R. corynebacteroides</em></td>
<td>Powalla et al. (1989)</td>
</tr>
<tr>
<td></td>
<td><em>R. erythropolis</em></td>
<td>Uchida et al. (1989)</td>
</tr>
<tr>
<td></td>
<td><em>Rhodococcus sp.</em></td>
<td>Tokumoto et al. (2009)</td>
</tr>
<tr>
<td></td>
<td><em>R. wratislaviensis</em></td>
<td>Tuleva et al. (2008)</td>
</tr>
<tr>
<td>Peptidolipids (lipopeptides), mycolylpeptidolipids and peptidoglycolipids</td>
<td><em>R. erythropolis</em></td>
<td>Koronelli (1988)</td>
</tr>
<tr>
<td></td>
<td><em>Rhodococcus sp.</em></td>
<td>Chiba et al. (1999); Peng et al. (2008)</td>
</tr>
</tbody>
</table>
Similarly, an extracellular polysaccharide has also been isolated and characterised from *R. rhodochrous* as having a tetrasaccharide repeating unit containing D-glucuronic acid, D-glucose, D-galactose and D-mannose (Urai et al. 2006b). Small quantities of C16 and C18 fatty acids were found to be esterified to this polymer, suggesting that these might represent a mechanism for anchoring the polysaccharide to the cell envelope. A third polymer, named mucoidan, was identified in *R. erythropolis* PR4 and characterised as having a pentasaccharide repeat unit containing D-glucuronic acid, two D-glucose, N-acetylglucosamine and L-fucose (Urai et al. 2007b). The same strain also produces another polysaccharide, named PR4 FACEPS (fatty acid-containing extracellular polysaccharide), with a tetrasaccharide repeating unit containing D-glucuronic acid, D-glucose, D-galactose and pyruvylated D-mannose, which is esterified with small quantities of fatty acids, as in the *R. rhodochrous* polysaccharide (Urai et al. 2007a). The polysaccharide component of PR4 FACEPS is notably identical to the previously described extracellular polysaccharide of *Rhodococcus* sp. 33 (Urai et al. 2006a).

The above rhodococcal cell envelope polysaccharides share some structural features in common with the capsular polysaccharides of *R. equi*. Interestingly, the structural motif of an acetal-linked pyruvic acid (1-carboxyethylidene) substituent, which is present in the *R. equi* serotype 1, 2 and 7 capsules (Richards 1994), was also identified in the polysaccharide from *Rhodococcus* sp. 33 and PR4 FACEPS from *R. erythropolis* (Urai et al. 2006a, 2007a). To date, 27 antigenically distinct capsular types have been defined by serotyping in *R. equi* alone (Nakazawa et al. 1983), so it is likely that the structural diversity of cell envelope and capsular polysaccharides produced by rhodococci is high. Therefore, this remains an interesting area for future study, particularly as these surface polymers may facilitate the ability of the bacteria to utilise hydrophobic substrates; (Urai et al. 2006b; Perry et al. 2007).

### 4.5 Lipoproteins

Bacteria are capable of covalently modifying proteins by attachment of a lipid group to a cysteine residue, which becomes the N-terminus of the mature protein, that is, synthesising lipoproteins (Hutchings et al. 2009). This provides an important mechanism for localising proteins to bacterial cell membranes. Bacterial lipoproteins are readily identifiable by bioinformatic methods, and analyses of sequenced bacterial genomes have revealed that putative lipoproteins typically represent ca. 2% of the predicted proteomes of Gram-positive bacteria (Sutcliffe and Harrington 2004; Babu et al. 2006; Rahman et al. 2008). As such, bacterial lipoproteins are a functionally diverse and numerically significant class of cell envelope proteins in Actinobacteria. In Gram-positive bacteria, lipoproteins are de facto anchored to the outer leaflet of the plasma membrane, and this is likely to be the major destination of lipoproteins in mycolic-acid containing Actinobacteria, although it remains possible that some lipoproteins are also associated with the mycolate layer (Fig. 3; Sutcliffe and Harrington 2004).
As in *M. tuberculosis* (Sutcliffe and Harrington 2004), bioinformatic analyses of the *R. jostii* genome indicates that ca. 2.0% (>100 proteins) of the predicted proteome are putative lipoproteins (our unpublished observations). As in other Gram-positive bacteria, substrate binding proteins of ABC transport systems for diverse substrates are well represented. Related to this is the recent demonstration that the mce4 operon of *R. jostii* constitutes a complex ABC transport system variant for cholesterol uptake (Mohn et al. 2008) and it is notable that the Mce4E proteins of both *R. jostii* (RHA1_ro04702) and *M. tuberculosis* are predicted lipoproteins (unpublished observations; Sutcliffe and Harrington 2004). Indeed, the mceE proteins associated with each of the multiple mce loci of *M. tuberculosis*, *Nocardia farcinica* and *R. jostii* are putative lipoproteins (unpublished observations; Sutcliffe and Harrington 2004). Collectively, mce operons may encode putative ABC-related transport systems for various (probably hydrophobic) substrates (Casali and Riley 2007). The putative lipoproteins may thus interact with the other membrane-associated/secreted components to form a cell envelope complex involved in substrate scavenging and delivery to the membrane permease.

Consistent with their location in proximity to both the cell membrane and wall, various putative lipoprotein enzymes including cell wall active enzymes can be distinguished. As noted in other Gram-positive bacteria (Hutchings et al. 2009), several putative lipoproteins predicted to be involved in membrane-associated redox processes can be identified (e.g. *R. jostii* RHA1_ro02035, a ResA homologue likely to be involved in cytochrome c biogenesis, and RHA1_ro01137, the cytochrome c oxidase subunit II CtaC). Moreover, at least two putative lipoproteins (RHA1_ro06090 and RHA1_ro06326) appear to be involved in ‘three component systems’ involved in cell envelope sensing and signalling processes (Hoskisson and Hutchings 2006; Ortiz de Orué Lucana and Groves 2009). Finally, as in other bacteria, significant numbers of conserved hypothetical proteins of unknown function were identified as putative lipoproteins.

In addition to these canonical lipoproteins, the immunodominant VapA virulence factor of *R. equi* (Jain et al. 2003) has also been reported to be an acylated protein (Tan et al. 1995), which may explain its association with the rhodococcal cell surface (Sutcliffe 1997). The VapA protein lacks any cysteine and thus cannot be a conventional lipoprotein of the above described type, as a cysteine containing signal peptide is central to the lipid modification pathway (Hutchings et al. 2009). Whether VapA is a unique post-translationally acylated protein or represents the prototype of a novel family of lipid-modified proteins (e.g. in other mycolata) remains an important question for future study. It is notable that the other members of the VapA family are not thought to be lipid-modified but to be surface-associated or secreted proteins (Byrne et al. 2001; Meijer and Prescott 2004).

5 Biosynthesis of Key Cell Envelope Components

As described in Sects. 2 and 4, the mycolate cell envelope is dominated by several distinctive covalently and non-covalently associated components. The biosynthesis and coordinated assembly of these components is reviewed in the following sections.
5.1 Mycolic Acid Biosynthesis

Few studies have directly addressed the production of rhodococcal mycolic acids but the fundamental processes involved in biosynthesis have been extensively investigated in mycobacteria. The synthesis of the component parts of the mycolic acids is, in the main, a straightforward fatty acid biosynthesis, which occurs via the repetition of a cycle of four reactions, where each cycle accomplishes an extension of the alkyl chain by a two-carbon unit.

Two types of fatty acid synthase (FAS) are known. The mammalian-like type FAS-I system is a homo-dimer containing all the necessary functions to achieve de novo fatty acid synthesis (Smith et al. 2003). In contrast, most bacteria utilise a FAS-II system wherein the growing fatty acyl chain is transferred between the active sites of dissociable component enzymes as an acyl thioester of a highly acidic acyl carrier protein (ACP). As with most other mycolata, *R. jostii* is unusual in that both FAS systems are present. FAS-I (*fas*) has been identified as RHA1_ro01426 showing 1965/3100 (63%) amino acid sequence identity to its counterpart in *M. tuberculosis* H37Rv. This FAS-I will be responsible for the de novo fatty acid synthesis, producing fatty acids of C14–C24 carbon chain length. For meromycolic acid biosynthesis, the further extension of the fatty acids produced by FAS-I is performed by a dissociable FAS-II system (Kremer et al. 2001b; Takayama et al. 2005). Like *M. tuberculosis* AcpM, the *R. jostii* AcpM (RHA1_ro01200) that serves FAS-II contains a C-terminal extension relative to other bacterial ACPs (data not shown). The significance of this C-terminal extension is still unknown but a sequence alignment of AcpM from representative mycolata reveals a correlation between the larger mycolates and the size of this extended region (data not shown). Thus it could be speculated that the length of this extension plays a role in the ability of the bacterium to produce longer meromycolates.

The key enzyme that links FAS-I and FAS-II, the β-ketoacyl-ACP synthase III FabH, can be identified as RHA1_ro05206 showing 56% amino acid sequence identity to *M. tuberculosis* H37Rv mtFabH (Choi et al. 2000; Brown et al. 2005). FabH elongates the acyl-CoA primers derived from FAS-I by condensing these with a malonyl-thioester of AcpM to form a β-keto-acyl-AcpM thioester product (Choi et al. 2000; Brown et al. 2005). The malonyl-AcpM substrate is produced by the acyl-CoA/ACP transacylase FabD (RHA1_ro01199), which is encoded within a syntenic gene cluster in *M. tuberculosis* H37Rv (Fig. 4; Kremer et al. 2001b). The β-keto-acyl-AcpM product of FabH is reduced by the β-keto-acyl-reductase, FabG (MabA, RHA1_ro07213) (Banerjee et al. 1998) and its β-hydroxy-acyl-AcpM product is dehydrated by a FabZ-type protein complex. Recently, Rv0635–Rv0637 (FabZ', FabZ, FabZ'', respectively) in *M. tuberculosis* H37Rv were identified as the three component subunits required to perform the dehydration reaction in this species (Brown et al. 2007; Sacco et al. 2007). Interestingly, *R. jostii* carries only homologues of FabZ' (RHA1_ro01983) and FabZ (RHA1_ro01984). The core unit of the dehydratase complex, FabZ, associates with the chain length specific subunits FabZ' and FabZ'' and therefore the absence of FabZ'', which is associated with the later stages of
meromycolate chain extension, is consistent with the shorter mycolate chain lengths observed in *R. jostii* RHA1. The trans-2-enoyl-AcpM product of the FabZZ complex then participates in the final step of the FAS-II reaction cycle, catalysed by the enoyl-ACP reductase FabI (InhA, RHA_ro07214), which is encoded adjacent to fabG, in mycobacteria (Kikuchi and Kusaka 1984; Banerjee et al. 1994). Completing the cycle thus produces an aliphatic acyl-ACP two carbons longer than its acyl primer (Banerjee et al. 1994).

In *M. tuberculosis*, the subsequent rounds of acyl extension by FAS-II are thought to be initiated by the highly similar β-keto-acyl-AcpM synthases, KasA and KasB (Kremer et al. 2000, 2002a; Schaeffer et al. 2001). These enzymes extend acyl-AcpM thioesters, rather than acyl-CoAs, by condensing them with malonyl-AcpM. Both enzymes require acyl-AcpM primers of at least 16 carbons, consistent with a role of FAS-II in extending FAS-I products towards the biosynthesis of long chain fatty acids (Kremer et al. 2002a). KasA, which is responsible for the extension intermediate chain length meromycolate precursors (Kremer et al. 2000), is present in *R. jostii* (RHA1_ro01201, 67% amino acid sequence identity to *M. tuberculosis* KasA). Bhatt et al. (2007) confirmed that KasB functions predominantly in the extension of long-chain length meromycolate precursors. A ΔkasB null mutant in *M. tuberculosis* synthesised shorter mycolic acids compared to the parent strain. Significantly, the only gene missing from the *R. jostii* kasA gene cluster compared to that observed in all mycobacteria (Fig. 4) is a KasB homologue, further supporting the hypothesis that the production of intermediate chain length meromycolates in rhodococci is due to the absence of the requisite machinery to perform further elongation cycles.

Introduction of C=C double bonds into fatty acids and mycolic acids requires fatty acid desaturases. Two putative long-chain fatty acyl ACP desaturases are encoded in the genome of *M. tuberculosis* H37Rv, Rv0824c (DesA1) and Rv1094.
(DesA2), respectively (Cole et al. 1998). DesA1 was originally detected as an exported component of an M. tuberculosis PhoA fusion library processed in M. smegmatis (Lim et al. 1995). The protein contains two copies of the characteristic (D/E)ENXH motif (Jackson et al. 1997) of the class II diiron-oxo proteins to which acyl-ACP desaturases belong (Fox et al. 1994). Two homologues of DesA1 are present in R. jostii, RHA1_ro02258 and RHA1_ro04869. Both exhibit 58% amino acid sequence identity to DesA1, and the gene for the latter is situated in a conserved locus comparable with that of M. tuberculosis DesA1. M. tuberculosis DesA2 does not contain the (D/E)ENXH motifs observed in other acyl-ACP desaturases, although it does possess an EEHXH motif as well as showing a high degree of homology throughout with stearoyl-ACP desaturases. RHA1_ro05863 appears to be an orthologue of DesA2 (37% amino acid sequence identity) in R. jostii and like the M. tuberculosis DesA2 retains only an EENXH motif. Neither of the M. tuberculosis gene products have yet been characterised in terms of desaturase activity and the significance of DesA1 secretion remains unknown.

A membrane-associated fatty acyl-CoA desaturase gene is encoded in the genome of M. tuberculosis H37Rv (Rv3229c, DesA3; Phetsuksiri et al. 2003). Phetsuksiri et al. (2003) demonstrated that DesA3 was involved in the production of olate from stearoyl-CoA and therefore it was designated as a Δ9-desaturase. R. jostii has six other DesA homologues, five of which (RHA1_ro06336, RHA1_ro03422, RHA1_ro01720, RHA1_ro6335 and RHA1_ro3346) show greater than 55% amino acid sequence homology to M. tuberculosis DesA3 and may thus play roles in fatty acid and/or mycolic acid desaturation. The sixth DesA3 homologue, RHA1_ro04464, is noted to contain a significant N-terminal deletion and so may be inactive.

The presence of complex mycolates in mycobacteria can be attributed to the numerous methyltransferases that are involved in functional group formation at proximal and distal modifications sites initially occupied by an unsaturated bond (Dover et al. 2004; Takayama et al. 2005). The absence of modifications in the relatively short mycolic acids of C. diphtheriae has been attributed to the absence of similar modification enzymes as well as to the absence of any fatty acyl desaturase DesA homologues that would provide the requisite unsaturation for further modification by the methyltransferases (Dover et al. 2004). Rhodococcal mycolates are intermediate in terms of both length and complexity compared to mycobacterial and corynebacterial mycolates (Fig. 1), containing up to four double bonds (Alshamony et al. 1976; Barton et al. 1989; Stratton et al. 1999; Nishiuchi et al. 2000) in the distal part of the meromycolate. It is tempting to speculate that the multiple DesA homologues identified above may be involved in the formation of multiply unsaturated mycolates. As in C. diphtheriae, the absence of methoxyl mycolic acid synthases and cyclopropane mycolic acid synthases from the genome of R. jostii RHA1 is consistent with the simpler mycolate profiles of rhodococci.

The penultimate step in the synthesis of mycolic acids involves the Claisen-type condensation of an acyl-S-CoA (that contributes the alkyl branch) with a meromycolyl-AMP (Takayama et al. 2005; Gokhale et al. 2007). Recently a polyketide synthase (Pks13, Rv3800 in M. tuberculosis) has been implicated in this process
Gene disruption experiments of Cg-pks in C. glutamicum generated a viable mutant devoid of corynomycolates (Gande et al. 2004). In R. jostii, RHA1_ro04065 exhibits 1005/1751 (57%) amino acid sequence identity to Rv3800 in M. tuberculosis. The region encompassing pks13 is highly conserved throughout the mycolata (see Sect. 5; Vissa and Brennan 2001; Dover et al. 2004), including R. jostii, due to the essential functions these gene products perform in cell wall biosynthesis (Fig. 5). The pks13 locus also appears conserved in R. rhodochrous (Portevin et al. 2004). As in M. tuberculosis, the specific fatty acyl-AMP ligase (FadD32) responsible for the conversion of the meromycolyol-S-AcpM derived from the FAS-II system to meromycolyl-AMP (Trivedi et al. 2004) is present in R. jostii adjacent to the pks13 gene (RHA1_ro04064). In mycobacteria, the precursor of the 2-alkyl branch is carboxylated by an acyl-CoA carboxylase composed of AccD4 and AccD5, in complex with an ε-subunit and AccBC, to yield 2-carboxyl-acyl-CoA (Gande et al. 2007). Bioinformatic searches have revealed R. jostii RHA1 possesses all the genes required for this function; accD5 is situated alongside the ε-subunit (RHA1_ro06292 and RHA1_ro06291, respectively). However, it is unclear which of the two possible homologues of AccBC (RHA1_ro06282 and RHA1_ro03742) is most likely to be involved, although the proximity of RHA1_ro06282 to

![Fig. 5 Comparative genomic alignment of the pks13-arabinogalactan loci. The M. tuberculosis protein sequences within the pks-arabinogalactan locus were BlastP searched against the R. jostii proteome. The R. jostii protein sequences identified were then reciprocally BlastP searched against the M. tuberculosis proteome to confirm the correct selection of the protein. The figure demonstrates the genomic region of the pks-arabinogalactan locus where the grey bars represent the proteins of significant homology](image-url)
RHA1_ro06291 - RHA1_ro06292 is noted. RHA1_ro04066 represents the likely AccD4 candidate showing 67% amino acid sequence identity to the *M. tuberculosis* protein. The meromycolyl-S-AMP and 2-carboxyl-acyl-CoA are transferred to the \(\beta\)-keto-acyl synthase domain of Pks13 for condensation of the two fatty acyl groups (Gande et al. 2004). In most cases, the \(\beta\)-keto-mycolate product of Pks13 is reduced to a (\(\beta\)-hydroxy)mycolate before export for integration into the cell envelope. Recently *M. tuberculosis* Rv2509 was implicated in the catalysis of this final step of mycolate synthesis (Lea-Smith et al. 2007; Bhatt et al. 2008). In *R. jostii*, RHA1_ro01416 exhibits 69% amino acid sequence identity to *M. tuberculosis* Rv2509 and so is the most likely candidate to perform this function.

### 5.2 Arabinogalactan Biosynthesis

The organisation of AG and its interactions with other wall components, such as peptidoglycan, are likely to prove crucial to the formation of a functional outer lipid permeability barrier in the mycolata by defining the relative spacing of the tethered mycolates on which it is based. Accordingly, the biosynthesis of AG appears to be highly conserved across the taxon, although some diversity, most notably in patterns of arabinan branching and the glycosyl linkages of the galactan domain, have been recorded (Daffe` et al. 1993; Eggeling et al. 2008). Much of our current understanding of the route to its production is derived from the study of various mycobacteria and, more recently, *C. glutamicum* (Eggeling et al. 2008). The dominant driving force behind this research has been the need to define the mechanisms of action of cell wall inhibitors used in current tuberculosis therapies and, following the emergence of extensively drug resistant *M. tuberculosis*, the need to define new targets in the biosynthesis of the *M. tuberculosis* wall (Dover et al. 2008b).

The first insight into AG biosynthesis was derived from the observation of a series of glycolipids elaborated by plasma membrane fractions of *M. smegmatis* and *M. tuberculosis*. Both preparations catalysed the incorporation of radioactivity from UDP-[\(^{14}\)C]-N-acetylglucosamine (GlcNAc) into two polyprenyl phosphate (Pol-P)-based glycolipids (GL1 and GL2). The initial step was identified as the formation of GL1, a Pol-P-P-GlcNAc unit (Mikušová et al. 1996). Incorporation of [\(^{14}\)C] Rhamnose (Rha) from dTDP-[\(^{14}\)C]Rha into GL2 exclusively identified it as Pol-P-P-GlcNAc-Rha (Mikušová et al. 1996). Addition of a cell wall enzyme preparation resulted in the formation of the increasingly polar glycolipids, GL3 and GL4. The inclusion of UDP-[\(^{14}\)C]Galactose (Gal) resulted in exclusive labelling of GL3 and GL4 indicating the initiation of a galactan chain on the GL2 primer (Mikušová et al. 1996). Subsequent analysis of the polymerised product resulting from these labelling experiments pointed to the formation of longer chain intermediates, eventually resulting in a polymer containing 35–50 residues (Besra and Brennan 1997; Mikušová et al. 2000). Glycosidic linkage analysis revealed that the bulk of the galactan polymer consisted of alternating 5- and 6-linked linear galactofuran residues, with a small amount of branching.
The incorporation of radiolabel from synthetic Pol-P-[\(^{14}\)C]-Arabinofuranose (Araf) (Lee et al. 1995) into this same polymer (Mikušová et al. 2000) suggested that the total synthesis of the AG arabinan domain might occur while it is linked to the Pol-P carrier. This switch to Pol-P-derived sugar donor substrates is likely indicative of a shift in the membrane topology of AG biogenesis. The exclusive use of sugar nucleotides until the incorporation of Araf into the polymer suggests that biosynthesis of galactan occurs at the cytosolic face of the plasma membrane. In contrast, the use of Pol-P-based Araf donors suggests that arabinan deposition occurs at the ‘periplasmic’ face of the membrane and implies that the Pol-P-P-LU-galactan is translocated across the plasma membrane before further modification.

### 5.2.1 Linker Unit Synthesis

The LU disaccharide is formed via the addition of first GlcNAc (to form GL1) and then Rha (to form GL2) at the cytoplasmic face of the plasma membrane. The first glycosyltransferase is often purported to be a homologue of *E. coli* Rfe (WecA) (Meier-Dieter et al. 1992) in *M. tuberculosis* (Rfe, Rv1302) though this designation remains presumptive. The RHA1_ro01480 and RHA1_ro01091 proteins are clearly members of the glycosyltransferase family 4 typified by the UDP-GlcNAc/MurNAc:polyprenol-P GlcNAc/MurNAc-1-P transferases (Pfam PF00953, http://pfam.sanger.ac.uk/family?acc=PF00953) (Lehrman 1994). The former displays 67% amino acid identity with *M. tuberculosis* Rfe, and as RHA1_ro01480 is located within a highly syntenic locus in *R. jostii*, it is likely to represent an orthologue. RHA1_ro01091 forms part of an operon that is devoted to the production of peptidoglycan precursors and is clearly the phospho-\(\text{N}^{-}\)-acetylmuramoyl-pentapeptide-transferase (MraY).

The complementation of a *wbbL* mutant of *E. coli*, which is deficient in Rha transfer for lipopolysaccharide biosynthesis, with Rv3265c (*wbbL1*) of *M. tuberculosis* implicates its product as the probable rhamnosyltransferase involved in GL2 synthesis (McNeil 1999). A WbbL1 homologue (63% amino acid identity) is encoded by RHA1_ro06306. Confidence regarding its designation as a rhamnosyltransferase and thus its orthology with *M. tuberculosis* WbbL1 is derived from analysis of its genetic context. The *M. tuberculosis* enzymes providing the dTDP-Rha donor substrate have all been identified and expressed in *E. coli* (Ma et al. 1997, 2001). RmlA to RmlD have been characterised as an \(\alpha\)-D-glucose-1-phosphate thymidylyltransferase, dTDP-\(\alpha\)-D-glucose 4,6-dehydratase, dTDP-4-keto-6-deoxy-\(\alpha\)-D-glucose 3,5 epimerase and dTDP-Rha synthase, respectively (Ma et al. 2001). Homologues of RmlA (RHA1_ro04097, 73% amino acid identity), RmlB (RHA1_ro04098, 70% amino acid identity), RmlC (RHA1_ro04096, 56% amino acid identity) and RmlD (RHA1_ro06305, 54% amino acyl identity) are apparent within the *R. jostii* genome. In both genomes, *rmlD* and *wbbL1* potentially form an operon supporting their coordinated function and ultimately a role in rhamnosyltransfer to the LU precursor.
5.2.2 Galactan Synthesis

The galactose (Gal) residues of AG occur in the relatively uncommon furanose (f) form (McNeil et al. 1987). The requisite UDP-Gal\textsuperscript{f} nucleotide sugar donor in \textit{M. tuberculosis} is provided via two sequential reactions from UDP-Glucose (p, pyranose form; UDP-Glc\textsubscript{p}). The first is catalysed by UDP-Glc\textsubscript{p} epimerase to form UDP-Gal\textsubscript{p}, which is then converted to UDP-Gal\textsubscript{f} by UDP-Gal\textsubscript{p} mutase. Weston et al. (1997) purified a protein with UDP-glucose 4-epimerase activity from \textit{M. smegmatis}. N-terminal sequence analysis suggested that the protein was related to the product of \textit{M. tuberculosis} Rv3634. A similar strategy was used to identify \textit{M. tuberculosis} Rv3809c (Glf) as an orthologue of \textit{M. smegmatis} UDP-Gal\textsubscript{p} mutase; the designation was confirmed by molecular cloning and analysis of crude extracts containing the recombinant protein (Weston et al. 1997).

Two galactosyltransferases involved in \textit{M. tuberculosis} galactan synthesis have now been identified. The product of Rv3808c appeared to be a good candidate in that it occupied the locus adjacent to glf. This putative transferase also contained the signature QXXRW motif, which is found only in processive enzymes, i.e. those which carry out multiple sugar transfers (Saxena et al. 1995). Over-expression of Rv3808c in \textit{M. smegmatis} caused an increased yield of a galactofuran polymer in the over-producing strain (Mikušová et al. 2000). Analyses of the incorporation of Gal\textsubscript{f} into artificial Gal\textsubscript{f} disaccharides by membranes of recombinant \textit{E. coli} expressing Rv3808c demonstrated that, consistent with the alternating β(1→5) and β(1→6) linkages of the native galactan, the incoming sugar adopted a (1→6) linkage when using a (1→5) linked disaccharide acceptor and vice versa (Kremer et al. 2001a). Furthermore, larger oligosaccharide products were also formed in these assays confirming that the product of Rv3808c, now designated GlfT2, is a processive enzyme and, consequently, is likely to produce the bulk of the galactan deposited in the \textit{M. tuberculosis} cell wall (Kremer et al. 2001a). The importance of galactan synthesis to mycobacteria was demonstrated by the disruption of glf in \textit{M. smegmatis}; growth was only supported when functional copies of both glf and glfT2 were provided on complementing plasmids (Pan et al. 2001).

\textit{R. jostii} orthologues of Glf (RHA1\_ro04053, 82% amino acid identity) and GlfT2 (RHA1\_ro04054, 69% amino acid identity) are apparent, and as in \textit{M. tuberculosis}, they are encoded by adjacent genes; glfT2 lies immediately downstream of and overlaps with glf by four nucleotides. Little is known regarding the mechanism by which GlfT2 introduces the distinctive alternating glycosyl linkage pattern that characterises \textit{M. tuberculosis} galactan and thus far the enzyme has proven intractable in structural studies. The genomes of all galactan-producing species sequenced to date contain GlfT2 homologues. A combination of galactan characterisation and a structural genomics survey of GlfT2 homologues or potential alternative galactofuranosyltransferases would provide structural details and illuminate the molecular basis for galactan heterogeneity in the mycolata (Sect. 2.2).

Biophysical analyses of recombinant \textit{M. tuberculosis} GlfT2 confirmed the intuition that, although capable of depositing the bulk of the Gal\textsubscript{f} residues, the enzyme would require a galactosyl primer to extend towards galactan; specifically
GlfT2 bound and donated Galf to both $\beta$5- and $\beta$6-linked Galf–Galf disaccharides but could not donate Galf to an artificial $\beta$-d-Gal-(1→4)-$\alpha$-L-Rha acceptor, which mimics the reducing terminus of galactan (Alderwick et al. 2008). Bioinformatic analyses led to the identification of a second *M. tuberculosis* galactosyltransferase (Rv3782) that participates in the biogenesis of GL4, the Pol-P-LU-Galf$_2$ (Mikušová et al. 2006; Alderwick et al. 2008). *E. coli* extracts containing recombinant Rv3782 (now designated GlfT1) transferred galactosyl residues to artificial acceptors designed to emulate LU and LU-Galf (Alderwick et al. 2008; Belaňová et al. 2008). These combined data suggest that GlfT1 might represent an even more versatile bifunctional protein than GlfT2, able not only to produce both $\beta$-(1→4) and $\beta$-(1→5) linkages but also to utilise diverse acceptor groups, that is a rhamnosyl acceptor in the initial reaction.

The *R. jostii* genome encodes a convincing GlfT1 orthologue in RHA1_ro04113 (68% amino acid identity with *M. tuberculosis* GlfT1) and, as in *M. tuberculosis*, the gene is clustered with two others encoding an apparent polysaccharide exporting ABC transporter (*RHA1_ro04114* and *RHA1_ro04115*). This transport complex represents an attractive candidate to facilitate the export of Pol-P-LU-galactan to the periplasm for arabinosylation.

### 5.2.3 Arabinan Synthesis

The structure of the arabinan portion of *M. tuberculosis* AG is much more complex than that of its galactan partner A series of branches contributes to the formation of the characteristic terminal pentaarabinofuranosyl motif that provides the esterification sites for AG-linked mycolates. Until the recent development of the genetically tractable *C. glutamicum* as a model for AG biosynthesis (Alderwick et al. 2005; Eggeling et al. 2008), much of our insight into arabinan biogenesis emerged from studies related to the mode of action and resistance against the important antitubercular drug EMB (reviewed in Dover et al. 2008a). In vivo pulse-chase labeling experiments in *M. smegmatis* suggested that the Araf residues ultimately deposited in AG derive directly from a Pol-P-Araf sugar donor (Wolucka et al. 1994). EMB, which inhibits biosynthesis of both AG and LAM (Takayama and Kilburn 1989), led to the accumulation of Pol-P-Araf (Wolucka et al. 1994) suggesting the drug caused a lesion in arabinosyltransfer. Application of a synthetic Pol-P-$\left[{^{14}}\text{C}\right]$Araf (Lee et al. 1995) in a cell-free assay system led to deposition of radiolabel in all recognised cell wall arabinan moieties, defining Pol-P-Araf as the major arabinosyl donor in mycobacteria (Xin et al. 1997). However, the possibility of both UDP-Ara (Singh and Hogan 1994) and GDP-Ara (Takayama and Kilburn 1989) in *M. smegmatis*, as well as an undefined soluble Araf donor in *C. glutamicum* (Tatituri et al. 2007), have all been proposed and cannot yet be ruled out as minor cell envelope Araf donors.

Pol-P-Araf appears to arise from 5-phosphoribose pyrophosphate (pRPP) with a 2′ epimerase mediating the ribose→arabinose conversion at an intermediate stage (Scherman et al. 1996). *M. tuberculosis* Rv3806c (UbiA) was identified as the...
pRPP/polypropenyl-phosphate 5-phosphoribosyltransferase and represents the first committed step towards Pol-P-Araf synthesis (Huang et al. 2005). Mikušová et al. (2005) hypothesised that Pol-P-β-d-5-phosphoribose is dephosphorylated to form Pol-P-β-d-ribose before epimerisation of the 2’ hydroxyl group is achieved in a two-stage process. First, oxidation of the hydroxyl probably forms Pol-P-2-keto-β-d-erythro-pentofuranose, which is subsequently reduced to generate Pol-P-β-d-Araf. Two candidate gene products were identified in M. tuberculosis through their similarity to Noe proteins implicated in the arabinosylation of the Azorhizobium caulidans nodulation factor. Rv3790 and Rv3791 were annotated as a putative FAD-dependent oxidoreductase and a probable short-chain dehydrogenase/reductase, respectively, both functions consistent with the reaction schemes hypothesised (Mikušová et al. 2005; Wolucka 2008). Together the purified recombinant proteins were able to catalyse the epimerisation reaction despite neither protein being sufficient to promote the initial oxidation step independently (Mikušová et al. 2005). The enzyme that catalyses the dephosphorylation of Pol-P-β-d-5-phosphoribose that precedes this epimerisation remains unidentified but a candidate is the putative phosphatase encoded by Rv3807c i.e. adjacent to ubiA (Wolucka 2008).

The genome of R. jostii encodes proteins that represent likely orthologues for each of these Pol-P-Araf biosynthetic enzymes (Table 2).

The products of the emb locus of Mycobacterium avium were identified as the targets for EMB. Overexpression of embA and embB from M. avium conferred EMB resistance in M. smegmatis (Belanger et al. 1996). Taken together with the immediate inhibition of [14C]Ara incorporation into both AG and LAM on EMB

### Table 2 Comparison of the enzymology for arabinogalactan biosynthesis in R. jostii and M. tuberculosis

<table>
<thead>
<tr>
<th>Function</th>
<th>M. tuberculosis archetype</th>
<th>RHA1 orthologue</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pol-P arabinose precursor synthesis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pRPP: Pol-P 5-phosphoribosyltransferase</td>
<td>Rv3806c</td>
<td>RHA1_ro04056</td>
<td>71</td>
</tr>
<tr>
<td>pRPP: Pol-P 5-phosphoribosyl phosphatase</td>
<td>Rv3807c</td>
<td>RHA1_ro04055</td>
<td>62</td>
</tr>
<tr>
<td>Pol-P-Ribose 2’ epimerisation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAD-dependent oxidoreductase</td>
<td>Rv3790</td>
<td>RHA1_ro04078</td>
<td>77</td>
</tr>
<tr>
<td>Short chain dehydrogenase</td>
<td>Rv3791</td>
<td>RHA1_ro04077</td>
<td>78</td>
</tr>
<tr>
<td><strong>Arabinosyltransferases (AraT)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α(1→5) AraT</td>
<td>EmbA (Rv3794)</td>
<td>Absent</td>
<td></td>
</tr>
<tr>
<td>α(1→5) AraT</td>
<td>EmbB (Rv3795)</td>
<td>RHA1_ro04068</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>EmbC (Rv3793)</td>
<td>RHA1_ro04069</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>EmbC (Rv3793)</td>
<td>RHA1_ro01774a</td>
<td>47</td>
</tr>
<tr>
<td>Galactan priming α(1→3) AraT</td>
<td>AftA (Rv3792)</td>
<td>RHA1_ro04076</td>
<td>56</td>
</tr>
<tr>
<td>Arabinan branching α(1→3) AraT</td>
<td>AftC (Rv2673)</td>
<td>RHA1_ro06863</td>
<td>54</td>
</tr>
<tr>
<td>Arabinan terminating β(1→2) AraT</td>
<td>AftB (Rv3805c)</td>
<td>RHA1_ro04057</td>
<td>50</td>
</tr>
</tbody>
</table>

*The R. jostii genome contains three clear Emb proteins. This one is located outside of the locus containing the EmbA/EmbB arabinosyl transferases likely to be involved in arabinogalactan biosynthesis (see Sect. 5.2.3)*
treatment of *M. smegmatis* (Takayama and Kilburn 1989) and the accumulation of Pol-P-Araf (Wolucka et al. 1994), a hypothesis that explains this resistance phenotype at the molecular level is that Emb proteins are the arabinosyltransferases contributing to AG biosynthesis. However, their possession of glycosyltransferase activities has yet to be demonstrated through their over-production in a heterologous organism.

As in *M. tuberculosis*, *M. smegmatis* possesses three closely related *emb* genes, clustered *emb*<sub>CAB</sub>, whilst despite possessing only one *emb* gene, *C. glutamicum* produces a similar AG to the mycobacteria (Eggeling et al. 2008). Gene knock out studies in *M. smegmatis* have shed light on the apparent redundancy in its *emb* locus (Escuyer et al. 2001). Individual mutants inactivated in *embC*, *embA* and *embB* were characterised. All three strains were viable but of them, the *embB*<sup>−</sup> mutant was most profoundly affected. Cell wall integrity seemed to be compromised as morphological changes were evident, and the cells displayed increased sensitivity to hydrophobic drugs and detergents. The arabinose content of the AG was diminished for both the *embA*<sup>−</sup> and *embB*<sup>−</sup> strains. Nuclear magnetic resonance studies showed that these mutations resulted in considerable effects upon the formation of the terminal pentaarabinofuranosyl motifs, specifically the addition of the β-d-Araf-(1→2)-β-d-Araf disaccharide to the 3 position of the 3,5-linked Araf residue resulting in a linear terminal motif. However, AG formation in the *embC*<sup>−</sup> strain seemed unaffected whereas arabinan deposition in LAM was abolished. These data support the hypothesis that Emb proteins are intimately involved in the process of cell envelope arabinan deposition and that EmbA and EmbB are crucial to the formation of the pentaarabinofuranosyl motifs of AG that are crucial for the deposition of mycolic acids.

Construction of a knock out mutant in the single *emb* gene of *C. glutamicum* (Alderwick et al. 2005) heralded a period of rapid progress towards the definition of arabinan biosynthesis. The mutant exhibited a slow growing phenotype and was significantly depleted in arabinan. Residual arabinosylation of galactan at the 3′ positions of its 5-linked 8th, 10th and 12th Galf residues by a single Araf residue was detected. This modification was not present in the galactan of a strain disrupted in *ubiA* that lacks Pol-P-Araf (Alderwick et al. 2005). Deletion of the gene immediately upstream of *C. glutamicum emb*, now designated *aftA*, which encodes a member of the GT-C glycosyl transferase superfamily, resulted in an arabinan deficient strain (Alderwick et al. 2006). Clearly, AftA represents a novel arabinosyltransferase that primes arabinan biosynthesis on galactan by addition of a single Araf residue that is presumably elaborated upon by EmbA/B or possibly another Ara transferase. Systematic deletion of other GT-C transferases that might contribute to the biosynthesis of cell envelope polysaccharides in *C. glutamicum* and mycobacteria has recently revealed two further conserved Araf transferases. AftC represents a α-(1→3)-Araf transferase that is essential for the branching of the arabinan towards its reducing end and may also contribute to the formation of the pentaarabinofuranosyl motif (Birch et al. 2008). AftB is another GT-C enzyme that forms the β-d-Araf-(1→2)-β-d-Araf structure that effectively terminates arabinan and also provides one of the sites for mycolylation (Seidel et al. 2007).
Although the *R. jostii* RHA1 AG has not been characterised, it appears that the bacterium possesses orthologues of all of the enzymes implicated in *M. tuberculosis* and *C. glutamicum* arabinan biosynthesis (Table 2). Two homologues of the Emb proteins, as well as the enzymes that initiate and terminate arabinan biosynthesis, AftA and AftB, respectively, are all encoded in a highly conserved cluster of 31 genes first recognised in *M. tuberculosis* by Belanger and Inamine (2000) incorporating Rv3779–Rv3809c and occupying 48.5 kb or ~1% of the chromosome (Fig. 5). Among these genes are *glf* and *glfT* (galactan polymerisation), *embCAB*, *aftA* and *aftB* (arabinan deposition), pks13 and associated enzymes (mycolyl condensation) and *fbpA* (mycolyltransfer). Significantly, the region is well conserved in *M. leprae*, the aetiological agent of leprosy (Vissa and Brennan 2001). This bacterium is an obligate intracellular pathogen and exemplifies an extreme case of reductive evolution as less than half of its genome contains functional genes (Cole et al. 2001). The retention of function over such a large syntenous genomic region in *M. leprae* clearly emphasises the essentiality of the cell wall to the pathogenic mycobacteria. Comparison of the *M. tuberculosis* cell wall locus with the equivalent from the more distantly related bacterium *C. diphtheriae* showed that the overall genetic arrangement remained well conserved but was split into two discontinuous segments resulting in the *emb* homologue of *C. diphtheriae* lying over 460 kb away from the *glfT* homologue (Dover et al. 2004). Likewise in *R. jostii* RHA1, two clusters are apparent, encompassing RHA1_ro04050 to RHA1_ro04079 and RHA1_ro04098 to RHA1_ro04118 (Fig. 5); each shows evidence of rearrangement and carry additional genes relative to *M. tuberculosis*, although it is not clear whether these represent rhodococcal acquisitions or mycobacterial losses or, indeed, whether they contribute to the construction of the rhodococcal cell envelope.

### 5.2.4 Macromolecular Ligation

Thus far, we have considered the independent biosyntheses of AG and the mycolic acids but these components must be brought together in the pseudoperiplasm and covalently combined to form the massive mycolyl–arabinogalactan–peptidoglycan complex. This process will require export of each of the structural components as well the enzymes responsible for mycolyl transfer to the terminal Araf residues of AG. Although their role in galactan export remains to be confirmed, *M. tuberculosis* *rfbDE* and *R. jostii* RHA1_ro04114 and RHA1_ro04115 appear to represent a conserved polysaccharide-exporting ABC transporter (Content and Peirs 2008) and, as both are clustered with a gene encoding the galactan-priming Galf transferase GlfT1(RHA1_ro04113), their coordinated function in galactan biosynthesis and export is likely. Once translocated, the arabinosylation of galactan can commence with AftA. On completion, AG units must be integrated into the growing murein sacculus; little is known regarding the process other than ligation requires simultaneous synthesis of both AG and peptidoglycan (Hancock et al. 2002). The enzymology of AG ligation remains enigmatic.
An interesting Pol-P-based mycolylated glycolipid, 6-O-mycolyl-β-D-mannopyranosyl-1-monophosphoryl-heptaprenol (Myc-PL), was purified from *M. smegmatis* and suggested to be the carrier of newly synthesised mycolic acid during translocation across the plasma membrane (Besra et al. 1994). A similar lipid had also been reported in *C. diphtheriae* (Datta and Takayama 1993) suggesting a conserved means for translocation of mycolates across the membrane for the synthesis of trehalose dimycolates and cell wall mycolates.

While prospecting for genes involved in mycolate biosynthesis and processing, Wang et al. (2006) isolated a slow-growing transposon insertion mutant of *Corynebacterium matruchotii* with an apparent impairment in corynomycolate production. The transposon had inserted within a probable orthologue of *C. diphtheriae* DIP1297, an integral membrane protein encoded by the first of a four gene cluster of which the latter three genes had been annotated as encoding an antibiotic transporter (Brabant et al. 2000). The genes of the equivalent cluster in *C. glutamicum* were apparently cotranscribed on a polycistronic mRNA suggesting coordinated function. Application of comparative genomics techniques demonstrated the conservation of the cluster in *M. tuberculosis* (rv1459c, rv1458c–rv1456c), other mycobacteria, corynebacteria and nocardiae and, by supposition, across the *Corynebacterineae* but not in other Actinobacteria (Wang et al. 2006). A similar cluster also occurs in *R. jostii* (RHAI_ro07191 to RHAI_ro07194). Analysis of mycolic acid chain length in the *C. matruchotii* mutant revealed that shorter chain-length corynomycolates (C₂₄–C₃₂ rather than C₃₄–C₃₆) were under-represented (Wang et al. 2006) leading the authors to suggest that this represented an export complex for short-chain mycolates (Wang et al. 2006). However, such short-chain mycolates are likely to be, at best, infrequent modifications to the cell wall of mycobacteria. As there was effective export of the larger mycolate subpopulation of corynomycolates in the *C. matruchotii* mutant, suggesting some redundancy in corynomycolate translocation, one might expect that other *Corynebacterineae* producing larger mycolates would possess this alternate system.

Another important factor in the processing of mycolic acids is the requirement for glucose or α-D-glucopyranosyl-containing oligosaccharides such as trehalose, which is essential for the growth of the *M. tuberculosis* but not corynebacteria. Despite *M. tuberculosis* possessing three potential routes to trehalose, inactivation of a component of the OtsAB pathway (OtsB2, Rv3372; trehalose-6-phosphate phosphatase) abrogated growth (Murphy et al. 2005). *R. jostii* possesses a single homologue of *M. tuberculosis* OtsB2 (RHA_ro00045, 56% amino acid identity) and two homologues of *M. tuberculosis* OtsA (trehalose-6-phosphate synthase; RHA_ro04708, 77% amino acid identity; RHA_ro04708, 69% amino acid identity).

In the absence of exogenous α-D-glucopyranosyl-containing oligosaccharides, a multiply-mutated *C. glutamicum* strain incapable of trehalose synthesis exhibited altered surface properties and lacked mycolic acids in its envelope. The mycolyl residues synthesised by the mutant grown with suitable oligosaccharides were transferred both onto the cell wall and free sugar acceptors. Furthermore, as the mutant had shown no capacity for trehalose uptake, radioactive labelling experiments with [¹⁴C]trehalose showed that the transfer of mycoloyl residues onto sugars
occurs outside the plasma membrane (Tropis et al. 2005b). Thus trehalose appears
to be an important extracytoplasmic carrier for mycolates, allowing their deposition
in the cell wall.

A mycolyltransferase capable of exchanging mycolyl residues between mycolyl-
trehalose and the free disaccharide was purified from M. smegmatis and a role in
mycolyl deposition suggested (Sathyamoorthy and Takayama 1987). Belisle et al.
subsequently demonstrated that three members of the M. tuberculosis antigen 85
complex, Ag85A, Ag85B and Ag85C2 (encoded by fbpA, fbpB and fbpC2 respec-
tively) were able to catalyse mycolyltransferase reactions (Belisle et al. 1997). In
order to shed light upon this apparent redundancy in mycolyltransferases and to
ascertain the biological roles of the individual enzymes, fbpC2, fbpA and fbpB have
all been disrupted (Jackson et al. 1999; Armitige et al. 2000). The disruption of
fbpC2 in M. tuberculosis decreased transfer of mycolates to the cell wall by 40%
without affecting the profile of mycolate types esterified to AG or occurring as free
glycolipids. Thus FbpC2 is involved, either directly or indirectly, in the transfer of
mycolates onto the cell wall and is probably not specific for a given type of
mycolate, or at least the remaining mycolyltransferases are able to maintain the
balance between the mycolate types through their own broad specificity (Jackson
et al. 1999). Although an fbpA mutant grew as well as the parent strain in laboratory
media and macrophage-like cell lines, the fbpB mutant only grew well in laboratory
media. In macrophage-like cell lines, the strain grew very poorly, if at all (Armitige
et al. 2000; Puech et al. 2002).

Corynebacteria possess genes with significant homology to those encoding the
antigen 85 complex (Joliff et al. 1992). Disruption of csp1 encoding the secreted
Fbp-like protein PS1 of C. glutamicum led to a 50% decrease in the amount of cell
wall-linked corynomycolates and an alteration in the cell wall permeability (Puech
et al. 2000). The expression of fbpA, fbpB and fbpC2 from M. tuberculosis in this
csp1-deficient strain restored the cell wall-linked mycolate content and the outer
permeability barrier of the mutant. The enormous structural differences between
corynomycolates and their mycobacterial counterparts (Fig. 1) suggest that these
enzymes possess a broad specificity (Puech et al. 2002). All three enzymes are able
to transfer mycolates to AG and display no preference for mycolyltransfer to the
terminal or 2-linked Araf residues of the pentaarabinosyl motifs of AG (Puech et al.
2002). Redundancy in mycolyltransferase activity is apparently a common theme
among the mycolata.

R. jostii RHA1 appears to possess 13 (RHA1_ro04059, ro04058, ro04960,
ro04060, ro04126, ro04189, ro05513, ro02206, ro02143, ro05007, ro05217,
ro05431, ro03469) potential mycolyltransferases (BLASTP query, M. tuberculosis
FbpA; cut off, E 10⁻²¹) and, consistent with a periplasmic location, signal peptides
were predicted for all 13 proteins and all retained a conserved triad of active site
residues (Belisle et al. 1997; Kremer et al. 2002c). Similarly, Sydor et al. (2008)
suggested that R. equi might possess up to 13 FbpA homologues. RHA1_ro04060 is
distinct from the mycobacterial mycolyltransferases because of its larger size (640
amino acids, i.e. almost double the size of M. tuberculosis FbpA and the other R. jostii
homologues, which are ca. 330 amino acids) and is likely to represent an orthologue of
PS1/Cop1 of *C. glutamicum* (Joliff et al. 1992; Brand et al. 2003). *RHA1_ro04058–RHA1_ro04060* are situated within one of the large cell wall biosynthetic clusters (Fig. 5) in a position analogous to *fbpAC2*, that is, immediately downstream from AftB which supplies terminal Ara residues to which mycolates are ultimately esterified by mycolyltransferases. Thus *RHA1_ro04058* and *RHA1_ro04059* are almost certainly mycolyltransferases. A thorough biochemical characterisation of this group of rhodococcal proteins will determine whether these and their *R. equi* counterparts represent astonishing redundancy in mycolyltransfer activity or whether some represent a series of paralogous secreted esterases that have significance in the metabolism of the rhodococci. As noted (Sect. 4.4), mutation of a *fbpA* homologue in *R. equi* affected capsule incorporation, but a detailed mycolate profile of mutant compared to wild type was not reported (Sydor et al. 2008).

Like *C. glutamicum* PS1, the N-terminus of *RHA1_ro04060* exhibits significant amino acid identity with *M. tuberculosis* *fbpA* over its full length (Joliff et al. 1992) with the remaining sequence representing a C-terminal extension that carries three LGFP repeats (Pfam08310) (Adindla et al. 2004). The four LGFP repeats of *C. glutamicum* PS1 are hypothesised to anchor the protein to the wall and may be important for maintaining cell wall integrity (Ramulu et al. 2006). Deletion of *C. glutamicum* PS1 results in a tenfold increase in cell volume and implicates the corresponding proteins in cell shape formation (Brand et al. 2003).

### 5.3 LAM Biosynthesis

As with the biosynthesis of other cell envelope polymers, understanding of LAM biosynthesis has been greatly advanced by comparative studies on mycobacteria and corynebacteria. Consistent with the structural elements of the lipoglycans, the biosynthetic pathway can be divided into distinct stages, with initial synthesis of phosphatidylinositolmannosides (PIM) at the cytoplasmic face of the plasma membrane preceding ‘flipping’ of the glycoprophospholipid prior to mannose chain extension and arabinosylation at the outer face of the plasma membrane. As in many other actinomycete genomes, and consistent with the widespread distribution of PIM glycoprophospholipids, an operon containing the phosphatidylinositol synthase, an acyltransferase and PimA mannosyltransferase required for the biosynthesis of acylated phosphatidylinositol monomannoside (PIM₁; Korduláková et al. 2002, 2003) is present in the *R. jostii* RHA1 genome (*RHA1_ro06880–RHA1_ro06882*). The mannose in PIM₁ is added to the inositol C2 position. The second mannose, added to the inositol C6 position in PIM₁, is added by the recently defined PimB’ mannosyltransferase, which generates PIM₂ (Lea-Smith et al. 2008; Mishra et al. 2008b, 2009). Both PimA and PimB are cytoplasmic enzymes that utilise GDP-mannose as the mannose donor. In *R. jostii*, RHA1_ro01122 can be clearly identified as PimB’ by its homology with *C. glutamicum* NCgl2106 and *M. tuberculosis* Rv2188c (Lea-Smith et al. 2008; Mishra et al. 2008b, 2009).
The mono- and dimannosylated ‘lower’ PIMs, notably acylated PIM₂, are readily observed as free lipids in the membranes of rhodococci and other mycolata (Minnikin et al. 1977; Barton et al. 1989). Further to PIM₂ biosynthesis, subsequent mannosyltransferase activities are needed to convert these ‘lower’ PIMs to the ‘higher’ PIMs that are found in mycobacteria (notably phosphatidylinositol hexamannoside, PIM₆). Some relevant mannosyltransferases have been identified but the extent to which there is redundancy in this pathway is not yet clear (Kremer et al. 2002c; Morita et al. 2006; Crellin et al. 2008). Moreover, at an as-yet undefined stage, PIMs are ‘flipped’ from the inner leaflet of the plasma membrane to the outside leaflet such that the final steps of PIM mannosylation are carried out by GT-C family glycosyltransferases, using Pol-P-linked mannose (see below) as the mannose donor (Berg et al. 2007). After PIM translocation, PIM₄ can be either shunted towards lipomannan/LAM biosynthesis by the LpqW lipoprotein (Kovacevic et al. 2006; Marland et al. 2006; Crellin et al. 2008) or, in mycobacteria, mannosylated with α₁→2 linked mannose to generate PIM₆ (Morita et al. 2006; Crellin et al. 2008). The PIM₄ precursor is mannosylated to generate the α₁→6 lipomannan core of LAM by the sequential action of the MptB (Mishra et al. 2008a) and MptA GT-C mannosyltransferases (Kaur et al. 2007; Mishra et al. 2007), each using Pol-P-mannose as mannose donor. Branching α₁→2 mannose units on the mannan core can be introduced by the Rv2181 GT-C mannosyltransferase (Kaur et al. 2008). Further to the generation of the lipomannan core unit, arabinosylation of mycobacterial LAM is carried out by the EmbC arabinosyltransferase (Zhang et al. 2003; Goude et al. 2008). However, as an embC mutant of M. smegmatis still incorporated two to three arabinosyl units per lipomannan (Zhang et al. 2003), it is likely that the initial ‘priming’ arabinose units are added by a separate arabinosyltransferase in a manner analogous to the priming by AftA in arabinogalactan synthesis (Alderwick et al. 2006). Thus a nearly complete pathway for mycobacterial LAM biosynthesis has been defined, with the crucial remaining questions including the nature and substrate(s) of the PIM ‘flippase’ and the mannosyltransferase(s) that convert PIM₂ to PIM₄, and the arabinosyl ‘priming’ activity.

From the above a near complete pathway for the biosynthesis of rhodococcal LAM can be reconstructed from the R. jostii genome. In addition to the above described acyl PIM₁ biosynthetic locus and PimB’, clear homologues of all the key enzymes identified in corynebacteria and/or mycobacteria can be identified (Table 3). In corynebacteria, higher PIMs are apparently not synthesised as free lipids, as indicated by the buildup of PIM₂ in C. glutamicum mutants unable to synthesise LAM or the Pol-P-mannose sugar donor (Gibson et al. 2003a; Mishra et al. 2008a). Thus PIM₂ is most likely flipped and elaborated into LAM (Mishra et al. 2008a). Intriguingly, the R. jostii genome contains a locus (RHA1_ro05934, RHA1_ro05929) comparable to that in mycobacterial genomes, which contains homologues of both the lipoprotein LpqW required to shunt PIM₄ towards LAM biosynthesis and the PimE mannosyl transferase required to synthesise PIM₆ from PIM₄ (Kovacevic et al. 2006; Marland et al. 2006; Morita et al. 2006; Crellin et al. 2008). Thus, R. jostii may be able to synthesise both a LAM-like lipoglycan (as in other rhodococci, see Sect. 4.2) and higher PIMs. PIMs larger than PIM₂ have not
been reported previously in rhodococci but have been reported in some mycolata other than mycobacteria (e.g. Khuller 1977; Furneaux et al. 2005). This may reflect the nature of the solvent systems used for extraction and analysis in early studies (Minnikin et al. 1977) and so a re-evaluation of the distribution of higher PIMs in rhodococci and other mycolate bacteria is warranted.

The extracytoplasmic stages of LAM biosynthesis rely on Pol-P-linked mannose and arabinose sugar donors. The requisite genes for the biosynthesis of each are present in the R. jostii genome (Sect. 5.2.3; Tables 2 and 3). Intriguingly, a ubiA mutant of C. glutamicum can still produce a truncated LAM-variant (Tatituri et al. 2007). In conjunction with the above described apparent residual arabinosylation of LAM in an M. smegmatis embC mutant (Zhang et al. 2003), it is possible to speculate that an alternative arabinose donor may be needed to prime the core lipomannan during LAM biosynthesis and that this might occur during the

<table>
<thead>
<tr>
<th>Step</th>
<th>M. tub.</th>
<th>C. glut.</th>
<th>RHA1</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polyprenyl-P-sugar precursor biosynthesis</td>
<td>Rv3806c</td>
<td>NCgI2781</td>
<td>Ro04056</td>
<td>216/302 (71%)</td>
</tr>
<tr>
<td>5’polyprenyl phosphoribose synthase</td>
<td>Rv3807c</td>
<td>NCgI2782</td>
<td>Ro04055</td>
<td>86/138 (62%)</td>
</tr>
<tr>
<td>Polyprenyl phosphoribose</td>
<td>Rv3790</td>
<td>NCgI0187</td>
<td>Ro04078</td>
<td>362/466 (77%)</td>
</tr>
<tr>
<td>2-epimerase (heterodimer)</td>
<td>Rv3791</td>
<td>NCgI0186</td>
<td>Ro04077</td>
<td>198/253 (78%)</td>
</tr>
<tr>
<td>Ppm1, Polyprenylphosphate mannosyl transferase</td>
<td>Rv2051c</td>
<td>NCgI1423</td>
<td>Ro00145</td>
<td>164/242 (67%)</td>
</tr>
</tbody>
</table>

**PIM biosynthesis**

<table>
<thead>
<tr>
<th>Step</th>
<th>M. tub.</th>
<th>C. glut.</th>
<th>RHA1</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>PimA, mannosyl transferase</td>
<td>Rv2610c</td>
<td>NCgI1603</td>
<td>Ro06882</td>
<td>254/367 (69%)</td>
</tr>
<tr>
<td>PIM1 acyltransferase</td>
<td>Rv2611c</td>
<td>NCgI1604</td>
<td>Ro06880</td>
<td>131/209 (62%)</td>
</tr>
<tr>
<td>PgsA, phosphatidylinositol synthase</td>
<td>Rv2612c</td>
<td>NCgI1605</td>
<td>Ro06881</td>
<td>192/300 (64%)</td>
</tr>
<tr>
<td>PimB’, mannosyl transferase</td>
<td>Rv2188c</td>
<td>NCgI2106</td>
<td>Ro01122</td>
<td>213/287 (74%)</td>
</tr>
<tr>
<td>PimC</td>
<td>MT1800</td>
<td>Absent</td>
<td>Ro04052</td>
<td>219/365 (60%)</td>
</tr>
</tbody>
</table>

**PIM extension to lipomannan**

<table>
<thead>
<tr>
<th>Step</th>
<th>M. tub.</th>
<th>C. glut.</th>
<th>RHA1</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>LpqW, lipoprotein delivering PIM to MptB</td>
<td>Rv1166</td>
<td>Absent</td>
<td>Ro05934</td>
<td>304/620 (49%)</td>
</tr>
<tr>
<td>MptB, mannosyl transferase</td>
<td>Rv1459c</td>
<td>NCgI1505</td>
<td>Ro07194</td>
<td>315/564 (55%)</td>
</tr>
<tr>
<td>MptA, mannosyl transferase</td>
<td>Rv2174</td>
<td>NCgI2093</td>
<td>Ro01108</td>
<td>255/460 (55%)</td>
</tr>
<tr>
<td>Branching mannosyl transferase</td>
<td>Rv2181</td>
<td>NCgI2100</td>
<td>Ro01114</td>
<td>175/394 (44%)</td>
</tr>
</tbody>
</table>

**Arabinosyltransferases**

<table>
<thead>
<tr>
<th>Step</th>
<th>M. tub.</th>
<th>C. glut.</th>
<th>RHA1</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priming arabinosyl transferase</td>
<td>Unidentified</td>
<td>Unidentified</td>
<td>Unidentified</td>
<td>508/1091 (46%)</td>
</tr>
<tr>
<td>EmbC</td>
<td>Rv3793</td>
<td>NCgI0184</td>
<td>Ro01774c</td>
<td>508/1091 (46%)</td>
</tr>
<tr>
<td>Capping mannosyltransferase</td>
<td>Rv1635c</td>
<td>Absent</td>
<td>Ro04110</td>
<td>200/507 (39%)</td>
</tr>
</tbody>
</table>

*PimC is a redundant mannosyltransferase capable of synthesising PIM1. However, this protein is absent from the genome of M. tuberculosis H37Rv (Kremer et al. 2002b)

bNo clear orthologue (cut off, E 10^{-35}) identified

cThe R. jostii genome contains three clear homologues of EmbC. This one is located outside of the locus containing the EmbA/EmbB arabinosyl transferases likely to be involved in arabinogalactan biosynthesis (see Sect. 5.2.3)
cytoplasmic stages of biosynthesis (Tatituri et al. 2007). This priming activity alone might therefore be sufficient to generate the truncated LAM types observed in R. equi and R. ruber (Sect. 4.2). In R. jostii, the presence of a third Emb family protein (RHA1_ro01774; Table 3) might be related to LAM or AG biosynthesis, although its gene is located outside of the cell wall biosynthetic loci (Fig. 5). Further investigation of both the LAM structure in R. jostii and the functional redundancy of the Emb proteins is needed. In this respect, it is intriguing that the R. jostii genome contains a clear orthologue (Table 3) of the Rv1635c mannosyltransferase that is involved in adding the mannan caps to the arabinan of mycobacterial LAM (Dinadayala et al. 2006; Appelmelk et al. 2008).

A PIM-anchored lipomannan as well as a second lipomannan most likely anchored by a mannosylglucosyluronic acid glycolipid have recently been identified in C. glutamicum (Tatituri et al. 2007; Lea-Smith et al. 2008). Synthesis of the mannosylglucosyluronic acid glycolipid from glucosyluronic acid-diacylglycerol depends on the mannosyltransferase MgtA (NCgl0452). An orthologue of MgtA, ro01995 (64%, 248/382 amino acid sequence identity) is encoded in the R. jostii genome raising the possibility that this species also synthesises novel mannosylglucosuronic acid based glycolipid(s). However, it is notable that extracts of R. equi and R. ruber that contain the truncated LAMs of these species do not contain a separate lipomannan fraction (Garton et al. 2002; Gibson et al. 2003b).

Finally, in addition to providing mannose for lipomannan biosynthesis, Pol-P-linked mannose can also be the sugar donor for protein glycosylation in Actinobacteria (VanderVen et al. 2005; Mahne et al. 2006; Wehmeier et al. 2009). R. jostii RHA1_ro05660 encodes a clear homologue of these protein mannosyltransferases, suggesting some cell envelope or secreted proteins are glycosylated.

6 Concluding Comments

The presence of a mycolic acid containing cell envelope is clearly one of the defining features that influences the biology of members of the genus Rhodococcus. Significant studies have confirmed the presence of all of the components typical of the cell envelopes of the mycolata, notably a peptidoglycan–arabinogalactan–mycolic acid complex, mycolyl glycolipids, channel-forming porins and LAM-like lipoglycans. As reviewed here and previously (Sutcliffe 1997, 1998), understanding of the general principles underlying the organisation of these components can be drawn from both theoretical models and experimental evidence obtained with other mycolata, notably members of the genera Corynebacterium and Mycobacterium. However, it is equally clear that there are likely to be genus, species and strain-specific variations in the fine detail of the organisation of these cell envelopes. Further to these models of cell envelope organisation, a comparative genomics approach should allow a rapid growth in knowledge of the pathways leading to the biosynthesis and assembly of cell envelope components, as illustrated herein by our
analyses of the genome of *R. jostii* RHA1. These developments are likely to herald a productive era in defining both the basic biology and the biotechnological potential of members of this fascinating genus.

References


The Rhodococcal Cell Envelope: Composition, Organisation and Biosynthesis 61


de Almeida ET, Ioneda T (1989) Composition and toxicity of lipids from Rhodococcus rhodochrous grown on medium containing galactose, glucose or mannose. Biochim Biophys Acta 1005:45–50


The Rhodococcal Cell Envelope: Composition, Organisation and Biosynthesis


