**Vítor Gonçalo Silva e Costa Mendes**

# New insights into the biosynthesis of the Mycobacterial Methylglucose Lipopolysaccharide

**Universidade de Coimbra 2011**

**Orientação científica:**

**Prof. Doutor Milton Simões da Costa**
**Doutor Nuno Miguel da Silva Empadinhas**

Dissertação apresentada à Faculdade de Ciências e Tecnologia da Universidade de Coimbra para prestação de provas de Doutoramento em Bioquímica, especialidade de Microbiologia

*Science is a way of thinking much more than it is a body of knowledge*

Carl Sagan

# Agradecimentos

Em primeiro, agradeço ao meu orientador, Professor Milton Simões da Costa, por me ter aceitado no seu laboratório e concedido a oportunidade de prosseguir a minha aprendizagem. O rigor e exigência pelos quais se rege e que me transmitiu ao longo deste trabalho, serão uma referência para a vida.

Nuno, por todo o apoio, orientação científica e encorajamento que me concedeste. Pela constante disponibilidade para discutir ideias, dentro e fora do laboratório. Pela confiança depositada e conhecimentos transmitidos, que me deram autonomia para perseguir as minhas ideias, ao longo destes quatro anos de trabalho, um grande Obrigado!

Agradeço à Fundação para a Ciência e Tecnologia (FCT) o apoio financeiro concedido, através de uma bolsa de doutoramento (SFRH/BD/36373/2007).

A todos os colegas de laboratório, quero agradecer a amizade, disponibilidade, brincadeiras, entreajuda, discussão de ideias, incentivo e espírito de equipa, que tornam este laboratório num local de trabalho especial. Em particular tenho ainda que agradecer à Ana Maranha e à Susana Alarico, pela ajuda valiosa em momentos de muito trabalho e ao Vítor Carmona, pela sua participação voluntariosa em tarefas pouco apetecíveis.

Um agradecimento especial, ao Professor António Veríssimo Pires, não só por me ter permitido iniciar a minha carreira científica neste laboratório, mas também pela forma de leccionar entusiasmante e que em muito contribuiu para despertar o meu gosto pela microbiologia.

E à minha família

João, agradeço-te todo o companheirismo, amizade, apoio e brincadeiras, que me fazem ver-te mais como um irmão do que como um tio.

Aos meus avós, quero agradecer o apoio incondicional e infinito que me deram e ainda por acreditarem continuamente em mim.

Aos meus pais, por toda a disponibilidade, incentivo, conselhos e esforço incansável para verem os meus sonhos realizados.

Pela força que me dás diariamente, pela paciência inesgotável, pelo sorriso que me faz brilhar, pela perspicácia, inteligência e sabedoria que me surpreendem continuamente, pela motivação permanente e finalmente pelo amor e amizade incondicionais… Carla, obrigado por seres a minha fonte de felicidade!

## Thesis Contents

## Abbreviations

| | |
|---|---|
| 3-PGA | 3-phosphoglycerate |
| AG | arabinogalactan |
| AM | arabinomannan |
| BTP | 2-[bis(hydroxyethyl)amino]-2-(hydroxymethyl)-propane-1,3-diol |
| CAPS | N-cyclohexyl-3-aminopropanesulfonic acid |
| dPGM | co-factor dependent phosphoglycerate mutase |
| DSMZ | Deutsche Sammlung von Mikroorganismen und Zellkulturen |
| ESX | ESAT-6 secretion system |
| FAS-I | fatty acid synthase complex I |
| GG | glucosylglycerate |
| GlgA | glycogen synthase |
| GlgB | glycogen branching enzyme |
| GlgC | glucose-1-phosphate adenylyltransferase (ADP-glucose pyrophosphorylase) |
| GlgE | maltosyltransferase |
| GPG | glucosyl-3-phosphoglycerate |
| GpgP | glucosyl-3-phosphoglycerate phosphatase |
| GpgS | glucosyl-3-phosphoglycerate synthase |
| GT81 | glycosyltransferase family 81 |
| HAD | Haloacid dehalogenase |
| IP | isoelectric point |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| LAM | lipoarabinomannan |
| LM | lipomannan |
| MA | mycolic acid |
| Mak | maltokinase |
| MES | 2-(N-morpholino)ethanesulfonic acid |
| MGG | mannosylglucosylglycerate |
| MGLP | methylglucose lipopolysaccharide |

| | |
|---|---|
| MGPG | mannosylglucosyl-3-phosphoglycerate |
| MMP | methylmannose polysaccharide |
| MPG | mannosyl-3-phosphoglycerate |
| MpgS | mannosyl-3-phosphoglycerate synthase |
| MpgP | mannosyl-3-phosphoglycerate phosphatase |
| MS | mass spectrometry |
| MW | molecular weight |
| NMR | nuclear magnetic resonance |
| OM | outer membrane |
| OtsA | trehalose-6-phosphate synthase |
| OtsB | trehalose-6-phosphate phosphatase |
| PAGE | polyacrylamide gel electrophoresis |
| PE | Proline-glutamate |
| PG | peptidoglycan |
| PGM | phosphoglycerate mutase |
| PI | phosphatidyl-myo-inositol |
| PIM | phosphatidyl-myo-inositolmannoside |
| PMPS | polymethylated polysaccharide |
| PPE | Proline-Proline-glutamate |
| RHG | arginine, histidine, glycine |
| SAM | S-adenosyl-methionine |
| SDS | sodium dodecyl sulphate |
| TB | Tuberculosis |
| TLC | thin-layer chromatography |
| TreS | trehalose synthase |
| TreY | maltooligosyl trehalose synthase |
| TreZ | maltooligosyl trehalose trehalohydrolase |
| Tris | tris(hydroxymethyl)aminomethane |
| WHO | World Health Organization |

## Abstract

*Mycobacterium tuberculosis* owes much of its pathogenicity to a thick hydrophobic envelope, rich in complex fatty acids (mycolic acids), that protects the cells from the host's defenses and modulates the immune response.

Mycobacteria are known to synthesize cytoplasmic polymethylated polysaccharides (PMPS) that modulate fatty acid synthesis and act as their carriers to the cell wall assembly. Two PMPS were identified; the methylglucose lipopolysaccharide (MGLP) and the methylmannose polysaccharide (MMP). While MGLP is a complex, branched and acylated structure that contains glyceric acid at the reducing end, MMP is a linear polysaccharide that was never identified in pathogenic mycobacteria.

We have identified two genes responsible for the synthesis of glucosylglycerate (GG), the putative primer for MGLP synthesis. One of the genes (Rv1208), was considered essential to the growth of *M. tuberculosis* and encodes a glucosyl-3-phosphoglycerate synthase (GpgS), belonging to the GT81 family of glycosyltransferases. This enzyme uses 3-phosphoglycerate and NDP-glucose to synthesize glucosyl-3-phosphoglycerate (GPG). The other (Rv2419c) encodes a glucosyl-3-phosphoglycerate phosphatase (GpgP) that promotes the dephosphorylation of GPG to GG. The mycobacterial GpgP, unlike all other known GpgPs that belong to the HAD-like hydrolase superfamily, belongs to the histidine phosphatase superfamily. Both genes are present in all mycobacteria with sequenced genome, including all species known to synthesize the MGLP and in other different actinobacteria. The two corresponding enzymes were produced in recombinant form, purified and characterized. Biochemical data revealed that both the enzymes were highly specific for their substrates.

Recently, maltose-1-phosphate was demonstrated to be the intermediate in a pathway linking trehalose with the synthesis of glycogen and other α-glucans, including the MGLP and, if accumulated, it was also found to be toxic to *M. tuberculosis*. In this study, we also produced, purified and biochemically characterized the mycobacterial

maltose-1-phosphate synthesizing enzyme (maltokinase). Although the *M. tuberculosis* H37Rv *mak* gene (Rv0127) was considered essential for growth, no mycobacterial maltokinase had been characterized, to date. Unlike other maltokinases studied to date, the mycobacterial enzyme showed high specificity towards the maltose acceptor but high flexibility for the phosphate donors, namely ATP, GTP and UTP. Although it was previously demonstrated that the maltokinase is involved in the conversion of trehalose into α-glucans, the proposed essentiality of this enzyme together with the non essentiality of the pathway points to other, yet unexplored roles of maltose-1-phosphate in mycobacterial physiology.

The proposed essentiality of the *gpgS* gene and of other genes putatively involved in MGLP synthesis in *M. tuberculosis*, points to a vital role of this polysaccharide in this organism; however, this remains to be experimentally established. This thesis unravels the identity of two of the genes and enzymes involved in the pathway leading to MGLP synthesis and furthers our understanding on how this important mycobacterial structure is synthesized, paving the way to future studies on the function of MGLP in mycobacterial physiology.

## Resumo

*Mycobacterium tuberculosis* deve grande parte da sua patogenicidade a uma parede espessa e hidrofóbica rica em ácidos gordos complexos (ácidos micólicos) que a protege das defesas do hospedeiro e que altera a resposta imunológica do hospedeiro.

As micobactérias sintetizam polissacarídeos polimetilados citoplasmáticos que regulam a síntese de ácidos gordos e actuam como transportadores destes para a síntese da parede celular. Dois polissacarídeos polimetilados foram identificados nestes organismos: o lipopolissacarídeo de metil glicose (MGLP) e o polissacarídeo de metil manose (MMP). Enquanto o primeiro é uma estrutura complexa, ramificada e acilada que contem ácido glicérico no terminal redutor o segundo é uma estrutura linear que nunca foi identificada em micobacterias patogénicas.

Neste trabalho foram identificados dois genes, responsáveis pela síntese de glicosilglicerato (GG), que é a o iniciador proposto para a via biosintética do MGLP. Um dos genes (Rv1208) foi considerado essencial para o crescimento de *M. tuberculosis* e codifica uma glicosil-3-fosfoglicerato sintetase (GpgS), que pertence à família GT81 das glicosiltransferases e que usa 3-fosfoglicerato e NDP-glicose para sintetizar glicosil-3-fosfoglicerato (GPG). O segundo gene, (Rv2419c), codifica uma glicosil-3-fosfoglicerato fosfatase (GpgP), que catalisa a desfosforilação do GPG a GG. A GpgP de micobacterias, ao contrário de todas as outras GpgPs encontradas até agora, que pertencem à superfamília das hidrolases semelhantes a dehalogenases haloácidas, pertence à superfamília das fosfatases com histidina. Ambos os genes estão presentes em todas as micobactérias com genoma sequenciado, em todas as espécies que sintetizam MGLP e em muitas outras actinobactérias. As duas enzimas correspondentes foram produzidas na forma recombinante, purificadas e caracterizadas. Os dados bioquímicos obtidos revelaram que ambas são extremamente específicas.

Recentemente foi demonstrado em *M. tuberculosis* que a maltose-1-fosfato é um intermediário que liga a trealose com a síntese de glicogénio e outros α-glucanos, incluindo o MGLP e também notado que esta molécula é letal para as micobactérias se acumulada. Neste trabalho também se produziu, purificou e caracterizou a enzima micobacteriana que sintetiza maltose-1-fosfato (maltocinase). Apesar do gene que codifica esta enzima ter sido considerado essencial para *M. tuberculosis* nenhuma maltocinase tinha sido caracterizada até agora.

Ao contrário de todas as outras maltocinases estudadas, a maltocinase micobacteriana mostrou ser bastante específica para a maltose e muito flexível, ao nível do dador de fosfato, nomeadamente para ATP, GTP e UTP.

Apesar de ter sido demonstrado previamente que a maltocinase está envolvida na conversão de trealose aα-glucanos, a essencialidade desta enzima juntamente com a não essencialidade desta via metabólica indica outras funções ainda não conhecidas da maltose-1-fosfato na fisiologia micobacteriana.

A essencialidade do gene que codifica a GpgS e de outros genes provavelmente envolvidos na síntese do MGLP em *M. tuberculosis*, indica uma função possivelmente vital deste polissacarídeo neste organismo, mas que no entanto ainda carece confirmação experimental.

Esta tese revela a identidade de dois genes e enzimas envolvidos na via metabólica de síntese do MGLP e aumenta o conhecimento sobre a síntese deste importante polissacarídeo, abrindo o caminho para estudos futuros sobre a função desta estrutura na fisiologia micobacteriana.

# Chapter 1

## General Introduction

# Contents

## 1. The genus *Mycobacterium*

The genus *Mycobacterium* includes 149 validated species, most of which are environmental saprophytic species with no apparent medical relevance. However, there is a sizeable number of clinically important species that inflict a startling toll in human morbidity and mortality. This genus, is best known for its two major pathogenic species; *Mycobacterium tuberculosis* and *Mycobacterium leprae*, the causative agents of two of the world's oldest and most insidious infectious diseases, throughout the human history; tuberculosis (TB) and leprosy (Balows, 1992).

The tubercle bacillus was first detected by Robert Koch, in 1882, in stained infected tissues. The generic name *Mycobacterium* was introduced later by Lehmann and Neumann in 1896, to include the tubercle and leprosy bacilli.

Mycobacteria are aerobic, acid fast actinomycetes that usually form straight or slightly curved nonmotile rods. Many species form white or cream colored colonies, but there are also several that form bright yellow or orange colonies, containing carotenoid pigments, especially among the fast growers (Kubica & Wayne, 1984). Most pigmented species form the pigments in the dark (scotochromogenic species) but in some cases the pigments are only formed in response to light (photochromogenic species) (Dworkin & Falkow, 2006). Mycobacteria are separated in two major groups on the basis of growth rate of each of the species; fast growing and slow growing, and this forms the basis of mycobacterial taxonomy. The pathogenicity of the species follows closely, although not scrupulously, this division with the slow growers being associated with human and animal diseases while there are only scattered reports of disease caused by the fast growers, in immunocompromised hosts (Dworkin & Falkow, 2006). However some authors argue that true non-pathogenic mycobacteria, in a strict sense, may not exist (Kubica & Wayne, 1984).

Mycobacteria have indeed shaped the course of human history. Until the 20$^{th}$ century, tuberculosis was one of the chief causes of death in Europe and in the Americas.

However with the advent of the "sanatorium practices" in the early 20th century and finally with the drug discovery advances in the mid 20th century, victory was declared over the disease, though this declaration would proved to be premature. Following an extensive period of decline, the incidence of TB in the United States and Europe began to rise again in the late 1980s.

Currently, the loss of human lives for TB remains largely unhindered as a result of the synergy with the AIDS pandemic and the emergence of multi-drug- and extensively-drug resistant strains. The Centers for Disease Control have reported that each year there are almost 2 million TB-related deaths around the world and nearly 9 million people become infected. Additionally, one third of the world's population is estimated to be infected with TB (Dye, 2009). The available therapies target few bacterial functions and the growing number of extensively drug-resistant strains urges for the identification of novel pathways, new drug targets and a deeper understanding of mycobacterial physiology (Dye, 2009). Although promising new antimycobacterial agents have been recently synthesized, they are far from being available for inclusion in therapies for both drug-sensitive and drug-resistant TB (Balganesh et al, 2008; Makarov et al, 2009).

Nevertheless, the renewed research interest on *M. tuberculosis* and other mycobacteria is providing valuable insight into the basic biology of these organisms and the lessons learned will hopefully provide novel treatments to fight the renewed threat of this old enemy.

## 2. The cell envelope of mycobacteria

Mycobacteria possess a unique cell envelope, rich in long-chain (C70–C90) fatty acids known as mycolic acids (MA), and much of the pathogenicity and drug resistance is associated with this distinctive structure (Daffé & Draper, 1998; Hett & Rubin, 2008;

Portevin et al, 2005). This structure is divided in three major segments: the plasma membrane, the cell wall and the outermost layer (Fig. 1.1) (Kaur et al, 2009a). The cell wall and the outermost layer are actually very distinctive features of the *Mycobacterium* genus in comparison to other prokaryotes and the MA composition of the cell wall further allows chemical distinction of this genus from other MA containing genera; *Corynebacterium*, *Nocardia* and *Rhodococcus* (Dworkin & Falkow, 2006). The cell wall core is made of peptidoglycan (PG) in covalent attachment with Arabinogalactan (AG) which is esterified by mycolic acids (Hett & Rubin, 2008; Kaur et al, 2009a). The exterior layer consists of an assortment of non-covalently bound glycolipids, polysaccharides, lipoglycans and proteins, including porins and ESX secreted proteins (Daffé & Draper, 1998; Sani et al, 2010). This outer layer, also called capsule, is essentially composed of polysaccharides and proteins with only minor amounts of lipids (Lemassu & Daffé, 1994; Sani et al, 2010). While there has been some debate about this structure on whether it actually existed, the latest cryo-electron microscopy techniques answered that question, showing it as a clearly visible and distinct segment of the mycobacterial cell envelope (Sani et al, 2010).

The impressive progress made, since the re-emergence of TB, in the chemical definition and structure of the mycobacterial cell wall allowed a significant review on the old model proposed in the early 80s (Kaur et al, 2009a). The latest advances in microscopy techniques permitted the observation of the several layers of the mycobacterial cell envelope in their native state (Hoffmann et al, 2008; Zuber et al, 2008). The study of vitreous sections of *Mycobacterium smegmatis*, *Mycobacterium bovis* BCG and of the closely related *Corynebacterium glutamicum*, using cryo-electron microscopy, confirmed the existence of an outer bilayer structure and of a periplasmic space in these organisms. As in the previous models, these studies confirmed mycolic acids as fundamental components of the outer membrane (OM), but it was found that this structure is apparently symmetric and much thinner than

what it was thought before (Kaur et al, 2009a). The new model (Fig. 1.1) suggests that similar lipids are found in both leaflets of the OM and that the meromycolate chain of bound MAs spans through the full height of hydrophobic region (Hoffmann et al, 2008; Kaur et al, 2009a). However, other model advocates that the meromycolate chains of MAs fold upon themselves to form a more packed structure (Kaur et al, 2009a; Zuber et al, 2008).

## 2.1 The cell wall

The mycobacterial cell wall consists mainly of AG, PG and MA covalently linked together and forming the MA-AG-PG complex. This construction expands outward in layers, starting from the plasma membrane with PG and ending with the MAs (Fig. 1.1). This complex is insoluble and referred to as the core of the mycobacterial cell wall (Brennan, 2003). Many of the drugs used to fight mycobacteria target the MA-AG-PG complex (Hett & Rubin, 2008).

As a whole, the high complex mycobacterial cell wall plays essential roles in controlling the growth of the bacteria, survival in the infected host and immunologic response (Brennan, 2003; Daffé & Draper, 1998; Hett & Rubin, 2008).

### 2.1.1 Arabinogalactan

AG is the major cell wall polysaccharide of mycobacteria. It is very important for cell wall integrity and responsible for anchoring the mycolic acid layer to the PG. The galactan region beyond where arabinan attaches is of unknown function, but it was suggested that it forms a gelatinous hydrophilic region between the PG and the MA layers (Kaur et al, 2009a).

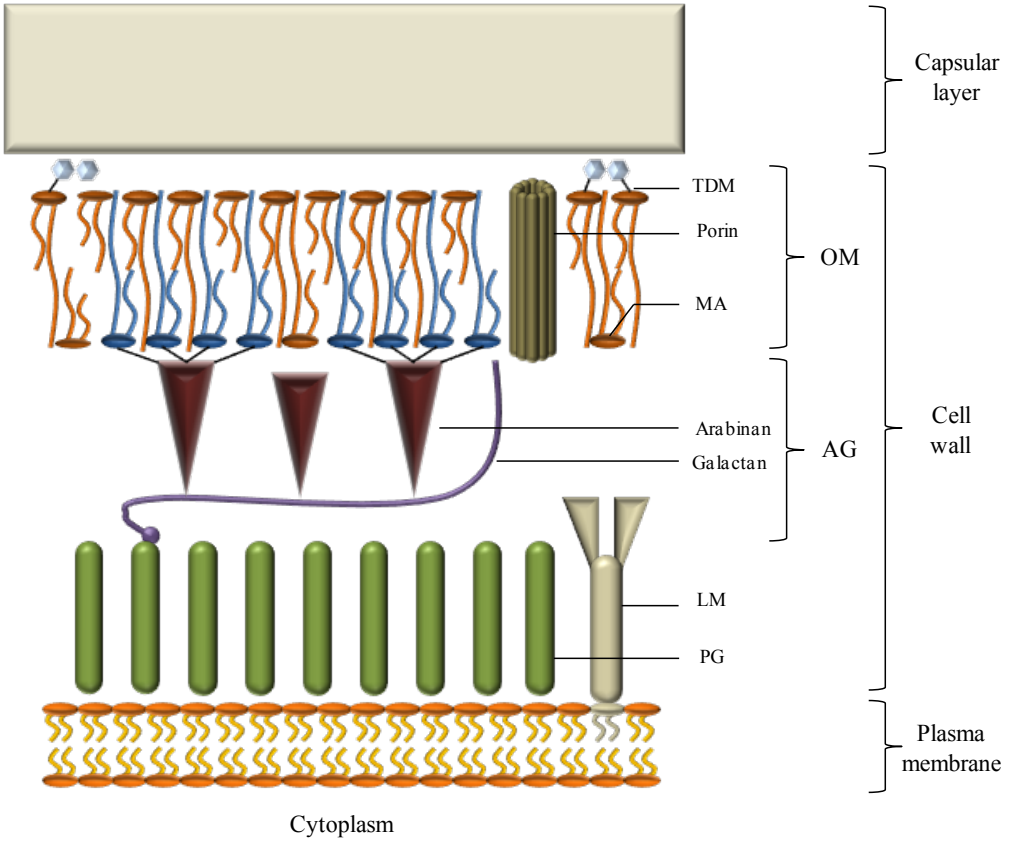



**Figure 1.1:** Schematic representation of the mycobacterial cell envelope. AG – Arabinogalactan. LM – Lipomannan. MA – Mycolic acid. OM – Outer membrane. PG – Peptidoglycan. TDM – Trehalose dimycolate.

The most recent model of AG indicates that it is composed of arabinan and galactan, both in the somewhat uncommon furanose form and with a specific linker unit ensuring a covalent attachment to PG, made of a rhamnosyl residue attached to an N-acetylglucosaminosyl-1-phosphate residue (Bhamidi et al, 2008). This linker is similar to the one used to attach teichoic acid to PG in gram-positive bacteria (Kojima et al, 1985). The galactose units of the galactan are organized in alternating (1→5) and

(1→6) linkages of roughly 30 residues, with arabinan connected to the galactan by (1→5) linkages (Hett & Rubin, 2008; Kaur et al, 2009a). The non-reducing terminus of the arabinan domain, serves as the anchoring point for the MAs (Kaur et al, 2009a).
The anti-mycobacterial activity of ethambutol, a first-line anti-TB drug, was shown to target the AG biosynthesis. This demonstrated the essentiality of AG in *M. tuberculosis* and other mycobacteria (Belanger et al, 1996). Since then, other studies have presented the genetic evidence of the essentiality of several steps in AG biosynthesis (Barry et al, 2007).

### 2.1.2 Peptidoglycan

PG is a intricate polymer, providing a rigid layer outside the plasma membrane, enough to maintain cellular shape and to resist osmotic pressure, serving also as an anchor for the other cell wall structures (Fig. 1.1) (Hett & Rubin, 2008; Kaur et al, 2009a). Almost all of the known species of bacteria contain PG, with the few exceptions including the *Mycoplasma* and, possibly, *Chlamydia* (Pavelka, 2007).
The mycobacterial PG is constituted by linear chains of N-acetyl-D-glucosamine and modified muramic acid substituted with peptide side chains that are heavily cross-linked (Kaur et al, 2009a). The total degree of cross-linking is around 70–80% in *Mycobacterium spp.* (Matsuhashi, 1966), while in *Escherichia coli* is about 50% (Vollmer & Holtje, 2004), therefore granting increased rigidity to mycobacteria.

### 2.1.3 Mycolic acids

MAs are organized in a bilayer structure with diverse glycolipids and glycoproteins intercalated with the MAs and also attached to both sides of this structure (Hoffmann

et al, 2008; Kaur et al, 2009a). These compounds in this structural arrangement form the outer membrane also known as mycomembrane (Sani et al, 2010). This structure is the principal factor for the low permeability of the mycobacterial cell wall (Hett & Rubin, 2008). MAs are found in mycobacterial species as a mixed group of structurally related molecules (Besra & Brennan, 1997; Rafidinarivo et al, 2008). They are β-hydroxy fatty acids with a long α-alkyl side chain (Takayama et al, 2005). The majority of these compounds are covalently linked to the AG by ester links (Hett & Rubin, 2008) and they have been associated with pathogenicity in numerous studies (Riley, 2006; Takayama et al, 2005). In addition, the trealose mycolate glycolipids, which are not covalently attached to the AG, are known to be important for *M. tuberculosis* physiology and virulence (Rafidinarivo et al, 2008; Riley, 2006).

### 2.1.4 Other important cell wall components

Several glycolipids like the phosphatidyl-myo-inositol mannosides (PIMs), lipomannan (LM) and lipoarabinomannan (LAM) are also present in high quantity in the mycobacterial cell wall. These molecules are attached in a non-covalently form, through their phosphatidyl-myo-inositol (PI) portion, to both leaflets of the outer membrane, but they are also found in the cytoplasmic membrane (Pitarque et al, 2008). These molecules are known to have roles in the entry of *M. tuberculosis* inside phagocytic cells, phagosome maturation and immune response modulation (Daffé & Reyrat, 2008). PIMs seem also to play central roles in the permeability of the cell envelope, cell membrane integrity and regulation of cell division (Daffé & Reyrat, 2008; Kaur et al, 2009a). Several porins and glycosylated proteins have also been reported in this structure (Hoffmann et al, 2008; Kaur et al, 2009a; Sani et al, 2010).

## 2.2 The Capsular layer

The external layer of the cell envelope of mycobacteria is a loosely bound structure, of different composition, in pathogenic and non-pathogenic species, called capsule (Chapman et al, 1959; Daffé & Draper, 1998; Hett & Rubin, 2008; Sani et al, 2010). The capsule is clearly visible as an electron transparent zone surrounding the rest of the mycobacterial cell envelope in conventional electron microscopy preparations (Frehel et al, 1986; Paul & Beveridge, 1994). This layer, however, was not visible in cryo-electron microscopy (Hoffmann et al, 2008; Zuber et al, 2008) adding to the debate on whether this structure actually existed. Further studies showed that this structure was not visible in cryo-electron microscopy due to the use of a cryoprotectant with a similar electron density of that of the capsule and if a different method was used, the capsule became clearly visible therefore demonstrating the existence of the layer (Sani et al, 2010).

The presence of this layer is highly influenced by the culturing conditions since mycobacteria are known to commonly release some of this material in the culture medium, during *in vitro* growth, if a detergent to prevent cell aggregation is added (Ortalo-Magne et al, 1995; Sani et al, 2010). Nevertheless it was shown that capsular components clearly cover *in vivo*-grown bacteria (Schwebach et al, 2002) and *in vitro*-grown bacteria in detergent devoided medium (Sani et al, 2010).

The capsular material is mainly composed by proteins and polysaccharides (97% of the total material) with only diminute amounts of lipids (Lemassu & Daffé, 1994; Ortalo-Magne et al, 1995). The small lipid content is mostly composed by PIMs and LAM (Kaur et al, 2009a; Sani et al, 2010). The protein content in this layer is very diverse and not fully unraveled; however, lipases, proteases, catalase, peroxidase, superoxide dismutase, ESX-secreted and PE/PPE proteins have been found in several mycobacteria (Daffé & Etienne, 1999; Raynaud et al, 1998; Sani et al, 2010). It has been reported that these ESX-secreted proteins and PE/PPE proteins have an essential

role in the interaction with the host (Gao et al, 2004; Pym et al, 2002; Stanley et al, 2003) and especially in bacterial translocation from the phagosome to the cytoplasm (van der Wel et al, 2007).

Three types of polysaccharides, devoided of any acylations, are found in the capsular material of *M. tuberculosis*; arabinomannan (AM), a D-mannan composed mainly of a α-(1→6)-linked D-mannose units and an high-molecular-weight (>100 kDa) α-D-glucan, very similar to glycogen, composed primarily of α-(1→4)-linked D-glucose units, branched every five or six residues by oligoglucosides (Dinadayala et al, 2004; Kaur et al, 2009a; Lemassu & Daffé, 1994; Ortalo-Magne et al, 1995).

The AM seems to be identical to LAM, apart from for the absence of the phosphatidyl-myo-inositol anchor.

The high molecular weight α-D-glucan is the main carbohydrate constituent of the capsule of *M. tuberculosis*, totaling up to 80% of the capsular polysaccharides (Lemassu & Daffé, 1994; Ortalo-Magne et al, 1995). Recent studies, based on the observation that this glucan possesses a glycogen-like structure have begun to unravel its biosynthesis. The *M. tuberculosis glg* genes, involved in the biosynthesis of glycogen, were identified and inactivated (Sambou et al, 2008). Analyses of these mutants indicated that the biosynthetic pathways of the glucan and glycogen share common enzymes, including the ADP-glucose pyrophosphorylase GlgC (Rv1213), the branching enzyme GlgB (Rv1326c), the α-(1→4)-glycosyltransferases Rv3032 and GlgA (Rv1212c), and the maltosyltransferase GlgE (Rv1327c) (Table 1.1) (Kalscheuer et al, 2010; Sambou et al, 2008). Interestingly some of these enzymes are also involved or have putative roles in the synthesis of a cytosolic polysaccharide; the methylglucose lipopolysaccharide (MGLP) (Kalscheuer et al, 2010; Stadthagen et al, 2007). The sharing of enzymes in the biosynthetical pathways of glycogen, MGLP and capsular α-D-glucan suggests that this capsular polysaccharide is synthesized in the cytoplasm before being moved to the cell surface, by unknown transporters (Kaur et al, 2009a). Several biological roles, including antiphagocytic and immunomodulatory

activities have been attributed to the polysaccharides present in the capsule (Cywes et al, 1996; Gagliardi et al, 2007; Geurtsen et al, 2009; Stokes et al, 2004). Specifically the α-D-glucan has been shown to induce monocytes to differentiate into altered dendritic cells, that fail to present lipid antigens to T cells (Gagliardi et al, 2007). The α-D-glucan was also proposed to be involved in *M. tuberculosis* evasion of the immune system (Lemassu & Daffé, 1994). Most of these findings however, were derived from *in vitro* studies, relying on purified capsular polysaccharides and cellular models, and consequently may not accurately reflect the physiological role of the polysaccharides. Therefore, Sambou and colleagues tried to further clarify the role of α-D-glucan in *M. tuberculosis* infection and persistence, by creating knock-out mutants of the genes putatively involved in its biosynthetic pathway and infecting mice with the mutants (Sambou et al, 2008). However, due to the overlap of the glycogen, MGLP and capsular glucan pathways, the authors were unable to clearly ascertain the role of the molecule.

Overall, the capsule being at the interface between the bacterium and host cells was proven to be involved in bacterial pathogenicity. It mediates the cell adhesion to and the infiltration of bacilli into the host's cells (Daffé & Etienne, 1999; Stokes et al, 2004). In addition, the capsule also represents a further passive barrier by preventing the diffusion of macromolecules towards the inner parts of the envelope (Daffé & Etienne, 1999). Finally, toxic lipids and constituents that exhibit immunomodulatory properties have been found in the capsule, thereby explaining part of the immunopathology of tuberculosis (Daffé & Etienne, 1999; Raynaud et al, 1998; Stokes et al, 2004).

**Table 1.1:** Genes and enzymes involved in the synthesis of MGLP, glycogen and α-glucan

| gene | designation | function | pathway |
|---|---|---|---|
| **Rv1212c** | ***glgA*** | α-(1→4)-glycosyltransferase | glycogen<br>α-glucan<br>MGLP |
| **Rv1213** | ***glgC*** | ADP-glucose pyrophosphorylase | glycogen<br>α-glucan<br>MGLP* |
| **Rv1326c** | ***glgB*** | Glycogen branching enzyme | glycogen<br>α-glucan |
| **Rv1327** | ***glgE*** | Maltosyltransferase | glycogen<br>α-glucan<br>MGLP* |
| **Rv3032** | - | α-(1→4)-glycosyltransferase | glycogen<br>α-glucan<br>MGLP |

*** - putative involvement**

## 3. Mycobacterial intracellular polysaccharides

Mycobacteria are known to synthesize an array of carbohydrates both intracellular and cell wall related, some of which unique to the order Actinomycetales (Berg et al, 2007). Two groups of intracellular polysaccharides have been studied in detail; glycogen and polymethylated polysaccharides (PMPS). However, the full physiological role of the PMPS is not yet fully understood nor its biosynthesis unraveled (Jackson & Brennan, 2009).

### 3.1 Glycogen

Glycogen serves as a major carbohydrate reserve in most bacteria and in mycobacteria was shown to be vital for growth (Belanger & Hatfull, 1999). Interestingly, the structures of the capsular α-D-glucan and glycogen are almost indistinguishable. Several of the enzymes involved in glycogen biosynthesis are also known to be involved in the synthesis of the capsular α-D-glucan and apparently, also in the synthesis of the methylglucose lipopolysaccharide (Elbein et al, 2010; Kalscheuer et al, 2010; Kaur et al, 2009a; Sambou et al, 2008)

### 3.2 Polymethylated Polysaccharides

PMPS were initially isolated from *Mycobacterium phlei* and *M. tuberculosis* in Dr. Clinton Ballou's laboratory (Gray & Ballou, 1971; Lee, 1966; Lee & Ballou, 1964) and, to a great extent, the available information about these molecules comes from those early studies.

PMPS are polysaccharides composed of 10 to 20 hexoses many of which are *O*-methylated, conferring to these molecules slight hydrophobic properties. Two different PMPS have been isolated from mycobacteria; the methyl mannose polysaccharide (MMP) and the methyl glucose lipopolysaccharide (MGLP). Although they are chemically different, their structure is similar and has been studied in detail (Fig. 1.2) (Maitra & Ballou, 1977; Tuffal et al, 1998; Yamada et al, 1979).

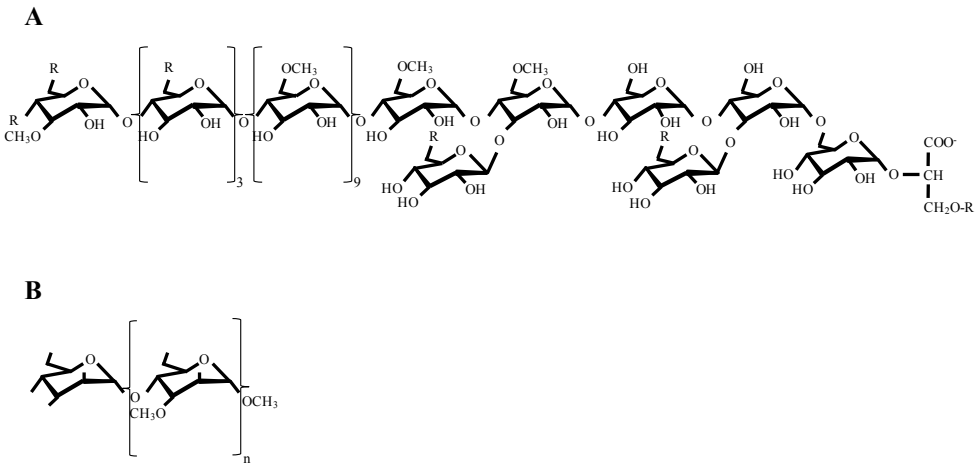



**Figure 1.2:** Structure of the mycobacterial polymethylated polysaccharides. **(A)** MGLP. **(B)** MMP

### 3.2.1 Methylmannose Polysaccharide

MMP is composed of 10–13 α-(1→4)-linked 3-*O*-methyl-D-mannoses terminated at the nonreducing end by a α-methyl aglycon. There is some heterogeneity in MMP structure as it occurs as a mixture of at least four isomers with differences in size and degree of *O*-methylation (Jackson & Brennan, 2009; Maitra & Ballou, 1977; Weisman & Ballou, 1984a; Weisman & Ballou, 1984b). MMP has been found in non pathogenic fast-growing mycobacteria and closely related compounds have also been reported in *Streptomyces griseus* (Harris & Gray, 1977; Jackson & Brennan, 2009; Weisman & Ballou, 1984b). 3-*O*-methyl-D-mannose has also been reported in polysaccharides of *Klebsiella spp*., *Escherichia coli* and *Leptosira biflexa* (Azuma et al, 1976; Nimmich, 1970). However, virtually nothing is known about these compounds or if they are structurally or functionally related to the mycobacterial MMP. Also there has been no

report of the production of this polysaccharide in slow-growing mycobacteria, and the inability to detect MMP in *M. tuberculosis* raises the possibility that it may be restricted only to fast-growing mycobacteria (Jackson & Brennan, 2009; Stadthagen et al, 2007).

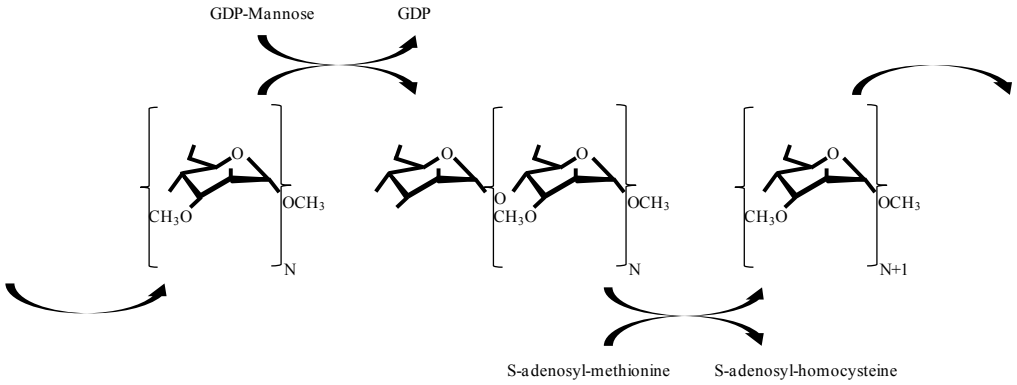



**Figure 1.3:** Proposed mechanism for the synthesis of MMP

Information about the synthesis of MMP is limited and restricted to the early studies performed in Ballou's laboratory. The purification of biosynthetic precursors of MMP and the biochemical characterization of a 3-*O*-methyltransferase and of an α-(1→4)-mannosyltransferase from *Mycobacterium smegmatis* cell-free extracts led to the proposal of a biosynthetic model in which MMP is extended by a linear sequential process of mannosylation and methylation (Fig. 1.3) (Maitra & Ballou, 1977; Weisman & Ballou, 1984a; Weisman & Ballou, 1984b; Yamada et al, 1979). GDP-mannose is the sugar source for the mannosyltransferase while *S*-adenosylmethionine (SAM) is the donor of methyl groups for the methyltransferase (Weisman & Ballou, 1984a; Weisman & Ballou, 1984b). The elongation reaction ends when the mannose

chain reaches a length suitable for fatty acid-binding; 11 to 13 3-*O*-methylmannoses (Jackson & Brennan, 2009). However, it is not clear how the initiation of the MMP chain is done since the enzymes involved in the elongation of the chain have no activity for the first precursors (Weisman & Ballou, 1984b). None of the genes involved in the biosynthesis of MMP have been identified and purification of the enzymes was not attempted by Ballou and associates due to severe enzyme instability (Weisman & Ballou, 1984b).

### 3.2.2 Methylglucose Lipopolysaccharide

While the MMP appears to be restricted to fast growing mycobacteria and *Streptomyces griseus*, MGLP has been isolated from several *Nocardia* species as well as *Mycobacterium bovis* BCG, *Mycobacterium leprae*, *Mycobacterium smegmatis*, *Mycobacterium xenopi* and *M. tuberculosis* (Jackson & Brennan, 2009; Lee, 1966; Lornitzo & Goldman, 1968; Pommier & Michel, 1986; Tuffal et al, 1998; Tuffal et al, 1995). MGLP is mainly composed by α–(1→4)-linked glucose and methylglucose units. It also has a α–(1→6)-linked glucosylglycerate (GG) to its reducing end, two β–(1→3)-linked glucoses at glucose 1 and 3 of the main α–(1→4)-chain (Tuffal et al, 1998) (Fig. 1.2). Like MMP this polysaccharide also presents some structure heterogeneity with the occurrence of a variable degree of acylations and acyl groups (Gray & Ballou, 1972). The α–(1→4)-linked *O*-methylated hexoses of both polysaccharides lead to an helical conformation characteristic of amylose, where the methyl groups are faced to the inner part of the coil forming a nonpolar and hydrophobic cavity (Peterson & Bloch, 1977) . It was also proposed that the acylation of MGLP may alter its fatty acid-binding specificity and serve to fine-tune its regulatory activity in the cell (Hindsgaul & Ballou, 1984; Jackson & Brennan, 2009).

Like in the case of MMP, Ballou's group also unraveled the foundations for the MGLP biosynthesis. However, since this compound is the only PMPS present in pathogenic mycobacteria, it received some attention in the genomic era allowing the identification of several genes involved in the process. The proposed model for MGLP synthesis emerged from Ballou's early work and is similar to the one suggested for MMP. Using similar approaches for the two polysaccharides, Ballou and co-workers proposed a model in which the elongation of the chain progresses, from the reducing toward the nonreducing end, through alternating glucosylation and methylation reactions (Fig. 1.4) (Ferguson & Ballou, 1970; Grellert & Ballou, 1972; Kamisango et al, 1987). Methyl groups were found to be transferred from S-adenosyl-methionine (SAM) to position 6 of some of the glucose units, and a membrane associated acyltransferase, capable of catalyzing the transfer of acetyl, propionyl, isobutyryl, octanoyl, and succinyl groups, from their respective acyl-CoA derivatives onto purified MGLP and partially acetylated α-(1→4)-linked D-gluco-oligosaccharides, was also reported (Grellert & Ballou, 1972; Jackson & Brennan, 2009; Tung & Ballou, 1973). It was also suggested that methylation and acylation take place together, the later controlling the former (Ferguson & Ballou, 1970; Grellert & Ballou, 1972). However, many of biosynthetic steps of MGLP, proposed by Ballou and associates (Kamisango et al, 1987) are still hypothetical.

Recently, a cluster of genes (Fig. 1.5) was reported to be involved in the synthesis of MGLP with the genes responsible for the 6-*O*-methylations of the glucose units (Rv3030) and for the α-(1→4) chain extension being identified (Rv3032 and Rv1212c) (Stadthagen et al, 2007). Interestingly, the later genes were also reported to be involved in the synthesis of glycogen and in the capsular α-D-glucan (Sambou et al, 2008). UDP-glucose serves as the sugar donor for both glycosyltransferases and SAM is the source of methyl groups for the methyltransferase (Sambou et al, 2008; Stadthagen et al, 2007). Gene disruption studies in *M. smegmatis* showed that a mutant of the Rv3030 ortholog had a defect in *O*-methylation of the glucose units of the

polysaccharide, that almost completely inhibited MGLP synthesis, confirming the involvement of this gene in the MGLP pathway (Stadthagen et al, 2007). Similarly, a disruption of the Rv3032 gene in *M. tuberculosis* decreased, but not totally vanished, the MGLP levels present in that organism (Stadthagen et al, 2007). However, a construction of a double Rv1212c and Rv3032 mutant proved to be unfeasible (Stadthagen et al, 2007) and it was latter reported that a functional copy of one of these genes is absolutely required for mycobacterial survival (Kalscheuer et al, 2010; Sambou et al, 2008). Interestingly another gene involved in glycogen biosynthesis encoding a maltosyltransferase (Rv1327) was postulated to have a role in MGLP α-(1→4) chain elongation (Kalscheuer et al, 2010). Close to the glucosyltransferase and methyltransferase genes, lies a putative acetyltransferase gene (Rv3034c), a putative SAM-dependent methyltransferase (Rv3037c), and a putative α-amylase/glycosyl hydrolase/GH-57 family branching enzyme gene (Rv3031). The Rv3031 gene was suggested to be the responsible for generating the α-(1→6)-glycosidic bond linking the first and second glucose residues, at the reducing end of the molecule, in a mechanism similar to the one involved in the branching of glycogen (Jackson & Brennan, 2009; Stadthagen et al, 2007). The Rv3037c gene was suggested to be the 3-*O*-methyltransferase responsible for the 3-*O*-methylation of the glucosyl unit of the non-reducing end of MGLP and the Rv3034c the gene responsible for at least some of the acylations (Jackson & Brennan, 2009). Despite being proposed to be involved in MGLP pathway, all of the functions of these genes still await experimental confirmation. The identity of the gene(s) encoding the glycosyltransferase(s) responsible for the glycosylation at position 3 of the second and fourth glucose residues close to the reducing end of MGLPs remain unknown.

At this point nothing was known about the early steps of MGLP biosynthesis, except that GG had been proposed to be the first precursor in this pathway (Kamisango et al, 1987). However, after the initial detection in *M. smegmatis* cell extracts, GG has been detected in diverse organisms from methanogenic archaea, proteobacteria,

hyperthermophilic bacteria to cyanobacteria were it appears to function as a compatible solute (Costa et al, 2007; Goude et al, 2004; Klahn et al, 2010; Robertson et al, 1992). Two different pathways for the synthesis of GG were found and described by our group; first in *Methanococcoides burtonii* and later in *Persephonella marina* (Costa et al, 2007; Costa et al, 2006; Fernandes et al, 2007). In a direct reaction a glucosylglycerate synthase condenses NDP-glucose and D-glycerate into GG (Fernandes et al, 2007). In another and more common reaction, a phosphorylated intermediate, glucosyl-3-phosphoglycerate (GPG), is first synthesized from NDP-glucose and 3-phosphoglycerate by a glucosyl-3-phosphoglycerate synthase (GpgS) and subsequently dephosphorylated by a specific phosphatase, glucosyl-3-phosphoglycerate phosphatase (GpgP) (Costa et al, 2007; Costa et al, 2006).

### 3.2.3 Biological functions of the PMPS

Extensive *in vitro* studies in cell free assays have highlighted the unique regulatory roles of MMP and MGLP on mycobacterial fatty acid synthase I (FAS-I). In these studies both PMPS raise the overall rate of synthesis of fatty acids by FAS-I and shift the length pattern of newly synthesized chains from long ($C_{20} - C_{24}$) to short ($C_{14} - C_{18}$) (Banis et al, 1977; Bloch, 1977; Flick & Bloch, 1974; Machida & Bloch, 1973). These effects are thought to result from the ability of the PMPS to form 1:1 complexes with fatty acyl chains and acyl-CoAs and sequestering the products of FAS-I (Machida & Bloch, 1973). PMPS markedly decrease the $K_m$ of the FAS-I for acyl-CoAs and also seem to enable a higher enzyme turnover, by promoting a faster release of the newly synthesized fatty acid chains, and terminating the elongation (Banis et al, 1977; Bloch & Vance, 1977; Flick & Bloch, 1974; Ilton et al, 1971; Machida & Bloch, 1973). In addition, while complexing the acyl-CoAs released by the FAS-I

complex, the PMPS are thought to mitigate product inhibition of the synthase (Bloch, 1977; Bloch & Vance, 1977; Flick & Bloch, 1975).

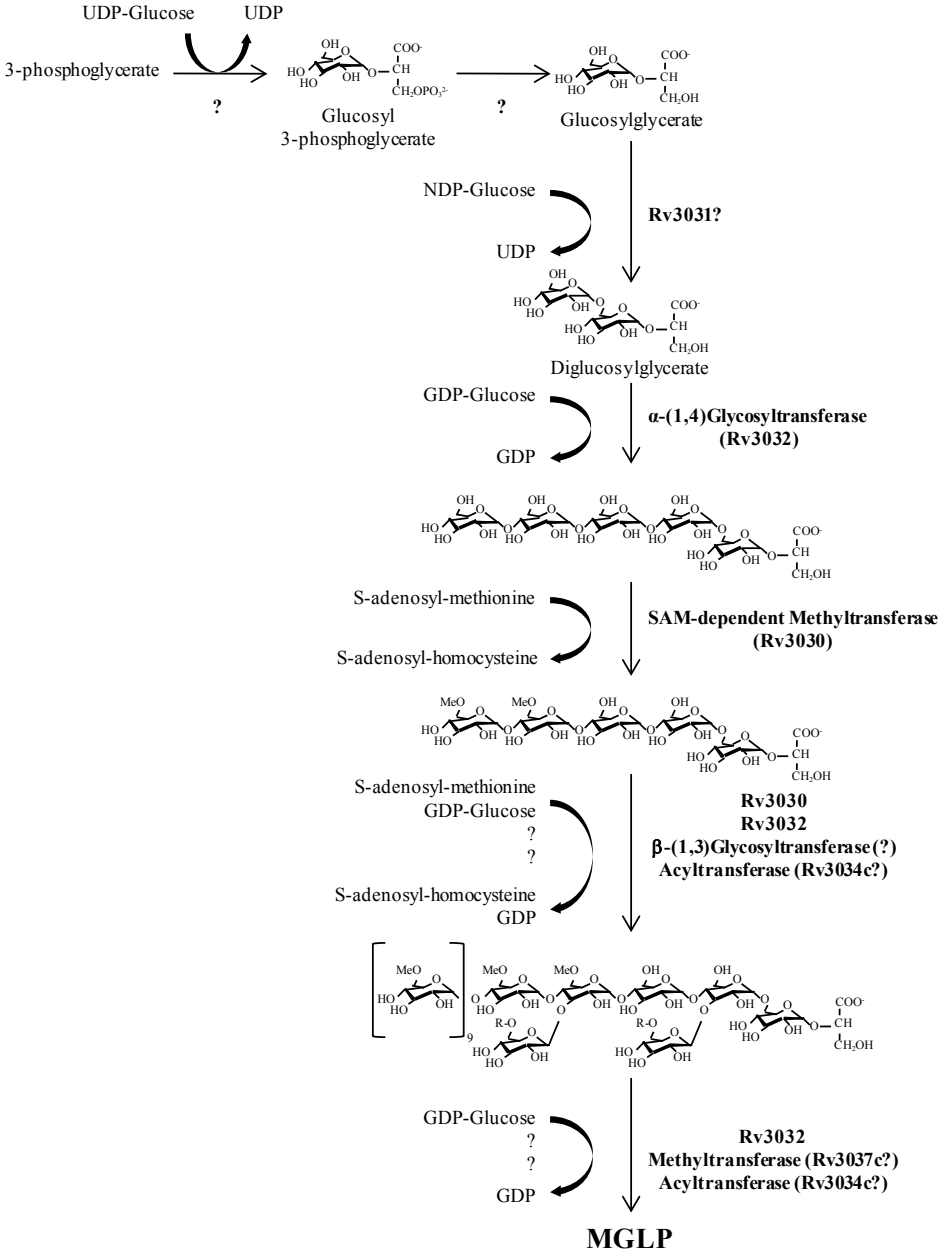


**Figure 1.4:** Proposed pathway for the biosynthesis of MGLP at the start of this work. Partially adapted from (Jackson & Brennan, 2009)

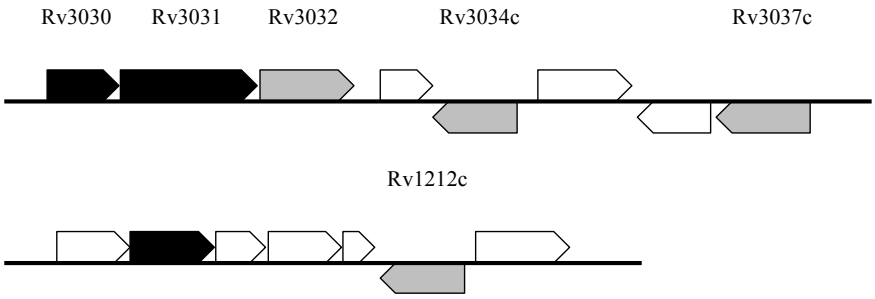



**Figure 1.5:** Known and putatively involved genes in MGLP synthesis in *Mycobacterium tuberculosis* at the start of this work. (■) genes proposed to be essential for growth (Sassetti et al, 2003). (■) non essential genes (Sassetti et al, 2003).

PMPS have been found to reach intracellular concentrations approaching 1 mM in *M. smegmatis* while the intracellular concentration of long-chain acyl-CoAs is about 0.3 mM (Yabusaki et al, 1979). With an projected dissociation constant of the polysaccharide-lipid complex of 0.1 μM, all of the cytoplasmic long-chain fatty acids may form complexes with PMPS under these circumstances (Yabusaki et al, 1979). Consequently PMPS can be viewed as the fatty acid carriers in mycobacteria, protecting long-chain fatty acyl-CoAs from the attack by degradative enzymes (Yabusaki & Ballou, 1979), preventing general metabolism disruption due to fatty acid accumulation (Kawaguchi & Bloch, 1974; Kawaguchi & Bloch, 1976; Mita & Yasumasu, 1981; Wititsuwannakul & Kim, 1977) and regulating their further processing for the synthesis of more complex fatty acids and lipids, including mycolic acids (Bloch, 1977; Forsberg et al, 1982; Yabusaki et al, 1979). While complexed with PMPSs, the fatty acid chain is included in the nonpolar cavity of the coiled polysaccharide chain.

So far, the physiological evidence for a role of PMPS in the regulation of mycobacterial lipid metabolism is still lacking. Only one study was published on the role of PMPS in mycobacterial fatty acid metabolism *in vivo* (Maloney & Ballou, 1980). A *M. smegmatis* spontaneous mutant, which contained only 50 and 7% of the wild-type levels of MGLP and MMP, respectively, did not have fatty acid synthesis dramatically altered, although the mutant accumulated more short-chain and less long-chain unsaturated fatty acids than the wild-type strain. Further mutagenesis studies with *M. smegmatis* and *M. tuberculosis* that targeted some of the genes involved in MGPL biosynthesis failed to shed further light onto this subject (Stadthagen et al, 2007). While MGLP levels in the *M. smegmatis* mutants were compromised, MMP could still replace MGLP. Moreover, disruption of Rv3032 in *Mycobacterium tuberculosis* was not lethal as this function was partially compensated by Rv1212c (Sambou et al, 2008; Stadthagen et al, 2007). Interestingly, Rv3032 and Rv3030 mutants of *M. tuberculosis* and *M. smegmatis* displaying reduced MGLP contents had their growth dramatically affected at high temperatures (Stadthagen et al, 2007), suggesting a possible role for the MGLP in thermal adaptation.

Despite the absence of physiological evidence, the reported essentiality of some of the genes known to be involved in MGLP biosynthesis in *M. tuberculosis* (Sassetti et al, 2003; Sassetti & Rubin, 2003) emphasizes that this structure is likely to have an important role in mycobacterial physiology.

## 4. Maltose-1-phosphate and Trehalose

Maltose-1-phosphate was initially identified in *Mycobacterium bovis* BCG cell extracts more than 40 years ago (Narumi & Tsumita, 1967), but no further research was carried out to clarify its role.

A maltose-1-phosphate synthesizing enzyme (maltokinase) was much later identified in *Actinoplanes missouriensis*, where it was constitutively expressed, independently of the sugars present in the growth medium (Drepper et al, 1996). A hypothetical role for maltokinase in maltose catabolism was excluded (Niehues et al, 2003) and the putative involvement of this enzyme in the phosphorylation of sugars during transport via phosphotransferase systems was also discarded since the transported disaccharides are exclusively phosphorylated in position 6 (Postma et al, 1993). Interestingly, maltose-1-phosphate could be synthesized without the requirement of ATP in *E. coli*, and it was proposed to be involved in the regulation of maltose metabolism (Decker et al, 1999). Also, worthy of note is that a maltose-1-phosphate related molecule, trehalose-6-phosphate, was recently acknowledged as an important signaling molecule found to regulate sugar metabolism in yeast, fungi, plants and some invertebrates (Paul et al, 2008). Moreover, maltose-1-phosphate was recently found to be toxic and lethal to *M. tuberculosis* if accumulated, triggering pleiotropic stress responses, eventually leading to cell death (Kalscheuer et al, 2010).

The maltokinase gene (*mak*) is present in most of the available mycobacterial genomes and it was considered essential for the growth of *M. tuberculosis* H37Rv (Sassetti et al, 2003). Furthermore, *mak* is always linked with the trehalose synthase gene (*treS*) and in some organisms fused into one bifunctional gene. This highly conserved sequential arrangement is found in many bacterial groups and suggests a shared metabolic route.

Trehalose is a ubiquitous disaccharide found in bacteria, archaea, fungi, plants and in some invertebrates, where it plays a multitude of biological roles: osmotic and thermal protection and even metabolic signaling (Iturriaga et al, 2009; Paul et al, 2008). In mycobacteria and related organisms, trehalose is synthesized by three different pathways; The OtsA/OtsB pathway, which uses glucose and glucose-6-phosphate, the TreY/TreZ pathway, which makes trehalose from glycogen, and the TreS enzyme which can convert maltose to trehalose (Woodruff et al, 2004).

Trehalose in mycobacteria has a functional role as a compatible solute, under osmotic or thermal stress, and a structural role in cell wall components, where it is linked to mycolic acids to form trehalose monomycolate or dimycolate (De Smet et al, 2000; Takayama et al, 2005; Woodruff et al, 2004) a glycolipid that modulates the host immune response (Riley, 2006) and that is also important for both the physiology and virulence of these organisms (Hunter et al, 2006).

*M. smegmatis* mutants defective in the three pathways are unable to grow unless trehalose is supplemented to the growth medium (Woodruff et al, 2004). Each pathway however, seems to have a specific role and hierarchy in closely related actinobacterial species (Murphy et al, 2005; Tzvetkov et al, 2003; Wolf et al, 2003; Woodruff et al, 2004).

The trehalose synthase (TreS) from *M. smegmatis* is capable of interconverting trehalose and maltose (Pan et al, 2004). TreS was also shown to catalyze the formation of trehalose from glycogen via its amylase activity and with maltose as the intermediate (Pan et al, 2008). Moreover, the synthesis of glycogen from trehalose has also recently been reported in *M. smegmatis* and *M. tuberculosis* (Elbein et al, 2010; Kalscheuer et al, 2010). In this proposed three-step pathway, TreS and Mak sequentially convert trehalose into maltose and further into maltose-1-phosphate, while a final maltosyltransferase (GlgE) transfers maltose from maltose-1-phosphate to glycogen (Fig. 1.6) (Kalscheuer et al, 2010). This gene, as referred previously, is also known to be involved in capsular α-D-glucan biosynthesis, and it has been suggested that it may also have a role in MGLP biosynthesis (Kalscheuer et al, 2010). Among the enzymes involved in this pathway, only the maltokinase was initially reported as essential for *M. tuberculosis* (Sassetti et al, 2003). Later however, the GlgE was also shown to be essential (Kalscheuer et al, 2010), but its essentiality ceases if the pathway becomes inactivated by TreS inactivation, which seems to be the only true non essential gene in this pathway (Kalscheuer et al, 2010). It is interesting to note that despite having several essential genes, the pathway itself is not essential

but becomes crucial if the redundant Rv3032 pathway is not functional (Kalscheuer et al, 2010). Simultaneous inhibition of both pathways is lethal and further confirms an overlap of both pathways and a possible involvement of Mak in MGLP synthesis (Kalscheuer et al, 2010).

Despite of the increasing numbers of genome sequences showing that *mak* genes are widely distributed throughout the bacterial world, there is still a tremendous lack of data on the corresponding enzymes. Overall, the role of maltose-1-phosphate and of the maltokinase in the organisms containing this gene is highly unexplored and of great interest, especially in *M. tuberculosis*, where the corresponding gene was reported to be essential for growth (Sassetti et al, 2003).

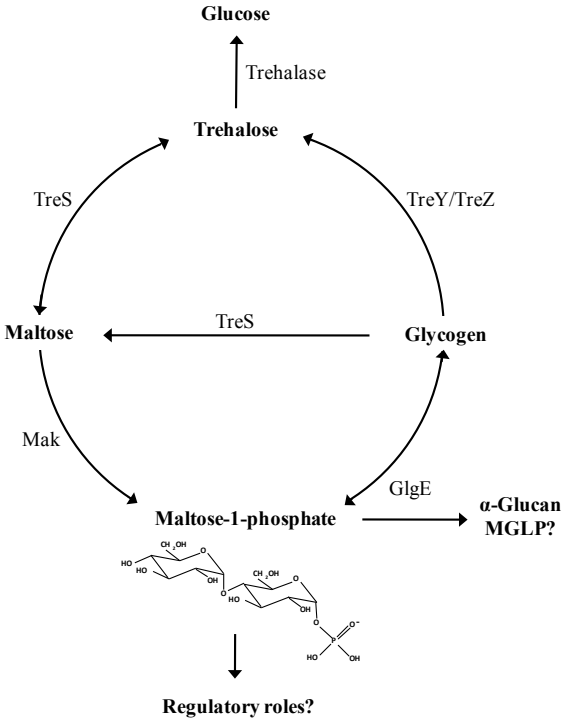



**Figure 1.6:** Metabolic circuit between trehalose, maltose, glycogen and other α-glucans

## 5. Histidine phosphatase superfamily

The histidine phosphatase superfamily is a large and functionally diverse group of proteins containing two deep rooted branches that share very little sequence similarity (Rigden, 2008). All the proteins of this superfamily share a conserved catalytic core, an RHG motif, located at the beginning of the sequence, centered on a histidine that becomes momentarily phosphorylated during the reaction (Foster et al, 2010; Rigden, 2008).

The earliest known, and possibly best studied member, is the cofactor-dependent phosphoglycerate mutase (dPGM), discovered in the 1930s (Fothergill-Gilmore & Watson, 1989). Since then, many different functions have been associated with this superfamily, but all of these new discoveries have phosphatase activities, illustrating that, the earliest member of the superfamily is an outsider in catalyzing a mutase reaction (Jedrzejas, 2000; Rigden, 2008).

The first and largest branch of this superfamily contains the oldest known members of this family (dPGM), but also the fructose-2,6-bisphosphatase and many other activities (Jedrzejas, 2000; Rigden, 2008). The second branch is smaller and composed mainly of acid phosphatases, most of unspecified function, and phytases (Rigden, 2008). The relation between these two branches is not evident by Blast analysis but becomes clear when structural approaches are made (Rigden, 2008).

This protein superfamily, like many others, has been subjected to extensive mis-annotations mostly derived, in this case, from the mutase-dominated superfamily history, while the phosphatase activity is actually the dominant (Rigden, 2008). Also, and due to their impressive sequence and functional diversity, the histidine phosphatases present significant challenges to automated genome annotation tools (Friedberg, 2006; Rigden, 2008).

The histidine phosphatase superfamily, despite having an increasingly list of known functions, without a doubt harbours many novel activities that still await disclosure

(Rigden, 2008). Structure and sensitive sequence analyses of this family portray a clear image of evolution of the two branches. In the first branch, close to two thirds of its members are bacterial proteins, usually intracellular and of broadly varied functions. The second branch includes some well studied enzymes like the phytases and other members like the acid phosphatases of uncertain function. This branch is mainly eukaryotic and extracellular (Rigden, 2008).

Structurally, all histidine phosphatases contain a conserved catalytic core, centered on the RHG motif, and conserved additional residues that help in preserving the proper orientation of this core (Jedrzejas, 2000). However, they present considerable differences in their proton donor and in substrate-binding residues (Rigden, 2008). This substantial structural variability, of the specificity determinants makes functional annotation from sequences in this superfamily even harder, and imply that structural knowledge will be vital for both reliable annotation of new sequences with known activities and for prediction of novel activities (Rigden, 2008). Nevertheless, experimental activity demonstration will play a definitive role in reliable annotation.

# Chapter 2

## Actinobacterial glucosyl-3-phosphoglycerate synthase

**This chapter is partially published in:**

Empadinhas N, Albuquerque L, Mendes V, Macedo-Ribeiro S, da Costa MS (2008). **Identification of the mycobacterial glucosyl-3-phosphoglycerate synthase.** *FEMS Microbiology Letters* **280**(2)**:** 195-202.

Fernandes C*****, Mendes V*****, Costa J, Empadinhas N, Jorge C, Lamosa P, Santos H, da Costa MS (2010). **Two alternative pathways for the synthesis of the rare compatible solute Mannosylglucosylglycerate in *Petrotoga mobilis*.** *Journal of Bacteriology* **192**(6)**:** 1624-1633.

***** These authors contributed equally to this work

# Contents

## Abstract

Mycobacteria are known to synthesize an array of fatty acids, including mycolic acids, and much of their pathogenicity is associated with the unique cell wall they assemble. These organisms synthesize unique polymethylated polysaccharides that regulate fatty acid synthesis, namely the methylglucose lipopolysaccharide (MGLP) and the methylmannose polysaccharide.

Glucosylglycerate (GG) is found at the reducing end of MGLP. The mycobacterial gene encoding a glucosyl-3-phosphoglycerate synthase (GpgS), belonging to the GT81 family of glycosyltransferases, primarily found in actinobacteria and sharing very low amino acid identity with known homo-functional GpgSs, has been identified. The *gpgS* genes from the fast-growing *Mycobacterium smegmatis* strain 1102, from the slow-growing *Mycobacterium bovis* BCG and from the thermophilic fast-grower *Mycobacterium hassiacum* were expressed in *Escherichia coli*, and the recombinant enzymes were purified and characterized. The actinobacterial-type GpgS from *Petrotoga mobilis* was also expressed in *E. coli* and the recombinant enzyme purified and characterized. In *Mycobacterium tuberculosis* H37Rv, the gene encoding GpgS (Rv1208) is 100% identical to the homologue in *M. bovis* BCG and was considered to be essential for growth.

The substrates for optimal activity for all the tested GpgSs were UDP-glucose and D-3-phosphoglycerate but ADP–glucose was also an efficient donor. All the enzymes were strictly dependent on divalent cations.

## Introduction

Despite the availability of an effective vaccine (BCG) and short-course chemotherapy, *Mycobacterium tuberculosis* continues to claim more lives than any other single infectious agent (Dye, 2009). The Centers for Disease Control has reported that each year, there are almost two million tuberculosis (TB) related deaths around the world and nearly nine million become sick. Moreover, one third of the world's population is infected with TB (Dye, 2009). In recent years the incidence of TB in both developing and industrialized countries has dramatically increased and the widespread emergence of drug-resistant strains together with a deadly synergy with the human immunodeficiency virus led the World Health Organisation (WHO) to declare tuberculosis a global emergency. Radical measures are needed now to prevent the grim predictions of the WHO becoming reality.

Mycobacteria synthesize two unique hexose based and highly methylated polysaccharides, thus conferring the molecules with a slight hydrophobicity (Yabusaki & Ballou, 1978; Yabusaki et al, 1979). One is the methylmannose polysaccharide (MMP), while the other is the methylglucose lipopolysaccharide (MGLP) (Gray & Ballou, 1971; Lee, 1966; Lee & Ballou, 1964). These polysaccharides form 1:1 ratio complex with palmitoyl-CoA (Yabusaki & Ballou, 1978; Yabusaki et al, 1979), preventing degradation of the fatty acids chains (Yabusaki & Ballou, 1979) and allowing mycobacteria to continuously synthesize palmitoyl-CoA without hampering cell metabolism (Flick & Bloch, 1975; Kawaguchi & Bloch, 1974; Kawaguchi & Bloch, 1976; Knoche et al, 1973; Mita & Yasumasu, 1981). It seems that both polysaccharides modulate fatty acids synthesis in mycobacteria. (Banis et al, 1977; Bloch, 1977; Flick & Bloch, 1974; Ilton et al, 1971).

MGLP is a complex compound of 16–20 glucose and methylglucose units mainly arranged in an α–(1→4)-linked chain, leading to a helical conformation (Peterson & Bloch, 1977). It also contains glucosylglycerate (GG) attached to its reducing end and

two β–(1→3)-linked glucoses branching from the main α–(1→4)-chain (Tuffal et al, 1998). It also possesses a variable degree of acylations (Gray & Ballou, 1972; Smith & Ballou, 1973). Some of the genes putatively assigned to MGLP synthesis were considered essential for mycobacterial growth (Sassetti et al, 2003; Stadthagen et al, 2007).

Free GG and diglucosylglycerate have been detected in small amounts in *Mycobacterium smegmatis* and considered to be the putative precursors for MGLP synthesis (Kamisango et al, 1987).

Interestingly, *Petrotoga miotherma* was found to accumulate a compatible solute α–D–mannopyranosyl–(1→2)–α–D–glucopyranosyl–(1→2)–glycerate (MGG) (Jorge et al, 2007), a compound very similar to the diglucosylglycerate found on the reducing end of the mycobacterial MGLP, whose precursors are GDP-mannose and GPG (Fernandes et al, 2010).

GG is a multipurpose molecule. It has been detected in a diverse range of organisms were it appears to function as a compatible solute that accumulates under osmotic stress or as an intermediate to glycolipids or polysaccharides (Costa et al, 2007; Goude et al, 2004; Kamisango et al, 1987; Klahn et al, 2010; Pommier & Michel, 1986; Robertson et al, 1992). Two pathways have been found for the synthesis of glucosylglycerate (Costa et al, 2006; Fernandes et al, 2007). One reaction involves the direct condensation of NDP–glucose with D-glycerate by glucosylglycerate synthase (Fernandes et al, 2007). However, the most common pathway involves the synthesis of the phosphorylated intermediate glucosyl-3-phosphoglycerate (GPG) from NDP-glucose and 3-phosphoglycerate by glucosyl-3-phosphoglycerate synthase (GpgS), and dephosphorylation by specific phosphatases (Costa et al, 2006). We have identified distantly related *gpgS* homologues in mycobacterial and *Petrotoga mobilis* genomes that belonged to the GT81 family of glycosyltransferases. We were compelled to recombinantly express the genes from three *Mycobacterium* species in *E. coli* for functional characterization. The two species from which the MGLPs have

been studied were chosen; the fast-growing *M. smegmatis* and the slow-growing *Mycobacterium bovis* BCG (Kamisango et al, 1987; Tuffal et al, 1998); and the thermophilic *Mycobacterium hassiacum*, confirming that the recombinant enzymes had GpgS activity *in vitro*. A fourth GpgS gene from *Petrotoga mobilis* was also expressed in *E. coli*. In addition, a comparison between several mycobacterial GpgSs and a mycobacterial-type GpgS from the distantly related bacterium *P. mobilis* was performed.

## Methods

### DNA, bacterial strains and growth conditions

DNA from *M. bovis* BCG (DSM 43990) was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ, Braunschweig, Germany) and DNA from *M. smegmatis* strain 1102 (ATCC 23037D-5) obtained from the American Type Culture Collection. *P. mobilis* (DSM 10674) and *M. hassiacum* (DSM 44199) were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (Germany).

For DNA extraction *M. hassiacum* was grown on a modified DSMZ GFPH medium (DSM 533 medium) at pH 7.2 containing per liter: 10 g glucose (Merck), 5 g tryptone (Difco), 5 g yeast extract (Difco), 5 g beef extract (Merck), 0.74 g calcium chloride (Merck) and 2 g of tween 80 (Sigma-Aldrich), in an orbital shaker, at 50ºC, for 24 h. *P. mobilis* was grown on DSMZ medium 718 modified by replacing yeast extract, trypticase and glucose with tryptone (1 g/L) and starch (5 g/L), and by increasing the concentration of $NH_4Cl$ to 4 g/L. *M. hassiacum* and *P. mobilis* chromosomal DNA were isolated according to (Belisle et al, 2009) and (Rainey et al, 1996) respectively.

*E. coli* BL-21 was the host for expression of the *gpgS* genes from *M. bovis* BCG, *M. smegmatis* 1102 and *M. hassiacum*, while *E. coli* DH5α was the host for expression of the *gpgS* gene from *P. mobilis*. The organism was grown in 1 L flasks in an orbital shaker at 37ºC, pH 7.0 in LB medium containing kanamycin at 30 μg/mL for the *E. coli* BL21 strain with the pET30a plasmid or ampicillin at 100 μg/mL *E. coli* DH5α strain with the pTRC99a plasmid. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added at a final concentration of 0.5 mM.

**Identification, sequence analysis, cloning and functional overexpression of the glucosyl-3-phosphoglycerate genes (*gpgS*) from *Mycobacterium bovis* BCG, *Mycobacterium smegmatis* and *Petrotoga mobilis***

To identify the *gpgS* gene in mycobacterial and *P. mobilis* genomes we used the amino acid sequences of the mannosyl-3-phosphoglycerate synthase (MpgS) from *Rubrobacter xylanophilus* and of the GpgS from *Persephonella marina* (GenPept accession numbers ACJ24528 and YP_002731348, respectively) in BLAST searches at the National Center for Biotechnology Information (NCBI, http://blast.ncbi.nlm.nih.gov/Blast.cgi) database. The clustalX2 (http://www.clustal.org) program was used for sequence alignments and the MEGA4 program (www.megasoftware.net) was used generate the phylogenetic trees.

The *gpgS* genes were amplified based on the available sequences from *M. smegmatis* strain mc$^2$155, *M. bovis* strain AF2122/97 and *P. mobilis* strain SJ95. NdeI restriction sites (underlined) were added to the forward primers for the amplification of the *M. smegmatis* gene (5'-ATTGATCAACATATGACAGCGATACCTCTGACCG-3') and for the *M. bovis* gene (5'-ATTGATCAACATATGACAGCATCGGAGCTGG-3'), while an NcoI site (underlined) was added in the case of *P. mobilis* (5'-AATCCATGGTGAAGGATAATATTTTAAAACG -3'). An EcoRI site (underlined)

was added to the reverse primer for the *M. smegmatis* gene (5'-ATAGAATTCCTGCGCGGACGCAGCGTGG-3') and a HindIII site (underlined) was added to the reverse primer for the *M. bovis* gene (5'-ATCAAGCTTGCGCGGCCGCATCACCTTC-3'). For the mycobacterial *gpgS* genes stop codons were removed to allow the translation of a C-terminal 6×His-tag encoded by the expression vector pET30a (Novagen), while for *P. mobilis* an HindIII site (underlined) was added but the stop codon was not removed (5'-AACAAGCTTTCATACCTTCTCCTCTTTTG-3') and the gene was cloned in pTRC99a producing a tagless recombinant protein. PCR were carried out with the AccuPrime Pfx DNA Polymerase (Invitrogen). DNA (200 ng) was denatured by 95ºC for 2 min followed by 30 cycles of 1 min denaturation step at 95ºC, 1 min annealing step at 60ºC, and 1.5 min extension step at 72ºC. The PCR products were cloned in the respective plasmids. The constructs were sequenced to confirm the identity of the insert (AGOWA, Berlin, Germany) and transformed into the respective *E. coli* strains, which were grown to the mid-exponential phase of growth ($OD_{610}$ nm = 0.8) at 37ºC. Expression was induced with 0.5 mM IPTG, and temperature was reduced to 30ºC for the last 8 h of growth.

**PCR amplification, cloning and expression of *gpgS* from *Mycobacterium hassiacum***

Since the *M. hassiacum* genome is not sequenced, the following degenerate oligonucleotides were designed, based on conserved regions of several mycobacterial *gpgS* genes available from public databases: forward primer (5'-TGA TCGTGCTGGAYTCSGG-3') and reverse primer (5'-CCSAGGTTGACCT GSGCGAT-3'). A 514 bp PCR fragment was amplified from the chromosomal DNA of *M. hassiacum* with these primers, using the AccuPrime GC rich DNA polymerase

(Invitrogen), in the same condition as before, but with 1 min extension. This fragment was subsequently used to design the reverse primer (5'-CGGTGTCGT CGGTGGAGCCG -3') and the forward primer (5'-GGCTACGGCGTGGA GATCGG-3') to be, respectively, combined with degenerate forward primer (5'-TCGGRGARACTCTGGGBGTG-3'), and with reverse primer (5'-GGCG CGAACGACCACAGCA-3'), both designed based on conserved regions of genes adjacent to mycobacterial *gpgS* genes. To amplify the complete sequence of *gpgS* gene from *M. hassiacum*, final 5' and 3' primers (5'-GCGCATATG ACACTGGTACCG-3') and (5'-ATAAGCTTGCCGCGCAGGGTG-3') were constructed with additional NdeI restriction site (underlined), and EcoRI restriction site (underlined), respectively. The stop codon was removed to allow the translation of a recombinant protein containing a C-terminal 6×His-tag. PCR amplification was carried out with AccuPrime GC-Rich DNA Polymerase (Invitrogen). The PCR product was purified from agarose gel (JET quick kit, Genomed), digested and cloned into pET30a (Novagen). The construct was sequenced (AGOWA, Berlin, Germany) and transformed into *E. coli* BL-21. Expression procedures were the same as for the other mycobacterial *gpgS* genes.

**Preparation of cell-free extracts**

Cells containing the recombinant GpgS from *P. mobilis* were harvested by centrifugation (7000 × *g*, 10 min, 4ºC) during the late exponential growth phase and the pellet suspended in 20 mM Tris-HCl (Merck) pH 7.5 containing 5 mM $MgCl_2$ (Merck), 10 µg/mL DNAse I and a cocktail of protease inhibitors (Roche). All the cells containing recombinant mycobacterial GpgS were harvested as above and the pellet suspended in 20 mM sodium phosphate buffer at pH 7.4 with 0.5 M NaCl and 20 mM imidazole. A protease inhibitor cocktail (Roche), 10 µg/mL DNAse I and 5

mM $MgCl_2$ were also added to the suspension. All cells were disrupted twice in a French-press followed by centrifugation (15000 × *g*, 30 min, at 4ºC).

**Enzyme assays**

The activity of the recombinant mycobacterial GpgSs and recombinant GpgS from *P. mobilis* in cellular extracts and during purification, was detected after 10 min, at 37ºC for the *M. bovis* and *M. smegmatis,* at 50ºC for the *M. hassiacum* and at 60ºC for the *P. mobilis* GpgSs, in reaction mixtures (50 µL) containing 10-25 µL of cell-free extract, 5.0 mM (each) of 3-phosphoglycerate (3-PGA) and UDP-glucose, and 20 mM $MgCl_2$ (Merck) in 50 mM BTP (Molekula), pH 7.5. After the initial incubation 2 U of alkaline phosphatase (Sigma-Aldrich) were added and the reactions were further incubated for 10 min at 37ºC.

The synthesis of GPG was monitored from the presence of GG by thin-layer chromatography (TLC) with a solvent system composed by chloroform/methanol/acetic acid/water (30:50:8:4, v/v) (Costa et al, 2007).

**Purification of the recombinant GpgSs**

The His-tagged recombinant mycobacterial GpgSs were purified in a prepacked Ni-Sepharose high-performance column (His-Prep FF 16/10) equilibrated with 20 mM sodium phosphate, pH 7.4, 0.5 M NaCl, and 20 mM imidazole. Elution was carried out with 500 mM imidazole and the purity of the fractions was determined by SDS-PAGE. The purest active fractions were pooled, concentrated and equilibrated with 50 mM BTP, pH 7.5, and 200 mM NaCl by ultracentrifugation in 30-kDa cutoff centricons (Amicon) and finally loaded onto a Superdex 200 column equilibrated with

the same buffer. Active fractions were concentrated, equilibrated with 50 mM BTP, pH 7.5, and the purity was determined by SDS-PAGE.

The recombinant GpgS from *P. mobilis* was purified with two consecutive DEAE-Sepharose columns, equilibrated with Tris-HCl (pH7.4). Elution was carried out with linear NaCl gradients of 0.0 to 1.0 M in the first column and 0.0 to 0.5 M in the second, for one hour. Two consecutive Q-Sepharose columns were then used in the conditions described for the previous columns. Finally, two Superdex 200 columns, equilibrated with 20 mM Tris-HCl and 200 mM NaCl (pH 7.4), were used to obtain pure recombinant GpgS. Active fractions were concentrated and equilibrated in 20 mM Tris-HCl (pH 7.4). In each step the purity of the fractions was determined by SDS-PAGE and only the purest active fractions were pooled, concentrated, equilibrated with the appropriate buffer and loaded in the next column.

Protein content of the samples was determined using the Bradford assay (Bradford, 1976).

**Characterization of the recombinant GpgSs**

The substrate specificity of the recombinant GpgSs was examined by TLC as described previously (Costa et al, 2007) using ADP-glucose, GDP-glucose, TDP-glucose, UDP-glucose, glucose-1-phosphate, ADP-mannose, GDP-mannose, UDP-mannose, mannose-1-phosphate and mannose-6-phosphate as sugar donors and 2-phosphoglycerate, 3-phosphoglycerate, D-glycerate, L-glycerate, DL-glycolate, DL-lactate as sugar acceptors (all from Sigma-Aldrich) at 37ºC for the *M. bovis* and *M. smegmatis*, at 50ºC for the *M. hassiacum* and at 60ºC for the *P. mobilis* GpgSs.

Temperature, pH profile, the effect of cations and stability of GpgSs were determined by the addition of either of the GpgSs from *M. smegmatis* (0.5 μg), from *P. mobilis* (0.5 μg), from *M. hassiacum* (1 μg) or from *M. bovis* (1.25 μg) to 50 μL reaction

mixtures containing the appropriate buffer and cation, 5 mM UDP-glucose and 2.5 mM 3-PGA, and the reactions were stopped at different times by cooling on ethanol–ice. The GpgSs were inactivated by the addition of 5 μL of 5 N HCl and neutralized by 5 μL of 5 N NaOH, followed by the addition of 240 μL of 50 mM Tris-HCl, pH 7.0, with 2.0 mM $MgCl_2$. The GPG formed was quantified by specific dephosphorylation with the *Persephonella marina* glucosyl-3-phosphoglycerate phosphatase (GenBank accession number EU277657) as previously described (Costa et al, 2007), using the Ames reaction (Ames, 1966).

The temperature profile for the *M. bovis* and for the *M. smegmatis* GpgSs was studied between 20 and 55ºC in 50 mM BTP at pH 8.0, with 20 mM $MgCl_2$. For the *M. hassiacum* GpgS it was determined between 30 and 100ºC in 50 mM BTP at pH 7.5, with 20 mM $MgCl_2$. For the *P. mobilis* GpgS the activity was studied between 30 and 75ºC in 20 mM Tris-HCl buffer pH 7.0, with 5 mM of $CoCl_2$. The effect of pH on the activity was determined in 20 mM acetate buffer at pH 4.0-5.5, 20 mM MES at pH 5.5-6.5 and 20 mM Tris-HCl or 20 mM BTP at pH 7.0-9.0, at 37ºC for the *M. bovis* and *M. smegmatis*, at 50ºC for the *M. hassiacum* and at 60ºC for the *P. mobilis* GpgSs.

The effect of cations was examined by incubating reaction mixtures, with the chloride salts of $Mg^{2+}$, $Mn^{2+}$, $Co^{2+}$, $Ca^{2+}$, $Fe^{2+}$, $Zn^{2+}$, $Sr^{2+}$, $Ni^{2+}$, $Cu^{2+}$ and with EDTA.

The kinetic parameters for the GpgSs were determined by measuring the amount of phosphate released (Ames, 1966). All experiments were performed in triplicate.

The behavior in solution of the recombinant GpgSs were estimated by gel filtration on a Superdex 200 column using molecular mass standards of albumin (67 kDa), aldolase (158 kDa), catalase (232 kDa) and ferritin (440 kDa). Blue Dextran 2000 (Amersham) was used to determine the void volume.

## Results

### Identification of the mycobacterial and *Petrotoga mobilis* GpgSs, and sequence analysis

BLAST analysis with the amino acid sequence of the *Rubrobacter xylanophilus* bifunctional MpgS/GpgS (accession number ACJ24528) and *Persephonella marina* GpgS (accession number YP_002731348), within mycobacterial databases resulted in the identification of a GpgS with 40-34% and 28-24% amino acid identity, respectively. This gene was annotated as a conserved hypothetical protein considered essential for the growth of *M. tuberculosis* (Sassetti et al, 2003). Blast analysis with the same sequences against the *Petrotoga mobilis* database also resulted in the identification of a GpgS with 63% and 26% amino acid identity respectively.

The *gpgS* genes from *M. bovis* BCG, *M. smegmatis*, *M. hassiacum* and *P. mobilis* contained, respectively, 975, 912, 945 and 978 bp coding for polypeptides with 324, 303, 314 and 325 amino acids with calculated molecular masses of 34.4, 32.3, 33.6 and 37.1 kDa and calculated isoelectric points of 4.9, 5.1, 4.9 and 5.6 (Table 2.1). The *M. bovis* BCG *gpgS* gene was identical to the homologues found in all species within the *M. tuberculosis* complex with genomes sequenced, namely *M. bovis*, *M. tuberculosis*, *M. africanum* and *M. microti*. The genomic context of the *gpgS* gene region is similar in all of the available mycobacterial genomes (Fig. 2.1).

A phylogenetic tree constructed based on several GpgS homologues available in the genomic databases shows two clusters; one being composed primarily of actinobacterial GpgSs with *P. mobilis* and two other unrelated species, while the remaining GpgS sequences form a different cluster that includes many unrelated bacteria and archaea (Fig. 2.2).

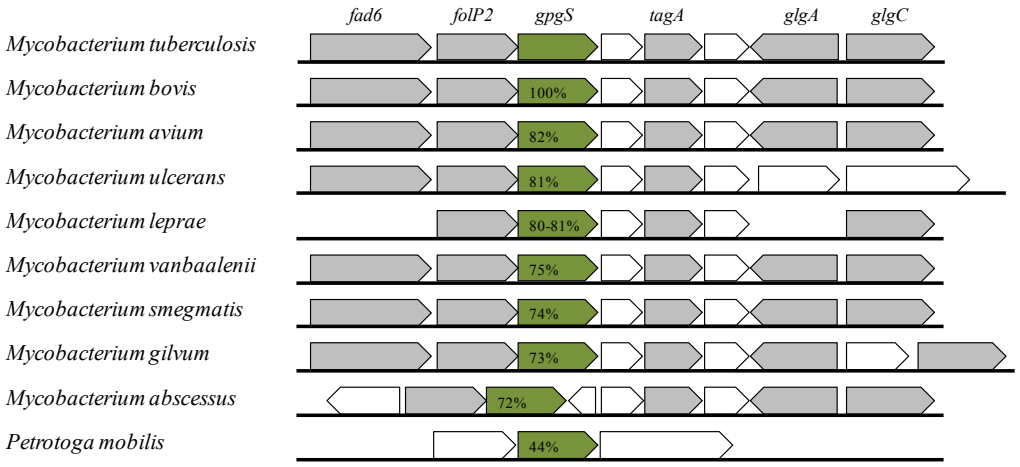



**Figure 2.1:** Genetic arrangement of the *gpgS* gene region. *fad6* - Probable fatty-acid-CoA ligase. *folP2* – Probable dihydropteroate synthase 2. *tagA* - Probable DNA-3-methyladenine glycosylase I. *glgA* – Glycogen synthase. *glgC* - ADP-glucose pyrophosphorylase.

## Expression and properties of the recombinant GpgSs

Expression of the *M. bovis* BCG, *M. smegmatis* and *M. hassiacum gpgS* genes in *E. coli* resulted in the high level production of recombinant His-tagged protein while expression of the *P. mobilis gpgS* gene resulted in high production of a tagless recombinant GpgS protein. Gel filtration experiments indicated that the recombinant mycobacterial His-tagged All the GpgSs behaved as dimeric proteins in solution (Table 2.1). The mycobacterial recombinant GpgSs used UDP-glucose, ADP-glucose and GDP-glucose as donor substrates with decreasing affinity and 3-PGA as acceptor (Table 2.1) while the *P. mobilis* GpgS could only use UDP-glucose and ADP-glucose. However, free ADP severely inhibited the *P. mobilis* GpgS even at very low concentrations rendering the kinetic parameters determination for this substrate impossible.

**Table 2.1:** Properties of the recombinant *M. bovis*, *M. smegmatis*, *M. hassiacum* and *P. mobilis* GpgSs

| | Protein size | Protein IP | Protein MW | Behavior in solution | Preferred cation | Optimal temperature | Optimal pH | Preferred substrates | $K_m$ (mM) | $V_{max}$ µmol/min/mg |
|---|---|---|---|---|---|---|---|---|---|---|
| *M. bovis* | 324 | 4.9 | 34.38 kDa | Dimer | $Mg^{2+}$ | 45ºC | pH 8.0 | UDP-glucose> | 2.43 ± 0.21 | 1.56 ± 1.09 |
| | | | | | | | | ADP-glucose> | 6.63 ± 0.76 | 3.56 ± 0.50 |
| | | | | | | | | GDP-glucose | 27.70 ± 4.62 | 1.69 ± 0.16 |
| | | | | | | | | 3-PGA | 0.04 ± 0.01 | 1.42 ± 0.03 |
| *M. smegmatis* | 303 | 5.1 | 32.31 kDa | Dimer | $Mg^{2+}$ | 45ºC | pH 8.5 | UDP-glucose> | 1.28 ± 0.02 | 27.51 ± 0.11 |
| | | | | | | | | ADP-glucose> | 6.08 ± 0.30 | 26.23 ± 0.72 |
| | | | | | | | | GDP-glucose | 32.04 ± 2.94 | 2.45 ± 0.11 |
| | | | | | | | | 3-PGA | 0.24 ± 0.04 | 25.91 ± 1.34 |
| *M. hassiacum* | 314 | 4.9 | 33.59 kDa | Dimer | $Mg^{2+}$>$Mn^{2+}$ | 60ºC | pH 7.5 | UDP-glucose> | 1.04±0.12 | 35.49 ± 1.40 |
| | | | | | | | | ADP-glucose> | 5.46±0.47 | 26.31 ± 1.85 |
| | | | | | | | | GDP-glucose | 25.55 ± 1.48 | 7.45 ± 0.26 |
| | | | | | | | | 3-PGA | 0.69 ± 0.06 | 42.19 ± 0.5 |
| *P. mobilis* | 325 | 5.6 | 37.14 kDa | Dimer | $Co^{2+}$>$Mn^{2+}$>$Ni^{2+}$ | 70ºC | pH 7.0 | UDP-glucose> | 1.00 ± 0.1 | 44.0 ± 1.1 |
| | | | | | | | | ADP-glucose | - | - |
| | | | | | | | | 3-PGA | 0.7 ± 0.1 | 43.9 ± 1.7 |

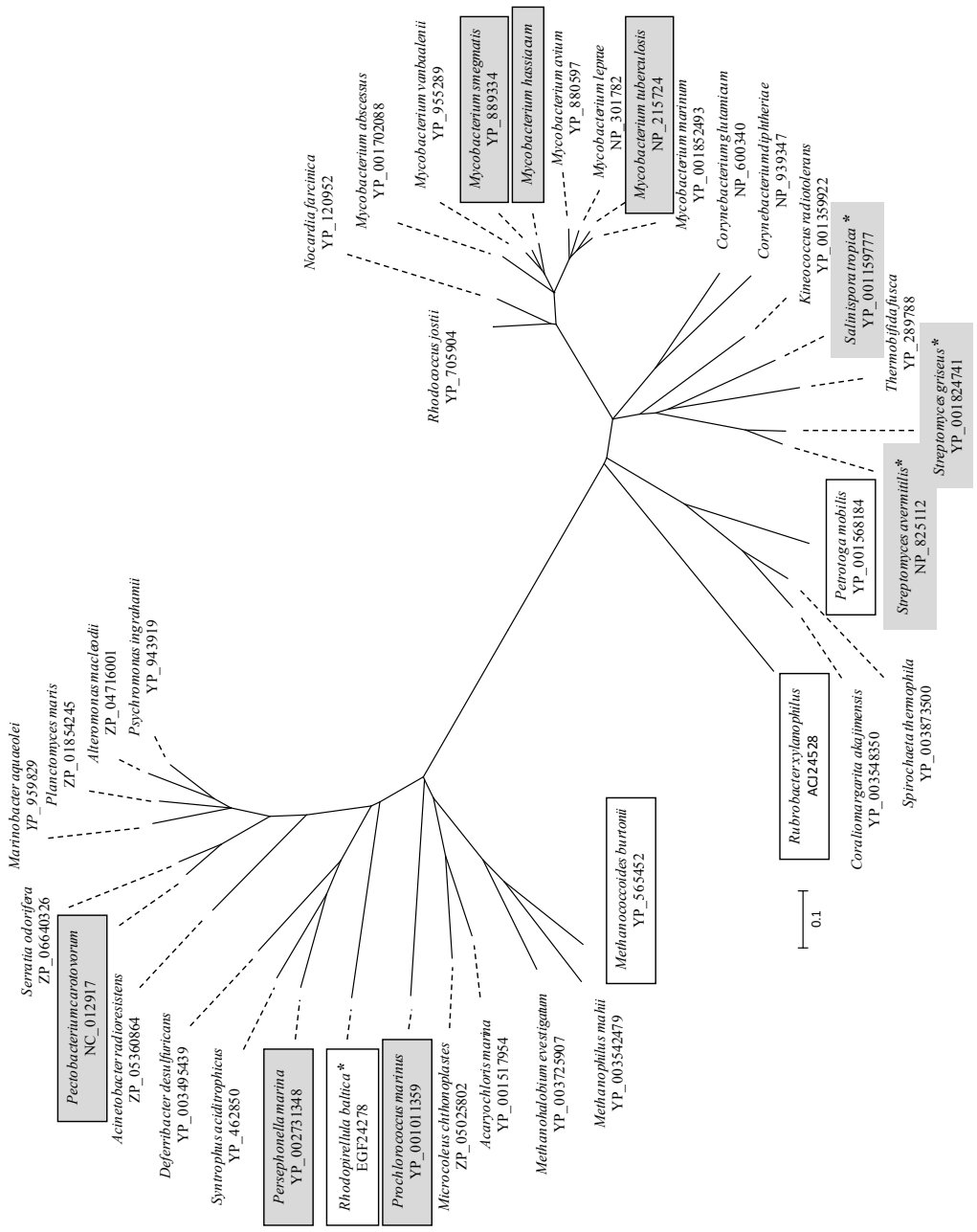



**Figure 2.2:** Unrooted phylogenetic tree based on available amino acid sequences of identified and putative GpgSs. Organisms where glucosylglycerate has been detected are shaded in gray.

Species with enzymes which have been demonstrated to synthesize GPG are surrounded by boxes. Organisms with an (*) have had their enzymes demonstrated to synthesize GPG or accumulation of GG has been observed by our group (our group unpublished results). Peptide accession numbers are below the organisms names, and have been retrieved from NCBI database. Scale bar, 0.1 changes per site.

It was also noteworthy that 3-PGA, at concentrations above 5 mM, was progressively inhibitory to the *P. mobilis* GpgS. D-glycerate could only be the acceptor with ADP-glucose and at a very low rate for *M. smegmatis* and *M. bovis* GpgS. All the GpgSs exhibited Michaelis–Menten kinetics at the respective organism optimal temperature (Table 2.1). The activity of the *M. bovis* and *M. smegmatis* GpgSs was undetectable below 20 and above 55ºC, with a maximum at 45ºC (Fig. 2.3 A). While the activity of the *M. hassiacum* and *P. mobilis* GpgSs was undetectable below 30 and above 80ºC, with maximum at 60 and 70ºC respectively (Fig. 2.3 A). The enzymes were active between pH 6.0 and 10.0, but the activities were maximal between pH 7 and 8.5 (Fig. 2.3 B and C). The recombinant GpgSs were dependent on divalent cations for activity. Of the cations tested, only $Mg^{2+}$ stimulated *M. smegmatis* and *M. bovis* GpgSs activity, while $Mg^{2+}$ and $Mn^{2+}$ stimulated the *M. hassiacum* GpgS and $Co^{2+}$, $Mn^{2+}$ and $Ni^{2+}$ stimulated the *P. mobilis* GpgS with decreasing activation rates (Table 2.1). The maximum activation of the GpgSs was observed between 20 and 75 mM of $MgCl_2$ for *M. bovis*, 20 mM $MgCl_2$ for *M. smegmatis*, 10 to 20 mM of $MgCl_2$ for *M. hassiacum* and 5 mM $CoCl_2$ for *P. mobilis*.

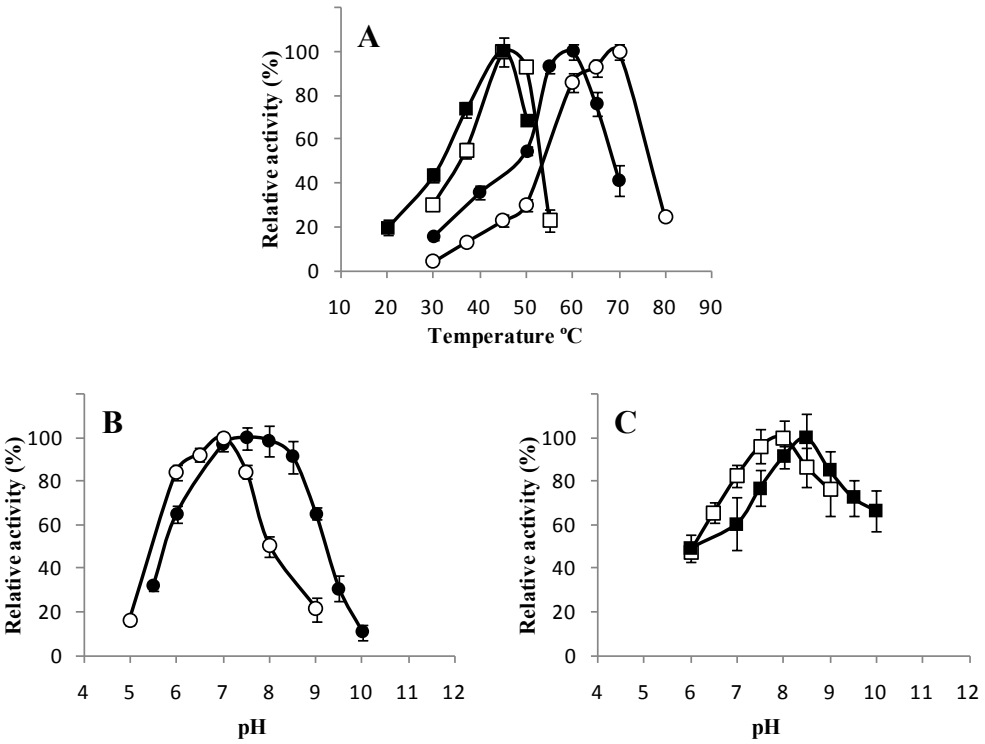



**Figure 2.3:** Temperature and pH profiles of the recombinant GpgSs from *M. bovis* (□), *M. smegmatis* (■), *M. hassiacum* (●) and *P. mobilis* (o). **A** – Temperature profiles of all GpgSs. **B** – pH profile of the *M. hassiacum* (●) and *P. mobilis* (o) GpgSs. **C** - pH profile of the *M. bovis* (□), *M. smegmatis* (■) GpgSs.

## Discussion

Free GG and diglucosylglycerate were reported in mycobacteria and proposed to be initial intermediates in the synthesis MGLP (Kamisango et al, 1987), a structurally complex polysaccharide implicated in fatty acid synthesis modulation and transport (Jackson & Brennan, 2009). The structural complexity of the methyl glucose

lipopolysaccharide (MGLP) (Fig. 2.4) suggests that several enzymes are involved in the biosynthesis of this molecule. An essential methyltransferase responsible for the 6-*O*-methylations of the glucose units of the polysaccharide (Rv3030), and two α-(1→4)-glycosyltransferases (Rv1212c and Rv3032) responsible for the synthesis of the α-(1→4) chain of the polysaccharide have already been identified and experimentally confirmed (Stadthagen et al, 2007).

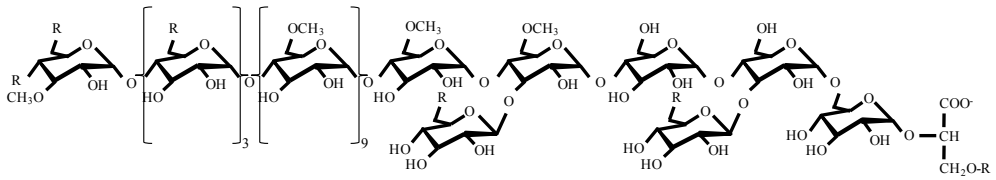


**Figure 2.4:** Structure of the mycobacterial MGLP.

*Petrotoga miotherma* was found to accumulate a compound very similar to the diglucosylglycerate found in mycobacteria, mannosylglucosylglycerate (MGG) (Jorge et al, 2007), whose precursors are GDP-mannose and GPG (Fernandes et al, 2010). This prompted us to identify the mycobacterial GpgS but also the *P. mobilis* GpgS whose genome was also available.

The *gpgS* gene has close homologues in all mycobacteria with available genome and in many actinobacteria, but with the exception of *Rubrobacter xylanophilus* where it encodes an enzyme with mannosyl-3-phosphoglycerate synthase (MpgS) and GpgS activities *in vitro* (Empadinhas et al, 2011), their function was yet to be experimentally confirmed. In this study, the *gpgS* genes from three mycobacteria; the fast-growing *M. smegmatis*, the thermophilic and fast-growing *M. hassiacum* and the slow-growing *M. bovis* BCG, and also from *P. mobilis* were expressed and the recombinant proteins

characterized. Although *P. mobilis* is phylogeneticaly unrelated to actinobacteria it harbours an actinobacterial-type GpgS (Fig. 2.3) most likely due to a lateral gene transfer event.

The *M. bovis* GpgS is 100% identical to Rv1208 from *M. tuberculosis* H37Rv and to the homologues from other strains of the species within the *M. tuberculosis* complex. Similarly, all the *M. smegmatis* strains share a 100% identical GpgS demonstrating that all these species and strains can synthesize GPG.

The three mycobacterial GpgSs characterized in this work presented an unusual sugar-nucleotide flexibility that resembles the *Persephonella marina* GpgS that was also capable to use three different sugar- nucleotide donors (Costa et al, 2007), but not the one from *Methanococcoides burtonii* that was only capable of using GDP-glucose (Costa et al, 2006). The *P. mobilis* GpgS also showed some sugar-nucleotide flexibility, however the very high inhibitory effect of free ADP, that prevented kinetic parameters determination for this substrate, is a characteristic so far unreported in other characterized GpgSs.

All the four GpgSs were dependent of divalent cations for activity in agreement to what was reported in the GpgSs from *M. burtonii* and *P. marina* and in the MpgS/GpgS from *R. xylanophilus* (Costa et al, 2007; Costa et al, 2006; Empadinhas et al, 2011). Interestingly, while the mycobacterial GpgSs were preferably activated by $Mg^{2+}$, the *P. mobilis* enzyme, like the *P. marina* and the *M. burtonii* GpgSs preferred low concentrations of $Co^{2+}$ and $Mn^{2+}$ (Costa et al, 2007; Costa et al, 2006). Temperature profile of the enzymes was in accordance with the respective organism optimum growth temperature, with the *M. smegmatis* and *M. bovis* GpgSs having the lowest optimal temperatures and the *M. hassiacum* and *P. mobilis* GpgSs with the highest optimal temperatures.

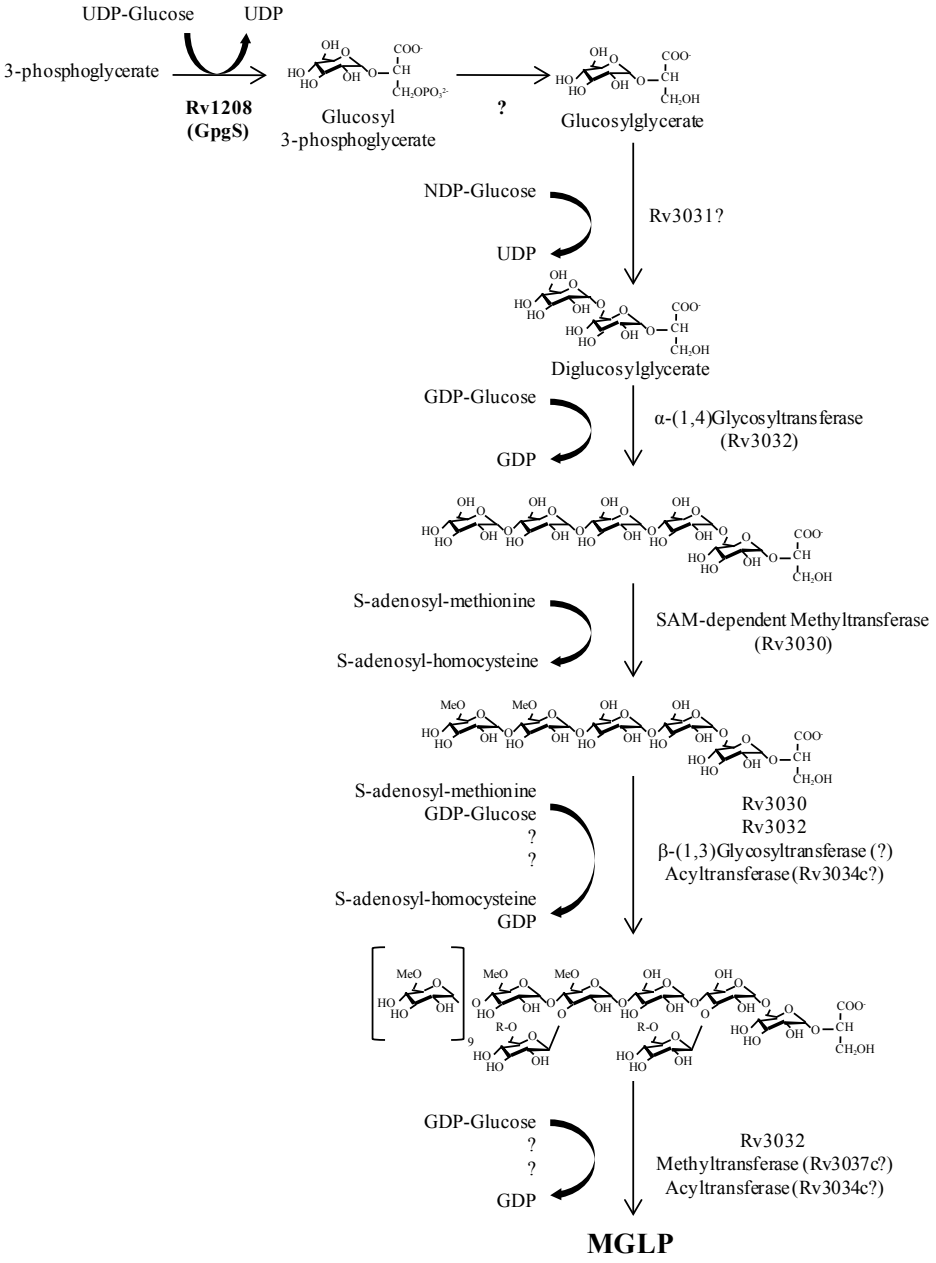



**Figure 2.5:** Proposed pathway for the synthesis of MGLP. Partially adapted from (Jackson & Brennan, 2009)

Numerous structural studies have been performed on the mycobacterial polymethylated polysaccharides, but many of the genes and enzymes involved in the biosynthetic pathways remain unknown (Gray & Ballou, 1971; Kamisango et al, 1987; Tuffal et al, 1998; Weisman & Ballou, 1984a; Weisman & Ballou, 1984b). However, it was established that GG was one of the first precursors of MGLP (Kamisango et al, 1987).

This work established that the phosphorylated form of GG, glucosyl-3-phosphoglycerate (GPG), is in fact the first precursor in the synthesis of this polysaccharide and that the *M. tuberculosis* Rv1208 gene encoding a GpgS, suggested to be essential for the growth of this organism (Sassetti et al., 2003), is responsible for this step. GPG will subsequently be dephosphorylated, before the addition of the first glucose unit of the α-(1→4)-chain of the polysaccharide (Fig. 2.5).

It was suggested that the glycerate group at the reducing end of the MGLP is a target for esterification by octanoate, which could help to anchor the polysaccharide to the cytoplasmic membrane, during step-wise addition of glucose units to the MGLP backbone (Jackson & Brennan, 2009; Smith & Ballou, 1973).

The ubiquitous presence of MGLP in mycobacteria suggests a fundamental role in these organisms physiology. With the identification of other genes involved in MGLP pathway that have been deemed essential for mycobacterial growth one may assume that this structure is in fact vital for mycobacteria (Sassetti et al, 2003; Stadthagen et al, 2007). It is reasonable to think that the lack or deficient production of this compound could interfere with fatty acid synthesis, cell wall assembly and possibly lead to cell death.

However GpgS deficient mutants constructed in *M. smegmatis* showed no lethal fenotype, despite having their MLGP content reduced in 80%, probably reflecting the ability of MMP to replace MGLP in the organisms that are able to synthesize both (Kaur et al, 2009b). Nevertheless, the presence of MGLP in GpgS knock out mutants raises the possibility of the presence of another yet unknown GPG synthesizing

enzyme in *M. smegmatis* with no sequence similarity to all the known GpgSs (Kaur et al, 2009b).

Nevertheless, since no structural analysis was performed on the MGLPs synthesized by the mutants, this also raises the possibility that GG could be absent from the reducing end of the polysaccharides or that another compound could have replaced the GG in the MGLP structure.

Finally, the reported essentiality the *gpgS* gene for *M. tuberculosis* growth (Sassetti et al, 2003), makes this protein an interesting target for potential antimycobacterial drug development.

## Work contributions and Acknowledgements

Nuno Empadinhas and Luciana Albuquerque identified, expressed and characterized the *M. bovis* and *M. smegmatis* GpgSs and Vítor Mendes performed the phylogenetic studies. Vítor Mendes cloned, expressed, purified and characterized the *P. mobilis* GpgS. Susana Alarico amplified, expressed and characterized the *M. hassiacum* GpgS. The authors gratefully acknowledge Helena Santos (ITQB, Oeiras, Portugal) for supplying pure glucosylglycerate. This work was partially supported by Fundação para a Ciência e a Tecnologia (FCT), Portugal/FEDER project POCI/BIA-MIC/56511/2004.

# Chapter 3

## The mycobacterial glucosyl-3-phosphoglycerate phosphatase

**This chapter will be the subject of a publication:**

**New insights into MGLPs biosynthesis arising from Rv2419c, the missing mycobacterial glucosyl-3-phosphoglycerate phosphatase**. In preparation

# Contents

## Abstract

The gene for the first step in the MGLP pathway (Rv1208 in *Mycobacterium tuberculosis*) has been functionally characterized as a glucosyl-3-phosphoglycerate synthase (GpgS), but a typical glucosyl-3-phosphoglycerate phosphatase (GpgP EC 3.1.3.70) required to dephosphorylate glucosyl-3-phosphoglycerate (GPG) to glucosylglycerate (GG), the primer for MGLP biosynthesis, is absent from mycobacterial genomes. To gain further insights into the early steps of MGLP biosynthesis we probed *Mycobacterium hassiacum* and *Mycobacterium vanbaalenii* extracts for GpgP activity, the proposed second step in the pathway. Herein, we report the identification and characterization of the mycobacterial GpgP, which is highly specific for glucosyl-3-phosphoglycerate but unrelated to any of the GpgPs characterized to date. In *M. tuberculosis* this enzyme is encoded by gene Rv2419c, incorrectly annotated as a putative phosphoglycerate mutase (PGM). The recombinant GpgP from *M. tuberculosis* and the native *M. hassiacum* GpgP are active between 20 and 60ºC with peak activity at 45 and 50ºC respectively and from pH 4.0 to 10.0 with optimum at pH 7.0. The *M. tuberculosis* GpgP had a $K_m$ for GPG at 37ºC of 0.35 ± 0.03 mM with $V_{max}$ of 67.19 ± 1.39 μmol/min.mg protein. The native GpgP from *M. hassiacum* had a $K_m$ for GPG at 50ºC of 0.85 ± 0.08 mM. The GpgP is a dimer in solution and unlike the non-mycobacterial GpgPs, the activity of the Rv2419c product is independent of cations.

The identification of this new function in the genome of *M. tuberculosis* significantly advances our understanding on how mycobacteria synthesize the unique MGLP.

## Introduction

Mycobacteria synthesize two, hexose based, polymethylated polysaccharides in which many of the hexose units are *O*-methylated. One is the methyl mannose polysaccharide (MMP) in which the methylated mannose units are 3-*O*-methylated and the other the methyl glucose lipopolysaccharide (MGLP) in which the methylated glucose units are 6-*O*-methylated (Jackson & Brennan, 2009).

Structurally the mycobacterial MGLP contains a backbone formed by α–(1→4)-linked glucose and methyl glucose units, leading to a helical conformation characteristic of amylose (Peterson & Bloch, 1977) with a glucosylglycerate (GG) unit attached to its reducing end (Tuffal et al, 1998). It further contains two branches composed by β–(1-3)-linked glucoses at glucose 1 and 3 of the main α–(1→4)-chain (Tuffal et al, 1998) and also presents a variable degree of acylations and acyl groups (Gray & Ballou, 1972). Since some of the genes putatively attributed to MGLP synthesis were considered essential for mycobacterial growth (Sassetti et al, 2003), this structure is likely essential for mycobacterial metabolism.

GG was originally detected in *Mycobacterium smegmatis* cell extracts (Kamisango et al, 1987). It has been detected, since, in many organisms including methanogenic archaea, proteobacteria, hyperthermophilic bacteria and cyanobacteria were it functions as a compatible solute during osmotic stress or as a precursor to larger molecules such as glycolipids and polysaccharides (Costa et al, 2007; Goude et al, 2004; Kamisango et al, 1987; Klahn et al, 2010; Pommier & Michel, 1986; Robertson et al, 1992). In mycobacteria GG functions as a precursor for the MGLP (Kamisango et al, 1987).

Two pathways for the synthesis of GG have been elucidated (Costa et al, 2006; Fernandes et al, 2007). In one reaction a glucosylglycerate synthase directly synthesizes GG from NDP-glucose and D-glycerate (Fernandes et al, 2007). A more widely distributed reaction involves the synthesis of a phosphorylated intermediate,

glucosyl-3-phosphoglycerate (GPG) from NDP-glucose and 3-PGA by a glucosyl-3-phosphoglycerate synthase (GpgS) with subsequent dephosphorylation by a specific phosphatase designated glucosyl-3-phosphoglycerate phosphatase (GpgP) (Costa et al, 2006).

Our group previously identified, characterized and determined the structure of the *Mycobacterium tuberculosis* GpgS (Empadinhas et al, 2008; Pereira et al, 2008), the enzyme for the first step in the synthesis of MGLP (Kaur et al, 2009b). The genes responsible for the main α–(1→4) chain extension and methylation of glucoses have also been identified (Stadthagen et al, 2007); however, the enzymes responsible for second are third steps remained elusive. The likely essentiality of MGLP for mycobacterial growth prompted us to elucidate the early biosynthetic steps for this polysaccharide.

We have detected high GpgP activity, the enzyme responsible for the second step in MGLP synthesis in *Mycobacterium hassiacum* and *Mycobacterium vanbaalenii* cell extracts. Since no relevant homologues of known GpgPs were found in mycobacterial genomes, we purified the native GpgP from these organisms and identified the corresponding gene. We then cloned and expressed the *gpgP* gene from *M. tuberculosis* in *E. coli*, purified the recombinant bioactive enzyme and determined its biochemical and kinetic properties.

## Methods

### Bacterial strains and growth conditions

*M. vanbaalenii* (DSM 7251) and *M. hassiacum* (DSM 44199) was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (Germany).

For enzymatic assays and purification of the native glucosyl phosphoglycerate phosphatase (GpgP), *M. vanbaalenii* was grown on a glycerol-based medium (Brennan & Ballou, 1967) in an orbital shaker, at 30ºC for 24 h. The medium contained per liter: 20 g glycerol (Merck), 5 g casamino acids (Difco), 1 g fumaric acid (Sigma-Aldrich), 1 g di-potassium hydrogen phosphate (Merck), 0.3 g magnesium sulfate (Molekula), 0.02 g iron sulfate and 2 g tween 80 (Sigma-Aldrich) at pH 7.0. *M. hassiacum* was grown on a modified GFPH medium (DSM 553 medium) at pH 7.2 containing per liter: 10 g glucose (Merck), 5 g tryptone (Difco), 5 g yeast extract (Difco), 5 g beef extract (Merck), 0.74 g calcium chloride (Merck) and 2 g tween 80 (Sigma-Aldrich), in an orbital shaker, at 50ºC, for 24 h.

**Preparation of cell-free extracts**

Cells were harvested by centrifugation (7000 × *g*, 10 min, 4ºC) during the late exponential growth phase and the pellet resuspended in 20 mM Bis-tris propane buffer (BTP) (Molekula) at pH 7.5, containing 5 mM $MgCl_2$ (Merck), 10 µg/mL DNAse I and a cocktail of protease inhibitors (Roche). Cells were disrupted twice in a French-press at 30000 PSI, followed by centrifugation (15000 × *g*, 30 min, at 4ºC).

**Enzyme assays**

The activity of the native GpgP from *M. vanbaalenii* and *M. hassiacum* and of the recombinant GpgP from *M. tuberculosis* in cell extracts and during purification, was detected after 10 min, at 30ºC, 50ºC and 37ºC, respectively, in reaction mixtures (25 µL) containing 10-15 µL of cell-free extract, 2 mM of glucosyl-3-phosphoglycerate (GPG), and 2 mM $MgCl_2$ (Merck) in 20 mM BTP (Molekula) at pH 7.0. The

dephosphorylation of GPG was monitored by thin-layer chromatography (TLC) with a solvent system composed by chloroform/methanol/acetic acid/ water (30:50:8:4, v/v). Glucosylglycerate (GG) and GPG standards were used for comparative purposes. Protein concentration was determined by the Bradford method (Bradford, 1976).

**Purification of the native GpgPs from *Mycobacterium hassiacum* and *Mycobacterium vanbaalenii***

To purify the native GpgP, 18 L of *M. hassiacum* and *M. vanbaalenii* were grown in the conditions described above. The GpgP was purified from cell extracts were purified using a DEAE-Sepharose equilibrated with 20 mM BTP at pH 7.5. Elution was carried out with a linear gradient of 0 to 1 M NaCl, and the GpgP activity in the fractions was determined by TLC as described above. The fractions containing GpgP activity were pooled, concentrated by centrifugation, in 30 kDa cutoff centricons (Amicon), loaded into a DEAE-Sepharose, equilibrated with 20 mM BTP at pH 6.5, and eluted by a linear gradient of NaCl of 0 to 0.7 M. Fractions were selected and pooled as above and loaded into a Q-Sepharose column equilibrated with 20 mM BTP, pH 7.5, and eluted by a linear gradient of NaCl of 0 to 0.7 M. A second Q-Sepharose column was used and equilibrated with 20 mM BTP at pH 6.5 and a linear gradient of 0 to 0.5 M was used. Two resource-Q columns were subsequently used; the first was equilibrated with 20 mM BTP, pH 7.5 and the second with 20 mM BTP, pH 6.5. A Superdex 200 column equilibrated with BTP 20 mM, NaCl 200 mM at pH 7.5 was then used. Two additional resource-Q columns and a final Superdex 200 column equilibrated with the above buffers were used. GpgP activity was assessed by TLC as previously described and the purity of the fractions was determined by SDS-PAGE. Each fraction (50 μL) was stored at -20ºC for protein identification by peptide mass fingerprinting. The two most active *M. hassiacum* fractions were pooled and selected

for kinetic assays. Protein content of the samples was determined by the Bradford assay (Bradford, 1976).

**Characterization of the native GpgP from *Mycobacterium hassiacum***

The substrate specificity of the native GpgP from *M. hassiacum* was determined using 2-phosphoglycerate, 3-phosphoglycerate, glucose-1-phosphate, glucose-6-phosphate, glucose-1,6-diphosphate, mannose-1-phosphate, mannose-6-phosphate, ATP, GTP, TTP, UTP, ADP, GDP, UDP (all from Sigma-Aldrich), glucosyl-3-phosphoglycerate (GPG), mannosyl-3-phosphoglycerate (MPG), mannosylglucosyl-3-phosphoglycerate (MGPG), as possible substrates. MPG, GPG and MGPG were synthesized as described previously (Fernandes et al, 2010).

The reaction mixtures (25 µL) containing partially purified native GpgP (0.3 µg), 1 mM of substrate and 2.5 mM $MgCl_2$ in 20 mM BTP at pH 7.0, were incubated at 50ºC for 10 min. The products were visualized by TLC as described above except for 2-phosphoglycerate and 3-phosphoglycerate where the Ames method was used instead.

Temperature, pH profile and cation dependence were determined by the addition of 0.3 µg of partially purified GpgP to 50 µL reaction mixtures containing the appropriate buffer, 1 mM GPG and stopped at different times by cooling on ethanol-ice. Free phosphate was quantified using the Ames reaction (Ames, 1966).

The effect of cations was examined by incubating the reaction mixture with the chloride salts of $Mg^{2+}$, $Mn^{2+}$, $Co^{2+}$, $Cu^{2+}$ (2.5 to 50 mM), EDTA 5 and 10 mM or with no addition of salts, at 37ºC.

The temperature profile was determined between 20 and 60ºC in 20 mM BTP, pH 7.0 with 2.5 mM $MgCl_2$. The effect of pH was determined at 50ºC in 20 mM acetate (pH 4.0 and 5.0), 20 mM BTP (pH 6.0 to 9.0) and 20 mM CAPS (pH 10.0), with 2.5 mM

$MgCl_2$. The kinetic parameters for the native GpgP were determined by measuring the amount of phosphate released, as described above. The $K_m$ value for GPG was determined at 50ºC, from Lineweaver–Burk plots. All experiments were performed in triplicate.

**Identification and sequence analyses of the mycobacterial GpgP**

Each of the *M. vanbaalenii* protein fractions (50 μL) obtained in the last purification step were loaded onto a large SDS-PAGE gel (20 cm × 20 cm) and 8 bands were selected for MS analysis (Proteomics Unit, Biocant, Cantanhede). Each of the peptide sequences obtained was analyzed with the BLAST tool from NCBI (http://blast.ncbi.nlm.nih.gov /Blast.cgi). From those analyses, the best *M. tuberculosis* candidate was selected for expression. The clustalX2 (http://www.clustal.org) program was used for sequence alignments and the MEGA4 program (www.megasoftware.net) was used to generate the phylogenetic trees.

**Cloning and functional overexpression, of the GpgP from *Mycobacterium tuberculosis***

The *gpgP* gene from *M. tuberculosis* was synthesized with the nucleotide sequence optimized for *E. coli* expression (GenScript, USA). This gene with optimized sequence was then cloned into pET30a (Novagen), using NdeI and HindIII restriction sites, and transformed into *E. coli* BL21 strain, which was used as host for expression. Recombinant *E. coli* was grown at 37ºC to mid-exponential phase of growth ($OD_{610}$ = 0.8) in a 5 L bioreactor with LB medium at pH 7.0, with continuous aeration and stirred at 180 rpm. Kanamycin (30 μg/mL) was added to the medium. IPTG (Zymo

Research) was added at a final concentration of 0.5 mM to induce gene expression and temperature was reduced to 20ºC. The cells were harvested 18 h later by centrifugation (8000×*g*, 10 min, 4ºC).

**Purification of the recombinant GpgP from *Mycobacterium tuberculosis***

The His-tagged recombinant GpgP from *M. tuberculosis* was purified in a prepacked Ni-Sepharose high-performance column (His-Prep FF 16/10) equilibrated with 20 mM sodium phosphate, pH 7.4, 0.5 M NaCl, and 20 mM imidazole. Elution was carried out with 500 mM imidazole and the purity of the fractions was determined by SDS-PAGE. The purest active fractions were pooled, concentrated by ultracentrifugation in 30 kDa cutoff, equilibrated with 20 mM BTP at pH 7.4 and loaded into a Q-Sepharose fast-flow column (Hi-Load FF 16/10), equilibrated with 20 mM BTP at pH 7.4 and eluted by a linear gradient of NaCl (0 to 1 M). The purity of the fractions was determined by SDS-PAGE and the purest active fractions were pooled, concentrated by ultracentrifugation in 30 kDa cutoff centricons (Amicon), equilibrated with 20 mM BTP at pH 7.4 with 200 mM NaCl, and loaded into a Superdex 200 fast-flow column equilibrated with the same buffer. After SDS-PAGE analysis the active pure fractions were concentrated and equilibrated with 50 mM BTP at pH 7 with 50 mM NaCl. Protein content of the samples was determined by the Bradford assay (Bradford, 1976).

**Characterization of the recombinant GpgP from *Mycobacterium tuberculosis***

The substrate specificity of the recombinant GpgP from *M. tuberculosis* was determined using 2-phosphoglycerate, 3-phosphoglycerate, 2,3-diphosphoglycerate,

glycerol-3-phosphate, phosphoenolpyruvate, α-glucose-1-phosphate, β-glucose-1-phosphate, glucose-6-phosphate, glucose-1,6-diphosphate, mannose-1-phosphate, mannose-6-phosphate, fructose-6-phosphate, trehalose-6-phosphate, ATP, UTP and ADP, (all from Sigma-Aldrich), glucosyl-3-phosphoglycerate (GPG), mannosyl-3-phosphoglycerate (MPG), mannosylglucosyl-3-phosphoglycerate (MGPG). MPG, GPG and MGPG were synthesized as described previously (Fernandes et al, 2010).

The reaction mixtures (25 μL) containing pure recombinant GpgP in known amounts, 3 mM of substrate and 2.5 mM $MgCl_2$ in 20 mM BTP at pH 7.0, were incubated at 37ºC for 10 min. The products were visualized by TLC as described above except for 2-phosphoglycerate, 3-phosphoglycerate, 2,3-diphosphoglycerate, glycerol-3-phosphate and phosphoenolpyruvate, where the Ames method was used instead.

Phosphoglycerate mutase activity was also assessed using the standard described method (Fig. 3.1) (Bergmeyer, 1974) and a positive control was performed with the recombinant phosphoglycerate mutase from human muscle (Affymetrix).

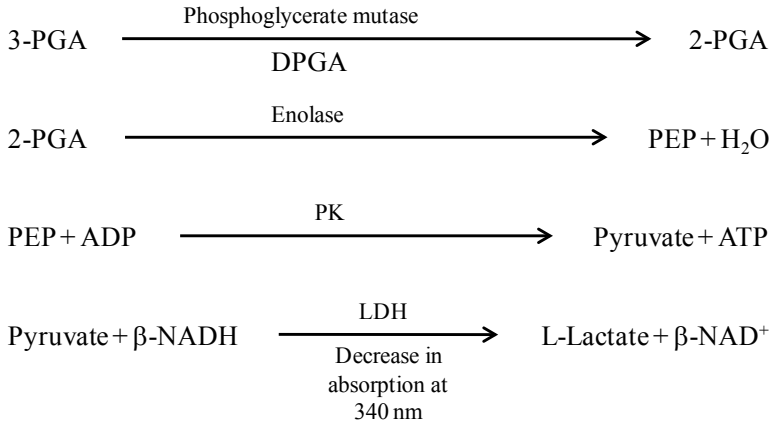



**Figure 3.1:** Schematic representation of the method used to detect phosphoglycerate mutase activity. Adapted from (Bergmeyer, 1974)

Temperature, pH profile and cation dependence were determined by the addition of known amounts of pure GpgP to 25 µL reaction mixtures containing the suitable buffer, 1 mM GPG and stopped at different times by cooling on ethanol-ice. Free phosphate was quantified using the Ames reaction (Ames, 1966).

The effect of cations was examined by incubating reaction mixture, with the chloride salts of $Mg^{2+}$, $Mn^{2+}$, $Co^{2+}$ (1 to 10 mM), EDTA 5 mM or with no addition of salts, at 37ºC. The temperature profile was determined between 20 and 60ºC in 20 mM BTP at pH 7.0. The effect of pH was determined at 37ºC in 20 mM acetate buffer (pH 4.0 and 5.0), 20 mM BTP (pH 6.0 to 9.0) and 20 mM CAPS buffer (pH 10.0). The kinetic parameters for the recombinant GpgP were determined by measuring the amount of phosphate released, as described above. The $K_m$ value for GPG was determined at 37ºC, from Lineweaver–Burk plots. All experiments were performed in triplicate.

## Results

### Identification of the mycobacterial GpgP and sequence analysis

MS analysis of the selected *Mycobacterium vanbaalenii* protein bands, obtained from SDS-PAGE, revealed a likely GpgP candidate with >99% confidence, in a band just under 25 kDa (Fig. 3.2). This *M. vanbaalenii* protein was annotated in the databases as a phosphoglycerate mutase and the gene was present in all sequenced mycobacterial species and in other actinobacteria (Fig. 3.3). The glucosyl-3-phosphoglycerate phosphatase gene (*gpgP*) from *M. vanbaalenii* contained 678 bp coding for a polypeptide with 225 amino acids with a calculated molecular mass of 24.2 kDa and a calculated isoelectric point of 5.6, while the *M. tuberculosis gpgP* gene contained 672

bp coding for a polypeptide with 223 amino acids with a calculated molecular mass of 24.2 kDa and a calculated isoelectric point of 6.1.

BLAST analyses with the *M. tuberculosis* GpgP sequence revealed highly similar homologues in all available mycobacterial genomes namely in *M. bovis* (100% amino acid identity), *M. leprae* (86%), *M. marinum* (84%), *M. intracellulare* (84%), *M. ulcerans* (83%), *M. avium* (83%), *M. parascrofulaceum* (82%), *M. kansasii* (80%), *M. smegmatis* (77%), *M. vanbaalenii* (77%), *M. abscessus* (75%) and *M. gilvum* (73%). Additional homologues were found in *Rhodococcus equi* (62% amino acid identity), *Nocardia farcinica* (61%), *Corynebacterium glutamicum* (47%), *Streptomyces griseus* (42%) and two in *Rubrobacter xylanophilus* (36 and 32%) (Fig. 3.3).

The genetic environment in the region containing the *gpgP* gene is highly conserved in all available mycobacterial genomes (Fig. 3.3). Similar genetic environments were also found in the genera *Nocardia*, *Corybacterium* and *Rhodococcus* (Fig. 3.3).

BLAST analysis with the *M. tuberculosis* GpgP amino acid sequence in this organism's genome revealed 5 proteins with relevant homology; the C-terminal domain of Rv2288c (43% amino acid identity), Rv3214-*gpm2* (33%), Rv3837c (32%), Rv0525 (32%) and Rv0489-*gpm1* (30%), all of which with putative PGM functions attributed. Although a sixth hypothetical PGM is annotated in *M. tuberculosis* genome (Rv2135c), no relevant homology with the GpgP was found in the BLAST analysis.

A comparison between the mycobacterial, other know or putative GpgPs and with the functionally related mannosyl-3-phosphoglycerate phosphatases (MpgPs) shows no relevant homology. Phylogenetic analysis of the same sequences clearly clusters the mycobacterial GpgP in the group classified as histidine phosphatase superfamily (branch 1) (Fig. 3.4). The other *M. tuberculosis* enzymes annotated as putative PGM also cluster in this group while other known GpgPs and MpgPs cluster in the HAD hydrolase superfamily group (Fig. 3.4). Further sequence analyses shows that unlike all the other known GpgP enzymes that are "DDDD" phosphohydrolases (Costa et al,

2007), the mycobacterial GpgP exhibited the well-known and characteristic "RHG" motif of the histidine phosphatase superfamily (Rigden, 2008) (Fig. 3.5).

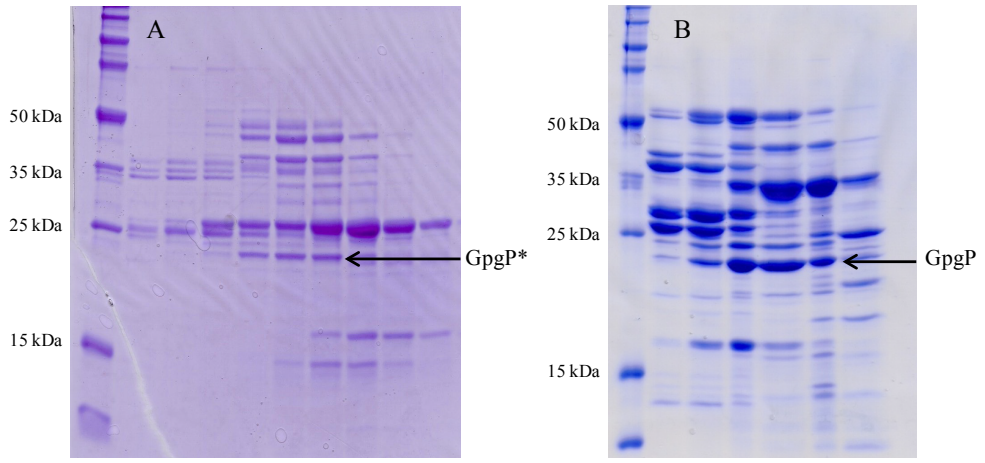



**Figure 3.2:** Large SDS-PAGE showing the partially purified native GpgPs from *M. hassiacum* **(A)** *and M. vanbaalenii* **(B)**. (*) - The identity of the *M. hassiacum* GpgP band was not confirmed and it is only a probable assumption.

### Properties of the partially purified native GpgP from *Mycobacterium hassiacum*

The purification of the native GpgP from *M. hassiacum* yielded a partially purified fraction with high GpgP activity, where the corresponding band was one of the major detected by SDS-PAGE (Fig. 3.2). The partially purified native GpgP used GPG as the preferred substrate and residual activity was also detected with MPG (<17%) and MGPG (<8%). No activity was detected for any of the other substrates tested (Fig. 3.6). Kinetic experiments showed that the partially purified GpgP exhibited

Michaelis-Menten kinetics at 50ºC with a $K_m$ for GPG of 0.85 ± 0.08 mM (Table 3.1 and Fig. 3.7 E).

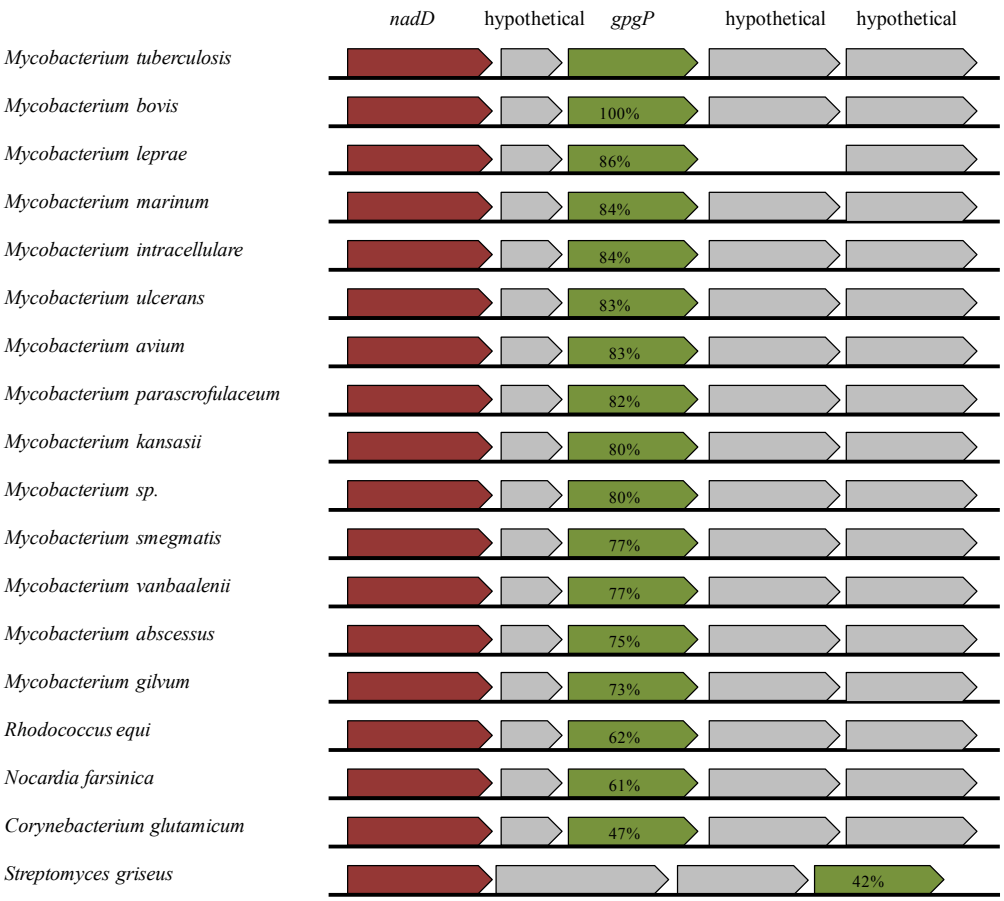



**Figure 3.3:** Genetic context of the mycobacterial and actinobacterial *gpgP* genes. *nadD* - Probable nicotinate-nucleotide adenylyltransferase. *gpgP* – Glucosyl-3-phosphoglycerate phosphatase.

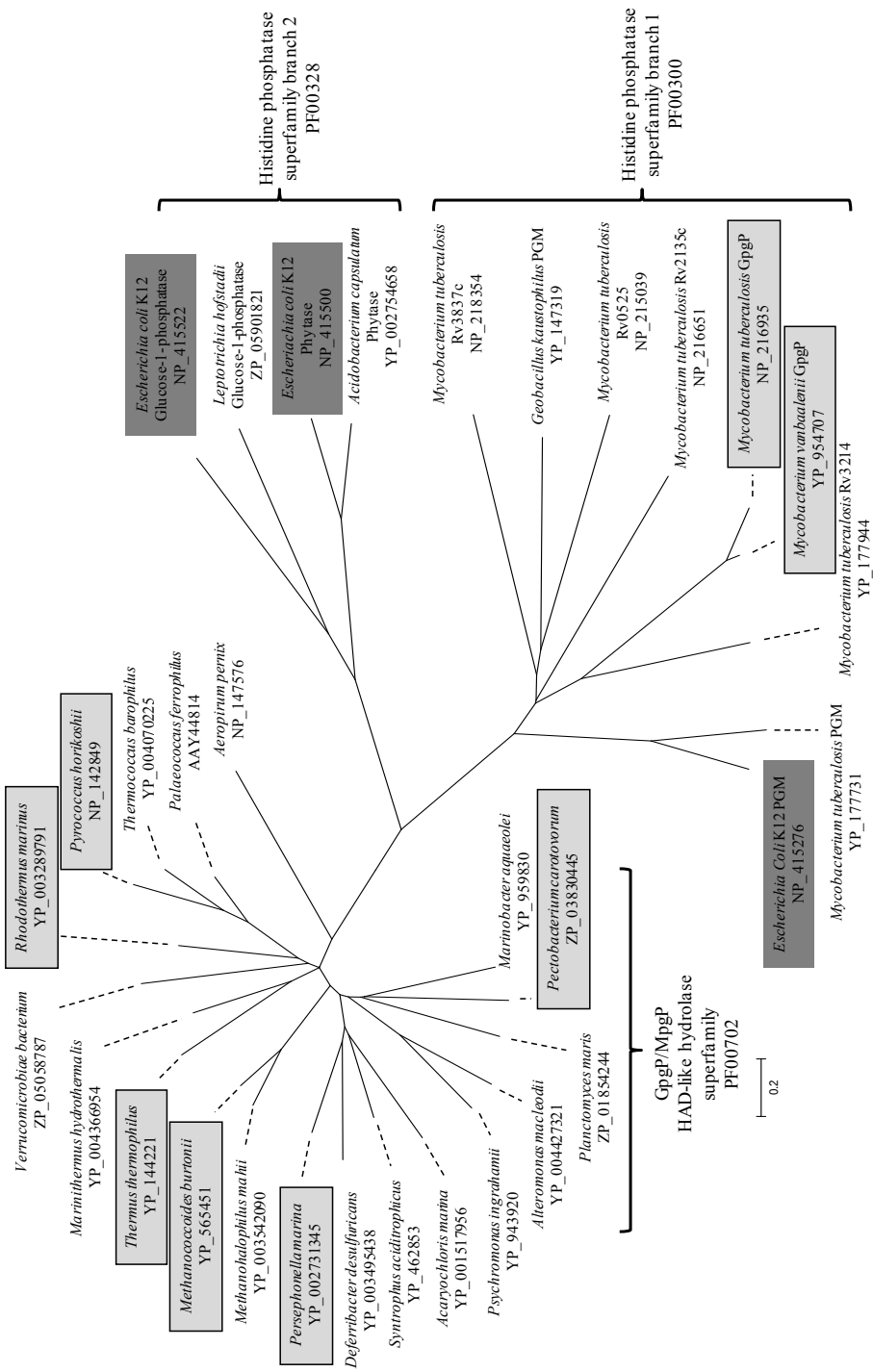

Histidine phosphatase superfamily branch 2 PF00328
Histidine phosphatase superfamily branch 1 PF00300
GpgP/MpgP HAD-like hydrolase superfamily PF00702
0.2
Escherichia coli K12 Glucose-1-phosphatase NP_415522
Leptotrichia hofstadii Glucose-1-phosphatase ZP_05901821
Escheriachia coli K12 Phytase NP_415500
Acidobacterium capsulatum Phytase YP_002754658
Mycobacterium tuberculosis Rv3837c NP_218354
Geobacillus kaustophilus PGM YP_147319
Mycobacterium tuberculosis Rv0525 NP_215039
Mycobacterium tuberculosis Rv2135c NP_216651
Mycobacterium tuberculosis GpgP NP_216935
Mycobacterium vanbaalenii GpgP YP_954707
Mycobacterium tuberculosis Rv3214 YP_177944
Mycobacterium tuberculosis PGM YP_177731
Escherichia Coli K12 PGM NP_415276
Pectobacteriumcarotovorum ZP_03830445
Marinobacter aquaeolei YP_959830
Planctomyces maris ZP_01854244
Alteromonas macleodii YP_004427321
Psychromonas ingrahamii YP_943920
Acaryochloris marina YP_001517956
Syntrophus aciditrophicus YP_462853
Deferribacter desulfuricans YP_003495438
Persephonella marina YP_002731345
Methanohalophilus mahii YP_003542090
Methanococcoides burtonii YP_565451
Thermus thermophilus YP_144221
Marinithermus hydrothermalis YP_004366954
Verrucomicrobiae bacterium ZP_05058787
Rhodothermus marinus YP_003289791
Pyrococcus horikoshii NP_142849
Thermococcus barophilus YP_004070225
Palaeococcus ferrophilus AAY44814
Aeropirum pernix NP_147576

**Figure 3.4:** Unrooted phylogenetic tree based on available amino acid sequences of identified and putative GpgPs/MpgPs (EC 3.1.3.70), PGMs (EC 5.4.2.1) and other enzymes belonging to the histidine phosphatase superfamily. Organisms where glucosylglycerate or mannosylglycerate has been detected are shaded in light gray and proteins which have been demonstrated to dephosphorylate GPG and MPG are surrounded by boxes. *E. coli* proteins highlighted in dark gray have confirmed function. Peptide accession numbers are below the organisms names, and have been retrieved from the NCBI database. Scale bar, 0.2 changes per site.

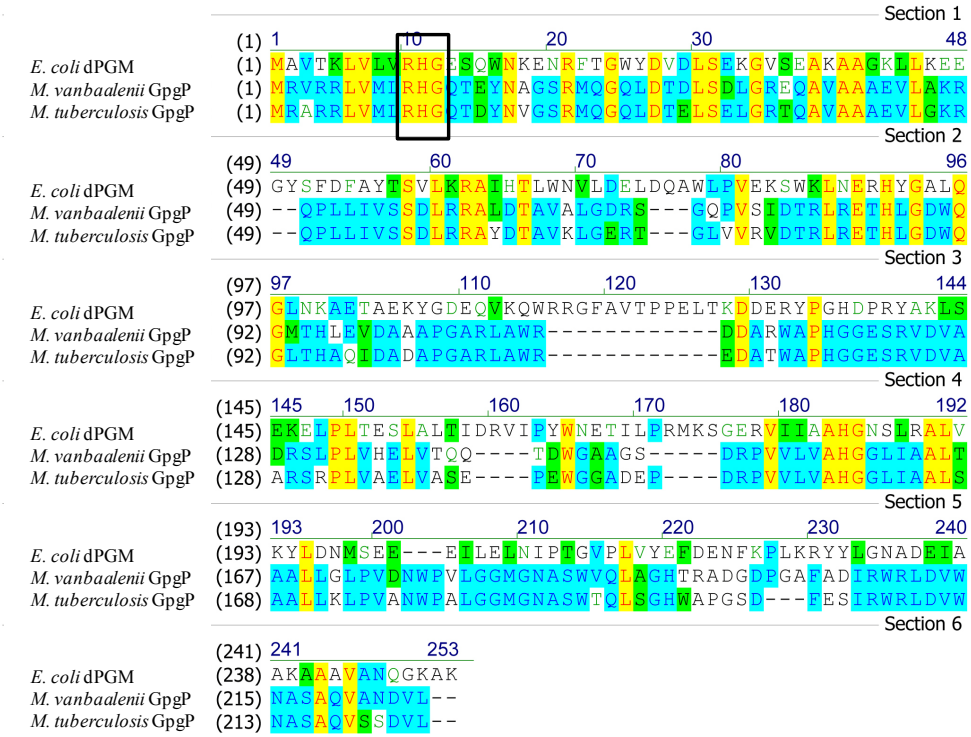



**Figure 3.5:** Alignment of the amino acid sequences of the *E. coli* dPGM and the *M. vanbaalenii* and *M. tuberculosis* GpgPs. The typical RHG motif of the histidine phosphatase superfamily is highlighted by a black box.

The enzyme was active between 20 and 60ºC, with maximal activity at about 50ºC (Fig. 3.7 A). At 50ºC, the enzyme was active between pH 4.0 and 10.0, optimally at pH 7.0 (Fig. 3.7 B). The native GpgP was not dependent on divalent cations. There were no differences in activity when using $Mg^{2+}$ (2.5-10 mM), $Mn^{2+}$ (2.5 mM), EDTA (2.5 and 10 mM) and deionized water. However, higher $Mg^{2+}$, $Mn^{2+}$ concentrations progressively inhibited the enzyme activity. The addition of $Co^{2+}$ (2.5 mM) led to partial inhibition and higher concentrations resulted in progressively high activity inhibition (5-50 mM). $Cu^{2+}$ totally inhibited the enzyme activity at all concentrations tested (2.5-50 mM).

**Expression and properties of the recombinant GpgP from *Mycobacterium tuberculosis***

Expression of the *M. tuberculosis gpgP* gene in *E. coli* resulted in the high level production of recombinant His-tagged protein (Fig. 3.8). Gel filtration experiments indicated that the recombinant His-tagged GpgP behaved as a dimeric protein in solution, with a molecular mass of about 46.0 ± 2.8 kDa.

Of the substrates tested, and like the native *M. hassiacum* GpgP, the recombinant GpgP preferentially used GPG and residual activity was also detected with MPG (<10%) and MGPG (<5%) (Fig 3.7 C). Due to the annotation of this protein as a phosphoglycerate mutase (PGM) a test was performed to assess whether the *M. tuberculosis* GpgP retained some PGM activity. A recombinant human dPGM was used as a positive control. While this enzyme was capable of converting 3-phosphoglycerate to 2-phosphoglycerate in the presence of 2,3-diphosphoglycerate, detected by a clear reduction in the absorption at 340 nm reflecting NADH oxidation (Fig. 3.1) that stabilized after 2 min of reaction, no activity was detected for the *M. tuberculosis* GpgP during a 5 min reaction.

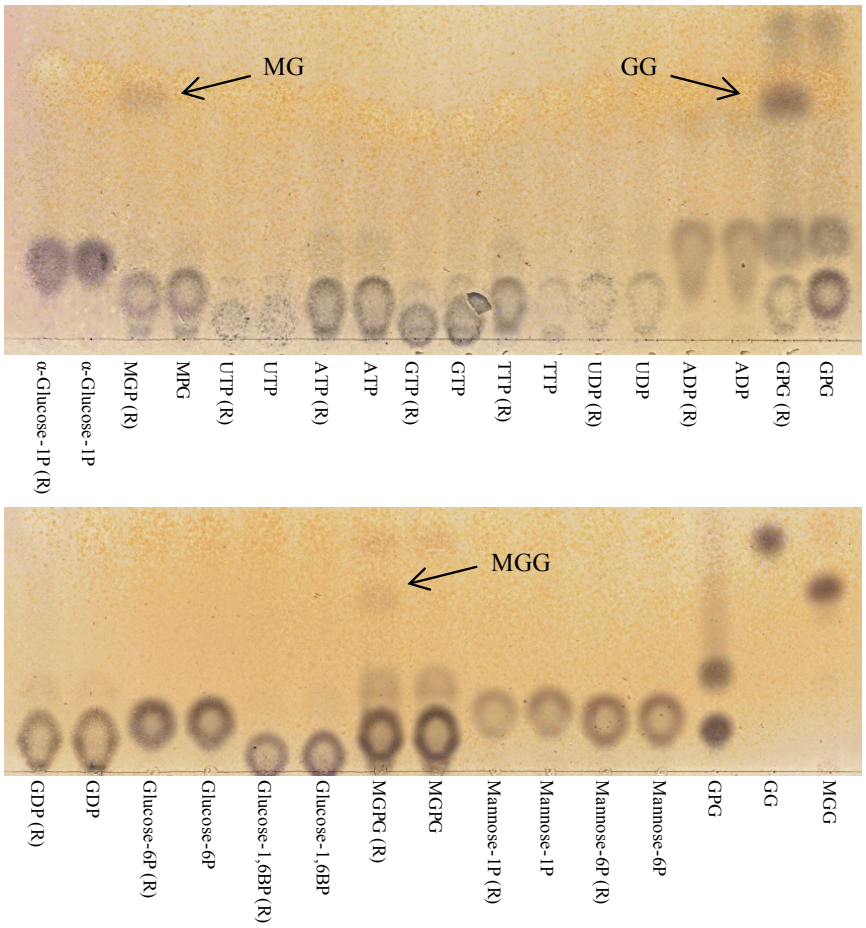
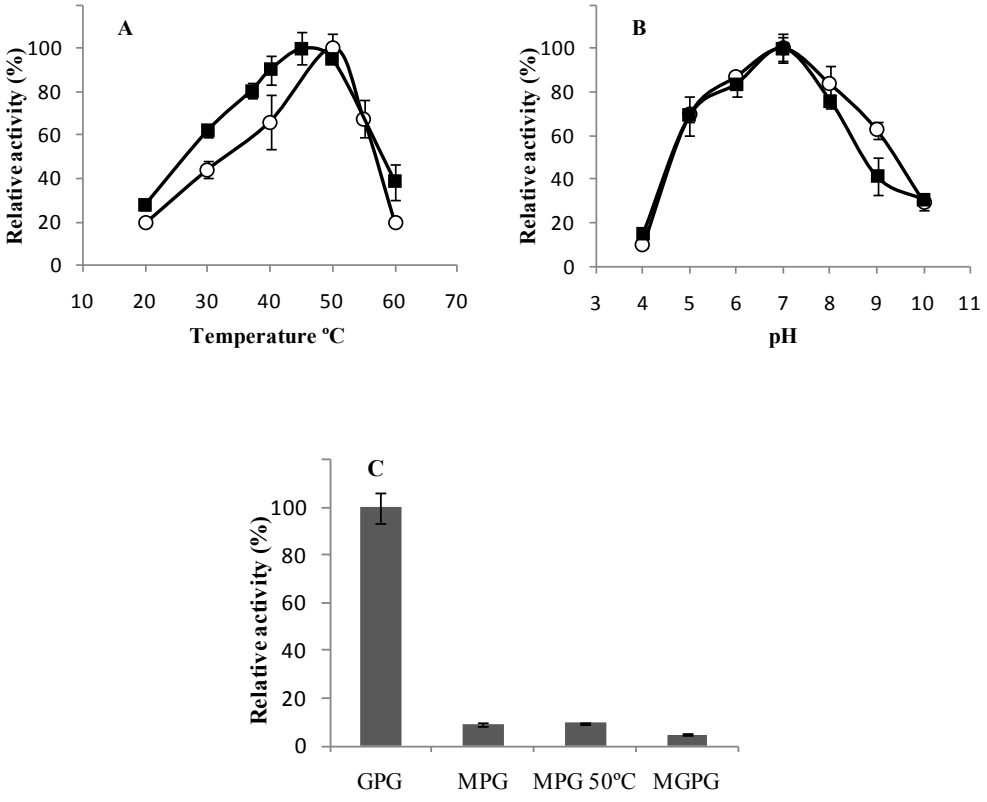



**Figure 3.6:** TLC analyses of the reaction products obtained with the native GpgP from *M. hassiacum* after 10 min incubation at 50ºC. The lanes with an (R) are the reactions, while the other lanes are controls with full reaction mixture but no enzyme.

Kinetic experiments showed that the recombinant GpgP exhibited Michaelis-Menten kinetics at 37ºC with a $K_m$ for GPG of 0.35 ± 0.03 mM and with $V_{max}$ of 67.19 ± 1.39 µmol/min.mg of protein (Table 3.1 and Fig. 3.7 D).

The enzyme was active between 20 and 50ºC, with maximal activity at about 45ºC (Fig. 3.7 A). At 37ºC, the enzyme was active between pH 4.0 and 10.0 and optimally

at pH 7.0 (Fig. 3.7 B). The recombinant GpgP was not dependent on added divalent cations, although the addition of EDTA (5 mM) reduced the activity to about 70%. There were no relevant differences in activity between $Mg^{2+}$ (1-10 mM), $Mn^{2+}$ (1-10 mM), $Co^{2+}$ (1-5 mM) and deionized water. However, higher $Co^{2+}$ concentration resulted in inhibitory effects while $Cu^{2+}$ totally inhibited the enzyme activity at any tested concentration. Stabilization assays showed that the enzyme retained full activity after 1 week storage at -20ºC in 50 mM BTP with 50 mM NaCl (pH 7.0).



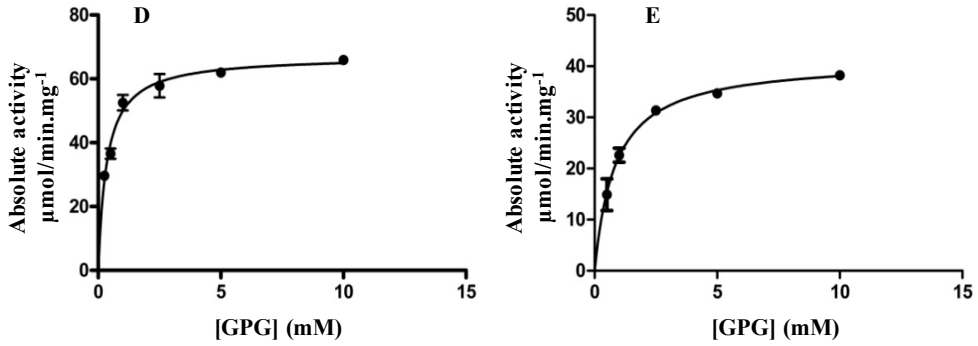



**Figure 3.7:** Temperature (**A**) and pH (**B**) profiles of the native GpgP from *M. hassiacum* (o) and the recombinant GpgP from *M. tuberculosis* (■). (**C**) – Substrate usage of the recombinant GpgP from *M. tuberculosis.* Kinetic studies of the recombinant GpgP from *M. tuberculosis* (**D**) and the native GpgP from *M. hassiacum* (**E**).

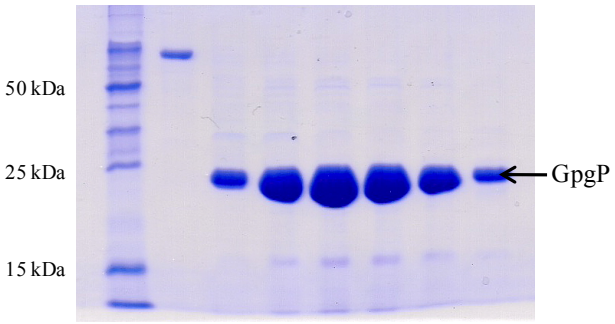
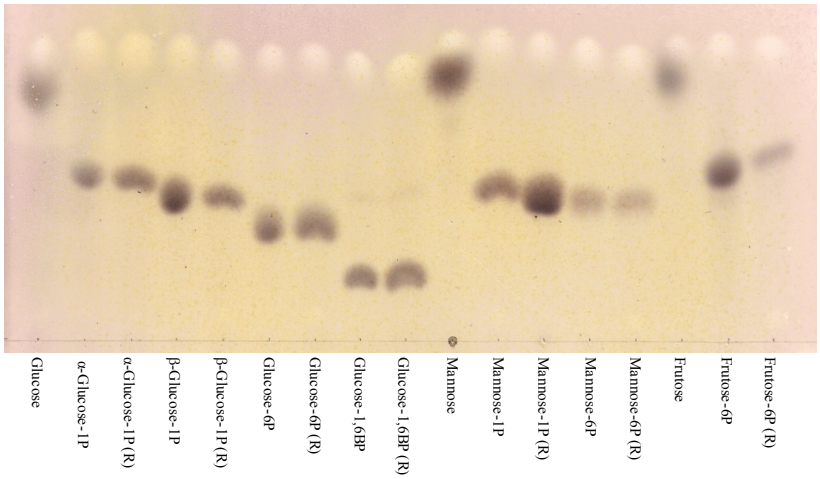



**Figure 3.8:** SDS-PAGE gel of the pure recombinant GpgP from *M. tuberculosis*.

**Table 1.** Kinetic parameters of the, partially purified, native GpgP, from *M. hassiacum* and the pure recombinant GpgP from *M. tuberculosis.*

| Substrate (GPG) | *M. tuberculosis* | *M. hassiacum* |
|---|---|---|
| $K_m$ (mM) | 0.35 ± 0.03 | 0.85 ± 0.08 |
| $V_{max}$(µmol/min.mg$^{-1}$) | 67.19 ± 1.39 | ___ |



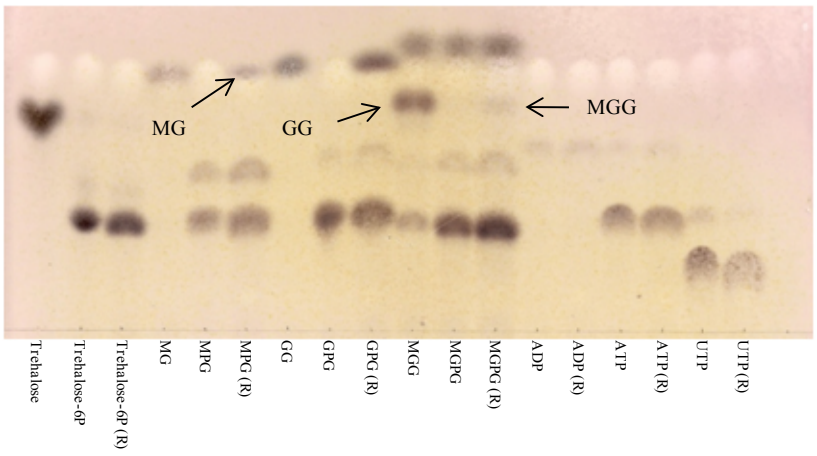



**Figure 3.9:** TLC analyses of the reaction products obtained with the recombinant GpgP from *M. tuberculosis* after 10 min incubation at 37ºC. The lanes with an (R) are the reactions, while the other lanes are controls with full reaction mixture but no enzyme.

## Discussion

Mycobacterial polymethylated polysaccharides had their structure extensively studied and recently their synthesis received renewed attention. However, most of the genes and enzymes involved in their biosynthetic pathways remain unknown (Gray & Ballou, 1971; Kamisango et al, 1987; Tuffal et al, 1998; Weisman & Ballou, 1984a; Weisman & Ballou, 1984b). While the methylmannose polysaccharide (MMP) has a simple structure, the structure complexity of the methylglucose lipopolysaccharide (MGLP) suggests that several enzymes are involved in the biosynthesis of this molecule. Furthermore, the reported essentiality of several genes (Sassetti et al, 2003), experimentally and putatively implicated in MGLP biosynthesis makes this biosynthetic pathways attractive targets for possible drug development. The enzyme responsible for the first MGLP precursor had already been identified by our group (Empadinhas et al, 2008) so we then focused on identifying the enzyme responsible for the next biosynthetic step.

The glucosylglycerate (GG) levels detected in *M. hassiacum* and *M. vanbaalenii* cell extracts (our unpublished results) prompted us to test both extracts for glucosyl-3-phosphoglycerate phosphatase (GpgP) activity. The discovery of high GpgP activity in both cell extracts encouraged us to purify the native enzymes of these organisms in order to identify this enzyme since no known GpgP homologues (EC 3.1.3.70) could be detected in the available mycobacterial genomes. While *M. hassiacum* was chosen for likely enhanced enzyme stability when compared to other mycobacteria, due to its thermophilic characteristics, the lack of a sequenced genome led us to choose a closely

related species of *Mycobacterium* for protein identification, namely *M. vanbaalenii* due to the close phylogenetic relatedness with *M. hassiacum*.

The mycobacterial GpgP identified by us belongs to the histidine phosphatase superfamily (branch 1) (Fig. 3.4). Since all the other known GpgPs belong to the HAD-like hydrolase superfamily (Costa et al, 2006), this was a surprising and unprecedented activity detected in this superfamily.

The mycobacterial *gpgP* gene was annotated as a putative phosphoglycerate mutase (PGM) in the databases; however, no PGM activity could be measured for the enzyme. Interestingly, several other *M. tuberculosis* proteins with some level of similarity to the *M. tuberculosis* GpgP were also initially annotated as putative PGMs. While some of these proteins have already been shown to have different activities (Watkins & Baker, 2006; Watkins & Baker, 2008), the most likely candidate for that activity in *M. tuberculosis* is the protein encoded by Rv0489 (Muller et al, 2005). The widespread mis-annotations in the histidine phosphatase superfamily seem to result from the mutase-dominated superfamily history, while the phosphatase activity is essentially the dominant (Rigden, 2008). As is the case for other superfamilies, the histidine phosphatases presents considerable challenges to automated genome annotation tools due to their remarkable sequence and functional diversity (Friedberg, 2006; Rigden, 2008). Apparently, most enzymes in this superfamily are phosphatases and the mutase activity is, in fact, the exception in this group (Rigden, 2008).

The mycobacterial GpgP had homologues in all available mycobacterial genomes and in many actinobacterial genomes (Fig 3.3). All the species known to synthesize MGLP had a homologue of this gene suggesting an involvement in the synthesis of MGLP. *Rubrobacter xylanophilus*, with apparently no known HAD-like hydrolase with MpgP/GpgP activity, and where there were no reports of MGLP synthesis, is known to synthesize mannosylglycerate (MG) by a two-step pathway (Empadinhas et al, 2007), and harbors 2 homologues of the mycobacterial GpgP. This finding suggests that at least one of the mycobacterial GpgP homologues in *R. xylanophilus* may be

responsible for the dephosphorylation of mannosyl-3-phosphoglycerate (MPG) to MG. This hypothesis is supported in part by the residual MPG-dephosphorylating activity found in both the mycobacterial GpgPs characterized in this work. Since amino acid sequence similarity of the *R. xylanophilus* homologues are only of 36% and 32%, it is possible that these differences may justify different substrate specificity and preference for MPG.

The substrate specificity, pH and temperature profiles of the *M. hassiacum* and *M. tuberculosis* GpgP were very similar. Both enzymes were also independent of cations for maximum activity although the *M. tuberculosis* GpgP was slightly inhibited by the addition of EDTA. However the $K_m$ of the *M. hassiacum* GpgP for GPG was over 2 times higher than that for the *M. tuberculosis* enzyme. This is not unexpected since the $K_m$ values tend to increase with temperature. While the $K_m$ for the *M. tuberculosis* GpgP was calculated at the organism optimal growth temperature of 37ºC, for the *M. hassiacum* enzyme the $K_m$ was calculated at 50ºC, this organism's optimum growth temperature.

Unlike most of the other described GpgPs/MpgPs that show comparable activities at high temperatures both for GPG and MPG (Costa et al, 2006), the mycobacterial enzymes exhibit much lower activity with MPG (<16% for *M. hassiacum* at 50ºC and <10% for *M. tuberculosis* at 37 and 50ºC). The two types of GpgPs also exhibit substantial differences in cation dependence since all the previously characterized GpgPs (and MpgPs) depend on divalent cations for activity (Costa et al, 2007; Costa et al, 2006; Empadinhas et al, 2003; Empadinhas et al, 2001) while the mycobacterial GpgP is not.

Interestingly, while the *M. tuberculosis* glucosyl-3-phosphoglycerate synthase (GpgS) gene (Rv1208) was proposed to be essential for growth (Sassetti et al, 2003), the GpgP gene (Rv2419c) was not. This suggests that one or more of the other GpgP homologues found in the *M. tuberculosis* genome or an unspecific phosphatase could exhibit GpgP activity under specific conditions and possibly be able to replace this

activity. However, this remains a hypothesis since no other protein fractions with this activity were detected during purification of the GpgPs from *M. hassiacum* and *M. vanbaalenii* cell extracts.

The apparent ubiquity of MGLP in pathogenic mycobacteria suggests a fundamental role for this structure in mycobacterial physiology. It seems reasonable to assume, also due to the reported essentiality of some of the genes involved in this biosynthetic pathway (Sassetti et al, 2003), that the absence or deficient assemblage of this structure could lead to cell death.

This work established that the phosphorylated form of GG, (GPG), is in fact, the primer for the synthesis of this polysaccharide and that the *M. tuberculosis* Rv1208 gene encoding a GpgS, suggested to be essential for the growth of this organism (Sassetti et al., 2003), is responsible for this step. This step is followed by the specific dephosphorylation of GPG, before the addition of the glucose unit that initiates the MGLP main chain by a yet unknown glycosyltransferase (Fig. 2.5).

This work provided the experimental confirmation of the proposed second biosynthetic step for the synthesis of the MGLP (Empadinhas et al, 2008; Jackson & Brennan, 2009) (Fig. 3.10), allowing the characterization of the corresponding enzyme. The identification of a new function in the genome of *M. tuberculosis* with the functional assignment of the Rv2419c to GpgP and its implication in MGLPs biosynthesis not only expands the already broad activity range of the histidine phosphatase superfamily but also furthers our understanding of this intricate metabolic trait, which might help us devise new strategies to fight tuberculosis.

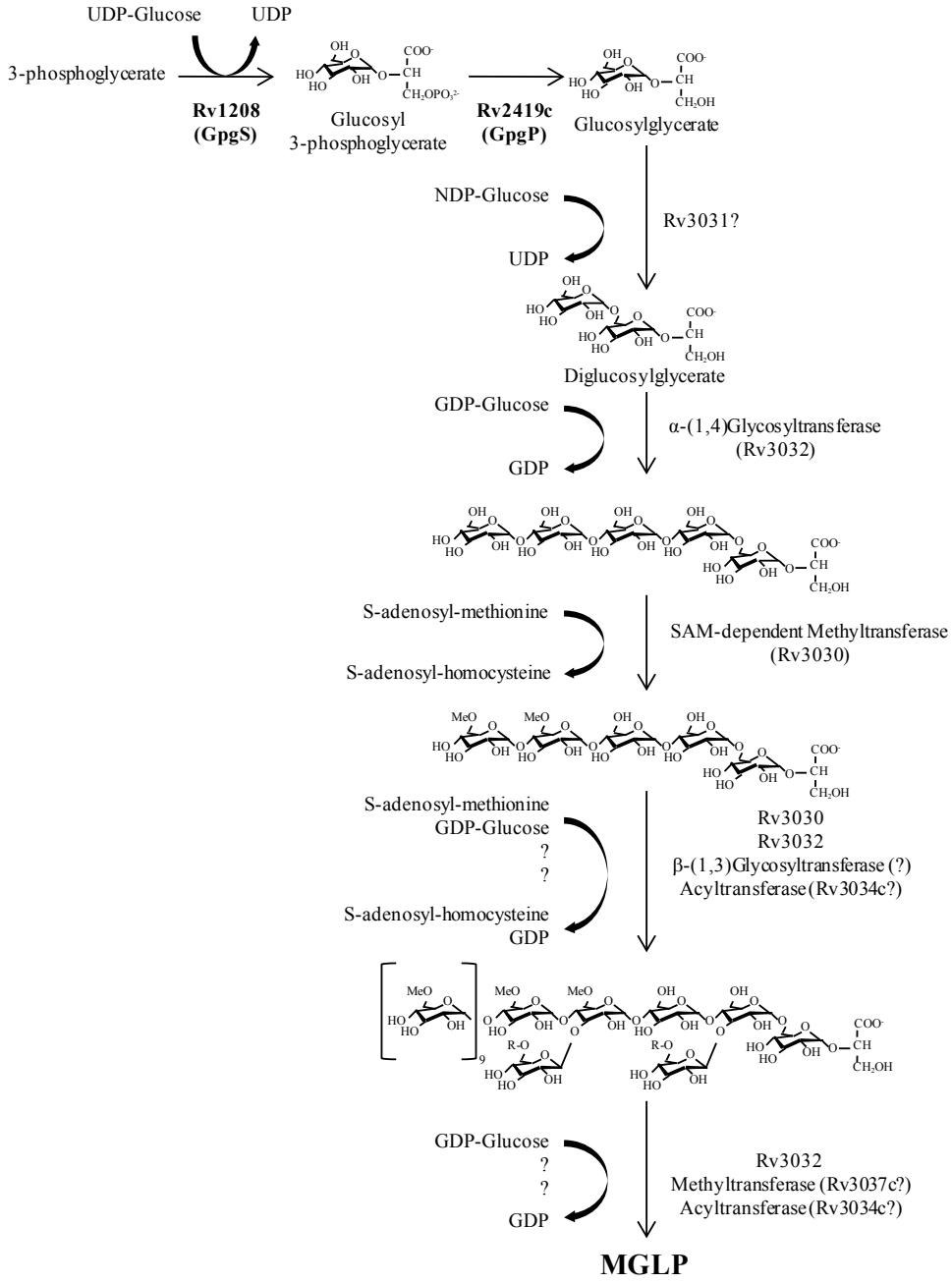



**Figure 3.10:** Proposed and updated pathway for the synthesis of MGLP. Partially adapted from (Jackson & Brennan, 2009).

## Acknowledgements and work contributions

This work was supported by Fundação para a Ciência e a Tecnologia, FEDER, through start-up funding by Center for Neuroscience and Cell Biology (Associate Laboratory) to Nuno Empadinhas and Project PTDC/BIA-PRO/110523/2009.

# Chapter 4

## Biochemical characterization of the maltokinase from *Mycobacterium bovis* BCG

**This chapter is published in:**

Mendes V, Maranha A, Lamosa P, da Costa MS, Empadinhas N (2010). **Biochemical characterization of the maltokinase from *Mycobacterium bovis* BCG**. *BMC Biochemistry* **11:** 21.

# Contents

## Abstract

Maltose-1-phosphate was detected in *Mycobacterium bovis* BCG extracts in the 1960's but a maltose-1-phosphate synthase (maltokinase, Mak) was only much later purified from *Actinoplanes missouriensis*, allowing the identification of the *mak* gene. Recently, this metabolite was established to be the intermediate in a pathway linking trehalose with the synthesis of glycogen in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis* by the action of a maltosyltransferase. Although the *M. tuberculosis* H37Rv *mak* gene (Rv0127) was considered essential for growth, no mycobacterial Mak has, to date, been characterized.

The sequence of the Mak from *M. bovis* BCG was identical to that from *M. tuberculosis* strains (99-100% amino acid identity). The enzyme was dependent on maltose and ATP, although GTP and UTP could be used to produce maltose-1-phosphate, which we identified by TLC and characterized by NMR. The $K_m$ for maltose was 2.52 ± 0.40 mM and 0.74 ± 0.12 mM for ATP; the $V_{max}$ was 21.05 ± 0.89 μmol/min.mg$^{-1}$. Divalent cations were required for activity and $Mg^{2+}$ was the best activator. The enzyme was a monomer in solution, had maximal activity at 60ºC, between pH 7 and 9 (at 37ºC) and was unstable on ice and upon freeze/thawing. The addition of 50 mM NaCl markedly enhanced Mak stability.

The unexplored role of maltokinases in mycobacterial metabolism, the possible involvement in the synthesis of MGLP and the lack of biochemical data led us to express the *mak* gene from *M. bovis* BCG for biochemical characterization. This is the first mycobacterial Mak to be characterized and its properties represent essential knowledge towards deeper understanding of mycobacterial physiology. Since Mak may be a potential drug target in *M. tuberculosis*, its high-level production and purification in bioactive form provide important tools for further functional and structural studies.

## Introduction

Maltose-1-phosphate was for the first time identified in the 60's, in *Mycobacterium bovis* BCG cell free extracts (Narumi & Tsumita, 1967). Much later an enzyme capable of catalysing the conversion of maltose and ATP to maltose-1-phosphate was identified in *Actinoplanes missouriensis* and named maltokinase (Mak) (Drepper et al, 1996). A maltokinase from *Streptomyces coelicolor* was also identified and partially characterized (Jarling et al, 2004).

The *A. Missouriensis* Mak was found to be constitutively expressed, regardless of the sugars present in the growth medium and therefore unlikely to play an essential role in maltose catabolism (Niehues et al, 2003). Maltose-1-phosphate also found to be synthesized in an ATP independent way in *E. coli*, where it was proposed to be involved in the regulation of maltose metabolism (Decker et al, 1999). The maltokinase gene (*mak*) is present in most of the available mycobacterial genomes and it was considered essential for the growth of *Mycobacterium tuberculosis* H37Rv (Sassetti et al, 2003). Moreover, *mak* is always associated with the trehalose synthase gene (*treS*) and in several organisms fused into one gene. This extremely conserved arrangement is found in many bacterial groups world and suggests a shared metabolic route.

The disaccharide trehalose has a ubiquitous occurrence in bacteria, archaea and fungi, but it is also found in plants and in some invertebrates, playing diverse biological roles (Iturriaga et al, 2009; Paul et al, 2008). Mycobacteria and closely related organisms possess three different pathways to synthesize trehalose reflecting the importance of this disaccharide that encompasses functional roles as a compatible solute under osmotic or thermal stress, and a structural role in cell wall components (De Smet et al, 2000; Takayama et al, 2005; Woodruff et al, 2004). Knock out of the three trehalose synthesizing pathways in *Mycobacterium smegmatis* rendered this organism unable to grow unless trehalose was supplemented to the growth medium (Woodruff et al,

2004). However, the pathways seem to have specific and different roles in closely related actinobacterial species. In *Corynebacterium glutamicum*, the TreY/TreZ pathway dominates trehalose synthesis, while the OtsA/OtsB plays only a minor role and the TreS pathway seems to be involved in trehalose degradation (Tzvetkov et al, 2003; Wolf et al, 2003). In *M. tuberculosis* the OtsA/OtsB pathway seems to be essential for the survival of this organism in infected mice, while inactivation of the TreY/TreZ pathway did not produce any detectable effects (Murphy et al, 2005). Interestingly, the inactivation of TreS in *M. tuberculosis* resulted in increased survival in mice (Murphy et al, 2005). Individual knock-out of each of the trehalose synthesizing pathways in *M. smegmatis* did not create any measurable effects indicating a redundancy of the three pathways in this organism (Woodruff et al, 2004). In *M. smegmatis* the TreS was shown to interconvert trehalose and maltose (Pan et al, 2004) and to catalyze the formation of trehalose from glycogen via its amylase activity (Pan et al, 2008). Recently, the synthesis of glycogen from trehalose was reported in *M. smegmatis* and *M. tuberculosis* (Elbein et al, 2010; Kalscheuer et al, 2010). In a three-step pathway, TreS and Mak consecutively convert trehalose into maltose and further into maltose-1-phosphate (Elbein et al, 2010; Kalscheuer et al, 2010). A maltosyltransferase (GlgE) transfers maltose from maltose-1-phosphate to glycogen and other α-glucans. Although *M. smegmatis* has an essential trehalase to degrade trehalose directly into glucose (Carroll et al, 2007), it was proposed that it is energetically advantageous to convert trehalose to glycogen rather than to hydrolyze it to glucose (Elbein et al, 2010), which cannot accumulate in cells.

Despite of the increasing numbers of genome sequences showing that *mak* genes are widely distributed throughout the bacterial world, there is still very little biochemical data available about the corresponding enzymes as only the Mak from *A. missouriensis* and *S. coelicolor* have been characterized (Jarling et al, 2004; Niehues et al, 2003). Therefore, we cloned and successfully expressed the *mak* gene from *M. bovis* BCG in *E. coli*, the organism where maltose-1-phosphate was firstly detected.

We purified the recombinant bioactive enzyme and determined its biochemical and kinetic properties to further understand its role in mycobacterial physiology.

## Methods

### Identification, sequence analysis, cloning and functional overexpression of the maltokinase gene (*mak*) from *Mycobacterium bovis* BCG, strains and growth conditions

To identify the *mak* gene in mycobacterial genomes we used the amino acid sequences of the *A. missouriensis* and *S. coelicolor* maltokinases (GenPept accession number AAQ01690 and CAA04602, respectively) in BLAST searches at the National Center for Biotechnology Information (NCBI, http://blast.ncbi.nlm.nih.gov/Blast.cgi) database. Promoter identification was carried out using the prokaryotic promoter prediction software from Berkeley University, available at (http://www.fruitfly.org). The *mak* gene was amplified from the chromosomal DNA from *M. bovis* BCG (DSM 43990) obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ, Braunschweig, Germany), with a pair of primers based on the gene sequence retrieved from the Institut Pasteur database (http://genolist.pasteur.fr/TubercuList/). An NdeI restriction site (underlined) was added to the forward primer MtuNde (5'-CTTACATATGACTCGGTCGGACACGC-3'). A HindIII restriction site (underlined) was added to the reverse primer MtuHind (5'-ATTAAGCTTGCTAGCGGTCAGGCGGG-3'). The stop codon was removed from the reverse primer to allow the translation of a C-terminal 6×His-tag encoded by the expression vector pET30a (Novagen). PCR was carried out with the AccuPrime GC-Rich DNA Polymerase (Invitrogen). DNA (200 ng) was denatured by 95ºC for 5 min followed by 30 cycles of 1 min denaturation step at 95ºC, 1 min annealing step at

62ºC, and 1.5 min extension step at 72ºC. The PCR product was purified from agarose gel (NZYTech, Portugal), digested and cloned into pET30a. The construct was sequenced to confirm the identity of the insert (AGOWA, Berlin, Germany), and transformed into *E. coli* BL-21, which was used as host for expression of the *mak* gene. Recombinant *E. coli* was grown in a 5 L bioreactor, with continuous aeration and stirred at 180 rpm at 37ºC, pH 7.0 in LB medium with kanamycin (30 μg/mL), to mid-exponential phase of growth ($OD_{610}$ = 0.8). IPTG was added at a final concentration of 0.5 mM to induce gene expression, and temperature was reduced to 20ºC. The cells were harvested 18 h later by centrifugation (7000×*g*, 10 min, 4ºC).

**Preparation of cell-free extracts**

*E. coli* cells carrying the recombinant maltokinase (Mak) from *M. bovis* BCG were suspended in 25 mM Bis-tris propane buffer (BTP) at pH 7.5 with 50 mM NaCl for enzyme assays, or in 20 mM sodium phosphate buffer at pH 7.4 with 0.5 M NaCl and 20 mM imidazole, for protein purification. A protease inhibitor cocktail (Roche), 10 μg/ml DNAse I and 5 mM $MgCl_2$ were added to the suspension. Cells were disrupted twice in a French-press cell followed by centrifugation (15000×*g*, 4ºC, 30 min).

**Enzyme assays**

The activity of the recombinant Mak from *M. bovis* BCG in *E. coli* extracts and during purification was detected after 15 min at 37ºC in reaction mixtures (50 μL) containing 25 μL of cell-free extract, 3.0 mM (each) of ATP and maltose, and 10 mM $MgCl_2$ in 50 mM BTP, pH 8.0. The synthesis of maltose-1-phosphate was monitored by thin-layer chromatography (TLC) with solvent systems composed by acetic acid/ethyl

acetate/water/ammonia 25% (6:6:2:1, v/v) and butanol/ethanol/water (5:3:2, v/v). Trehalose-6-phosphate, ATP, GTP, UTP, ADP, GDP, AMP and maltose standards were used for comparative purposes. The maltose-1-phosphate formed was identified by NMR as described below. *E. coli* cell-free extracts carrying an empty vector were used as negative controls. Protein concentration was determined by the Bradford method (Bradford, 1976).

**Purification of recombinant Mak from *Mycobacterium bovis* BCG**

The His-tagged recombinant Mak from *M. bovis* BCG was purified in a prepacked Ni-Sepharose high-performance column (His-Prep FF 16/10) equilibrated with 20 mM sodium phosphate, pH 7.4, 0.5 M NaCl, and 20 mM imidazole. Elution was carried out with 500 mM imidazole and the purity of the fractions was determined by SDS-PAGE. The purest active fractions were pooled, diluted ten times with 25 mM BTP at pH 7.4 and loaded into a Q-Sepharose fast-flow column (Hi-Load FF 16/10), equilibrated with 25 mM BTP at pH 7.4 with 50 mM NaCl, and eluted by a linear gradient of NaCl (50 to 500 mM). The purity of the fractions was determined by SDS-PAGE and the purest active fractions were pooled, concentrated by ultracentrifugation in 30 kDa cutoff centricons (Amicon), equilibrated with 25 mM BTP at pH 7.4 with 200 mM NaCl, and loaded into a Superdex 200 fast-flow column equilibrated with the same buffer. After SDS-PAGE analysis the active pure fractions were concentrated and equilibrated with 50 mM BTP at pH 7.4 with 50 mM NaCl. Protein content of the samples was determined by the Bradford assay (Bradford, 1976). The identity of the purified maltokinase from *M. bovis* BCG was confirmed by Peptide Mass Fingerprinting (IPATIMUP Proteomics Unit, Porto, Portugal).

**Characterization of the recombinant Mak from *Mycobacterium bovis* BCG**

The substrate specificity of the recombinant Mak from *M. bovis* BCG was determined using glucose, galactose, mannose, maltose, isomaltose, trehalose, maltotriose, maltotetraose, maltopentaose and maltoheptaose as possible acceptors and with ADP, CDP, GDP, TDP, UDP, ATP, CTP, GTP, TTP and UTP as possible phosphate donors (all from Sigma-Aldrich). Since the Mak protein had high sequence identity (>40%) with putative aminoglycoside phosphotransferases, this activity was also tested using the aminoglycoside antibiotics gentamicin, kanamycin, streptomycin and hygromycin B as possible phosphate acceptors (all from Sigma-Aldrich). The reaction mixtures (50 μL) containing (0.5 μg) pure recombinant Mak, 3 mM of each substrate, and 10 mM $MgCl_2$ in 50 mM BTP at pH 8, were incubated at 37ºC for 10 min. The products were visualized by TLC as described above.

Temperature and pH profile, effect of cations and thermal stability of Mak were determined by the addition of 0.5 μg of Mak to 50 μL reaction mixtures containing the appropriate buffer, 3 mM maltose and 3 mM NTP and stopped at different times by cooling on ethanol–ice. The Mak was inactivated by the addition of 5 μL of 1N HCl and neutralized by 5 μL of 1N NaOH. Controls were performed to account for possible NTP degradation following acid treatment. The amount of NDP released was determined at 340 nm after incubation of the sample with 3 U of pyruvate kinase and lactate dehydrogenase, 0.3 mM NADH and 2.5 mM phosphor-enolpyruvate (all from Sigma-Aldrich) in 1 mL mixture (total volume) for 10 min at 30ºC (Qu et al, 2004). The temperature profile was determined between 20 and 65ºC in 50 mM BTP at pH 7.0, with 10 mM $MgCl_2$. The effect of pH was determined at 37ºC in 50 mM BTP (pH 6.0 to 9.0) and in 50 mM CAPS (pH 9.0 to 11.0), with 10 mM $MgCl_2$. The effect of cations was examined by incubating reaction mixture containing the appropriate substrates, with the chloride salts of $Mg^{2+}$, $Mn^{2+}$, $Co^{2+}$, $Zn^{2+}$ (0.5 to 50 mM) or without cations, at 37ºC.

The kinetic parameters for the recombinant Mak were determined by measuring the amount of NDP released, as described above. The $K_m$ values for the substrates ATP, GTP, UTP and maltose were determined at 37ºC, from Lineweaver–Burk plots. All experiments were performed in triplicate.

The molecular mass of the recombinant Mak from *M. bovis* BCG was estimated by gel filtration on a Superdex 200 column and the molecular mass standards were aprotinin (6.5 kDa), ribonuclease (13.7 kDa), carbonic anhydrase (29 kDa), ovalbumine (43 kDa), conalbumine (75 kDa), aldolase (158 kDa). Blue Dextran 2000 was used to determine the void volume (Amersham).

**Enzyme stabilization assays**

Pure recombinant Mak was stored in several aliquots containing 50 mM BTP at pH 8.0, and one of the following substances (final concentration) was added: 20% glycerol, 50% glycerol, 10 mM maltose, 10 mM trehalose or 50 mM NaCl. The initial Mak activity was measured at 37ºC in 50 mM BTP at pH 8.0, with 10 mM $MgCl_2$ and 3 mM (each) of ATP and maltose. Aliquots were stored at 4 and -20ºC for seven days. Control aliquots were kept at 4 and -20ºC only in 50 mM BTP at pH 8.0. After seven days storage the residual activity was measured for each of the conditions described above.

**NMR spectroscopy**

For the NMR experiments, 10% (v/v) of $D_2O$ was added to the reaction mixture containing 5 mM (each) of ATP and maltose, 5 mM $MgCl_2$ in 10 mM BTP at pH 8.0. After spectral acquisition, 15 µg of enzyme were added to the NMR tube and

incubated for 10 min at 37ºC. All spectra were acquired on a Bruker AVANCE III 800 spectrometer (Bruker, Rheinstetten, Germany) working at a proton operating frequency of 800.33 MHz, equipped with a four channel 5 mm inverse detection probe head with pulse-field gradients along the Z axis.

Spectra were run at 25ºC using standard Bruker pulse programs. $^{1}H$ and $^{13}C$ chemical shifts are referenced to 3-(trimethylsilyl)propane sulfonic acid, and $^{31}P$ resonances to external 85% phosphoric acid. In the $^{1}H$-$^{13}C$ and the $^{1}H$-$^{31}P$ heteronuclear two-dimensional single quantum coherence (HSQC) spectra, delays of 3.44 and 71.4 ms were used for evolution of the $J_{H,C}$ and $J_{H,P}$ couplings, respectively; proton decoupling was, in both cases, achieved using the GARP4 sequence (Shaka et al, 1985).

## Results

### Identification of mycobacterial maltokinase (Mak) and sequence analysis

BLAST analysis with the *A. missouriensis* and *S. coelicolor* maltokinase sequences within mycobacterial genome databases showed that the mak gene was annotated as a hypothetical protein/trehalose synthase fused maltokinase gene with homologues in most of these genomes. The amino acid identities of these proteins to the *A. missouriensis* sequence were between 62% and 67%, and 39% to 41% to the *S. coelicolor* Mak sequence. The maltokinase gene (*mak*) from *M. bovis* BCG contained 1368 bp coding for a polypeptide with 455 amino acids with a calculated molecular mass of 49.9 kDa and a calculated isoelectric point of 5.2. The Mak protein was 99 to 100% identical to the homologues from all *M. tuberculosis* and *M. bovis* strains with the genomes available (Fig. 4.1). Additional Mak homologues were found in *Mycobacterium avium* (79% amino acid identity), *Mycobacterium kansasii* (79%), *Mycobacterium marinum* (79%), *Mycobacterium intracellulare* (78%),

*Mycobacterium ulcerans* (78%), *M. smegmatis* (64%), *Mycobacterium gilvum* (60%) and *Mycobacterium vanbaalenii* (59%), but not in *Mycobacterium leprae* nor in *Mycobacterium abscessus*, which also lack *treS* genes. In other organisms like *Rubrobacter xylanophilus* or *Pseudomonas fluorescens* the *mak* gene is fused with the *treS* gene (Fig. 4.1). In *M. bovis*, *M. tuberculosis* and *M. marinum* the *mak* and *treS* genes have intergenic regions of about 72 to 102 bp that could accommodate a promoter sequence, implying independent regulation of *mak* expression. Our attempts to detect consensus promoter motifs in this region indicated about 69 to 75% probability of these elements within the 3' end of the *treS* gene.

The Mak protein had significant sequence similarity with putative aminoglycoside phosphotransferases from several organisms of the phylum *Actinobacteria*, namely *Kribbella flavida* (NCBI accession number YP_003382767, 49% amino acid identity), *Actinosynnema mirum* (YP_003103727, 45% identity) and *Nakamurella multipartita* (YP_003202415, 45% identity).

**Expression and properties of the recombinant maltokinase**

Expression of the *M. bovis* BCG *mak* gene in *E. coli* resulted in the high level production of recombinant His-tagged protein (Fig 4.2). The Mak identity was confirmed by peptide mass fingerprinting. Gel filtration experiments indicated that the recombinant His-tagged Mak behaved as a monomeric protein in solution, with a molecular mass of about 50.7 ± 4.2 kDa.

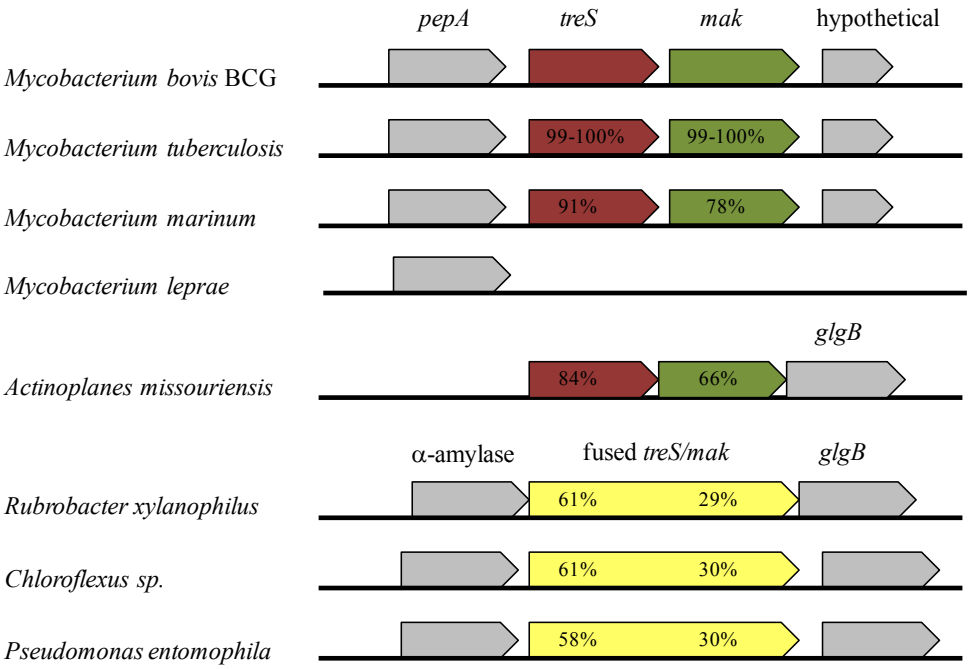



**Figure 4.1:** Genetic environment of the *mak* gene in different organisms. Organization of the region containing the *mak* gene in *M. bovis* BCG, *M. tuberculosis*, *M. marinum*, *M. leprae*, *A. missouriensis*, *R. xylanophilus*, *Chloroflexus* sp. and *P. entomophila*. Arrows represent genes and their orientation. (*pepA*) - Serine protease. (*treS*) - Trehalose synthase. (*mak*) – Maltokinase. (*glgB*) - Glycogen branching enzyme. (fused *treS/mak*) - Probable fused trehalose synthase/maltokinase.

The recombinant Mak used ATP, GTP and UTP as phosphate donors with comparable but decreasing efficiency (Table I). Residual activity was also detected with ADP and GDP (<7%). Among the sugar substrates tested, maltose was, by far, the preferred acceptor. Residual activity was also detected with maltotriose, maltotetraose,

maltopentaose and maltoheptaose (<2%), which may be due to maltose contamination. Due to the sequence similarities of Mak with putative aminoglycoside phosphotransferases, the aminoglycoside antibiotics gentamicin, hygromycin B, kanamycin and streptomycin were also tested as possible acceptors of phosphate but no activity was detected. Kinetic experiments showed that the recombinant Mak exhibited Michaelis-Menten kinetics at 37°C with ATP, GTP and UTP up to 5 mM (Table I and Fig. 4.3). Higher concentrations of these phosphate donors were progressively inhibitory, in the reaction mixture containing 20 mM maltose and 10 mM $MgCl_2$. While maximum Mak activity was observed with 5 mM ATP, the activity dropped to 75% with 10 mM ATP and to about 50% with 20 mM ATP.

The enzyme was active between 20 and 65ºC, with maximal activity at about 60°C (Fig. 4A). At 37ºC, the enzyme was active between pH 6 and 11, optimally between pH 7.0 and 9.0 with maltose and ATP, GTP or UTP (Fig. 4.4 B). The recombinant Mak was strictly dependent on divalent cations with $Mg^{2+}$ (10 mM) having the most pronounced stimulatory effect (Fig. 4.4 C). Other divalent cations like $Co^{2+}$ and $Mn^{2+}$ also activated Mak, but to lower extent (~41% and ~14% of maximal activity, respectively). The enzyme retained only about 40% of activity after 7 days storage at 4ºC in 50 mM BTP (pH 7.5). Most of the activity was also lost after freeze/thawing the enzyme at -20ºC in the same buffer. Glycerol strongly inhibited Mak activity, even at very low concentrations (1%). The addition of 10 mM maltose (final concentration) only slightly improved the stability at 4ºC to about 50% of maximal activity. However, the addition of 50 mM NaCl dramatically improved the stability of the enzyme, as the residual activity after 1 week at 4ºC was still above 90% of maximal activity.

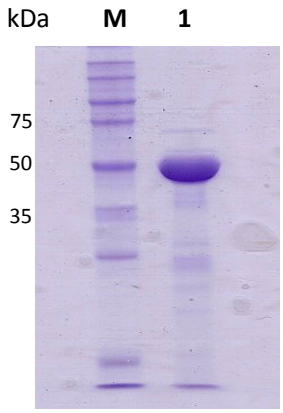



**Figure 4.2:** SDS-PAGE of the recombinant maltokinase (Mak) from *M. bovis* BCG. Lane 1 – Purified recombinant Mak. M - Protein MW markers.

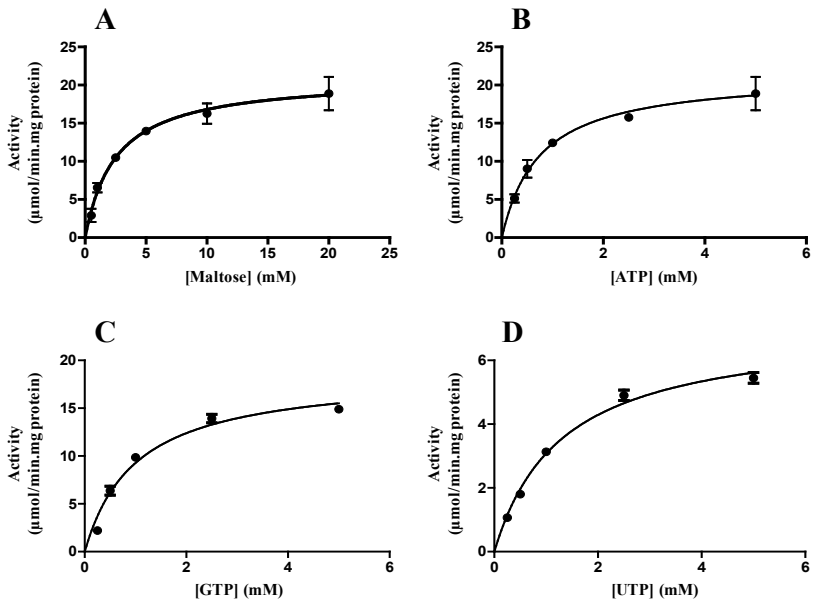



**Figure 4.3:** Kinetic studies of Mak activity. Dependence of Mak activity on the concentrations of maltose (**A**), ATP (**B**), GTP (**C**) and UTP (**D**).

**Identification of maltose-1-phosphate by NMR**

Maltose is composed of two glucose residues linked via an alpha-1,4-glycosidic bond. This means that one of the residues is blocked in the alpha configuration, while the other has a free aldehyde group at position one and therefore can adopt either the alpha or the beta configuration (Fig. 4.5 C). This explains the presence of the three signals in the $^{1}$H-NMR spectrum of maltose (Fig. 4.5 A), in the region where protons 1 of hexoses resonate; signal M corresponding to proton 1 of the blocked residue and signals $N_a$ and $N_b$ corresponding to proton 1 of unblocked residue in the alpha and beta configurations, respectively. After incubating the reaction mixture with the enzyme (Fig. 4.5 B), signals M, $N_a$, and $N_b$ almost disappear and only two new signals arise (signals O and P). These signals present H-H coupling constants $^{3}J_{1,2}$ around 3Hz, meaning that the respective residues are in the alpha configuration. Additionally, resonance O presents a second 7 Hz splitting of the signal, consistent with a H-P coupling constant. This phosphorylation was confirmed by a $^{1}$H-$^{31}$P HSQC spectrum that presents a correlation between signal O and a phosphorus resonance at 1.64 ppm (the region where phosphomonoesters resonate). This means that the reaction product has two hexose residues blocked in the alpha configuration and that one of them is phosphorylated at position 1, thus we conclude to be in the presence of alpha-D-maltose-1-phosphate. This result was further confirmed by the acquisition of a $^{1}$H-$^{13}$C HSQC spectrum and comparison of the $^{1}$H and $^{13}$C resonances with the literature (Drepper et al, 1996)

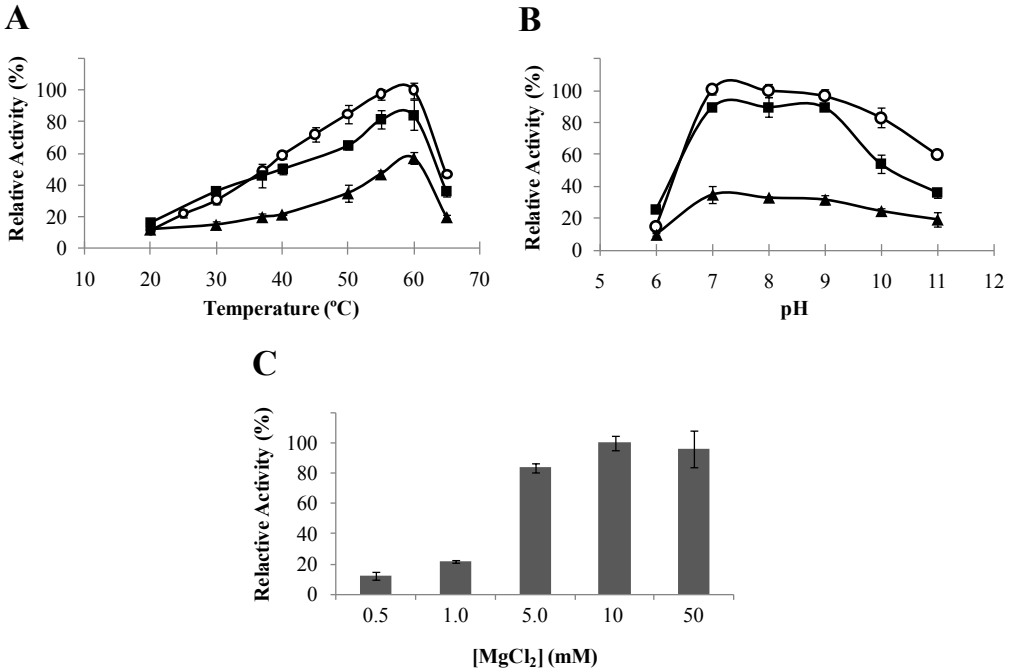



**Figure 4.4:** Effects of temperature, pH and $Mg^{2+}$ concentration on Mak activity. Temperature profile **(A)** and pH dependence **(B)** using ATP (o), GTP (■) and UTP ▲) as phosphate donors. **(C)** Effect of $Mg^{2+}$ concentration.

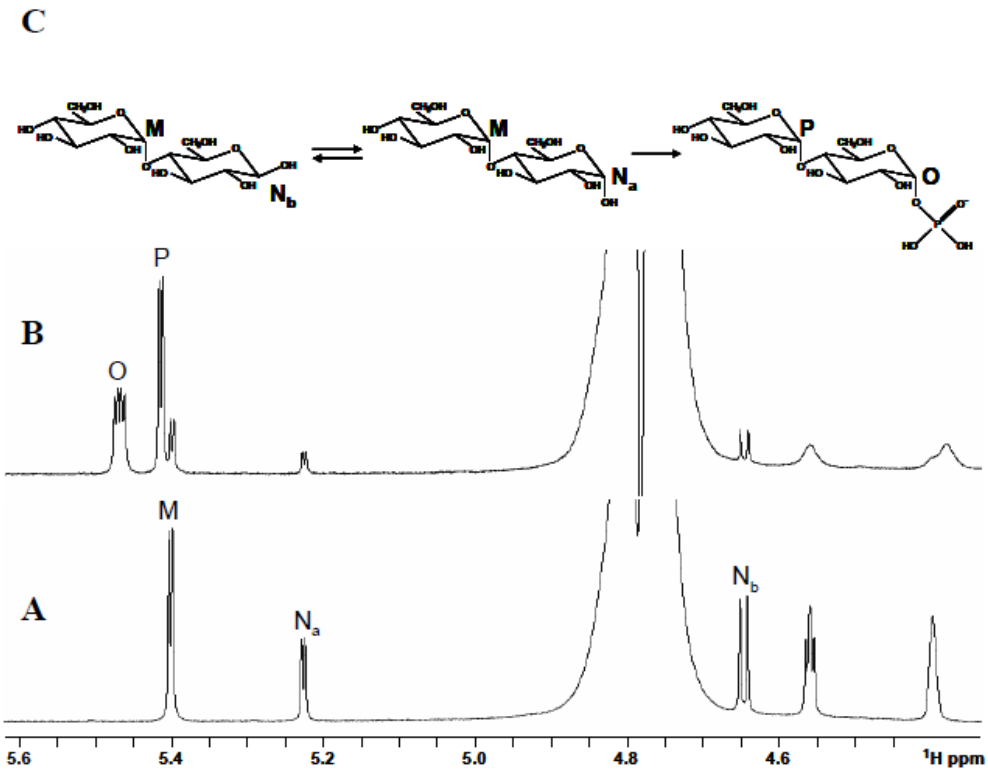



**Figure. 4.5:** $^{1}$H-NMR spectra of the Mak reaction mixture. Mixture contained 5 mM each ATP and maltose, 5 mM $MgCl_2$ in 10 mM BTP at pH 8, prior **(A)** and after **(B)** the addition of 15 µg of Mak and incubation for 10 min at 37ºC. The top panel **(C)** represents the structures of maltose and maltose-1-phosphate and is labeled according to their assigned resonances.

## Discussion

Maltose-1-phosphate was identified more than forty years ago in *Mycobacterium bovis* BCG (Narumi & Tsumita, 1967) the maltokinase, was only identified much later in *Actinoplanes missouriensis* (Drepper et al, 1996). Although the corresponding gene

(*mak*) can now be identified in many bacterial genomes, including those of most mycobacteria, no mycobacterial Mak has so far been characterized.

The *Mycobacterium tuberculosis* H37Rv *mak* gene (Rv0127) was considered essential for the pathogen's growth (Sassetti et al, 2003). This information and the absence of biochemical data on mycobacterial maltokinases prompted us to study the enzyme from the slow-growing species *M. bovis* BCG.

Our results show that, in contrast to the *A. missouriensis* and *Streptomyces coelicolor* enzymes, which use ATP as the sole phosphate donor (GTP and UTP are used at less than 3% by the *A. missouriensis* Mak and not used by the *S. coelicolor* enzyme), the *M. bovis* BCG maltokinase was able to use ATP, GTP and UTP with comparable efficiency. This substrate flexibility may reflect an absolute requirement for maltose-1-phosphate, corroborating the proposed essentiality of the *mak* gene in *M. tuberculosis* (Sassetti et al, 2003). However, lower flexibility of the Mak for NTP substrates may occur in vivo. Similar behavior has been detected between native and recombinant forms of enzymes involved in trehalose metabolism (Cardoso et al, 2007; Nobre et al, 2008). Interestingly some unprecedented residual activity (<7%) could also be detected with ADP and GDP as phosphate donors. Like the enzymes from *A. missouriensis* and *S. coelicolor*, only maltose served as the acceptor substrate. Although the $K_m$ values for maltose were similar for the *A. missouriensis* and *M. bovis* BCG Mak (no kinetic data is available for the *S. coelicolor* Mak), the later had a $K_m$ value for ATP (0.74 mM) slightly higher than that measured for the *A. missouriensis* Mak (0.54 mM).

A major obstacle for the characterization of the *M. bovis* BCG maltokinase was its very low stability. Our attempts to stabilize the enzyme with glycerol failed as it severely inhibited the activity. Unlike the native Mak from *A. missouriensis*, which could be stabilized by maltose, this disaccharide only moderately stabilized the mycobacterial Mak while trehalose was ineffective. Only the addition of 50 mM NaCl to the preparation stabilized the enzyme for characterization. The mycobacterial Mak

was, like the *A. missouriensis* and the *S. coelicolor* enzymes, a monomer in solution. While no data is available for the cation dependence of the *S. coelicolor* Mak, both the *A. missouriensis* and the *M. bovis* BCG enzymes were dependent on $Mg^{2+}$ ions for maximal activity. The cations $Co^{2+}$ or $Mn^{2+}$ could partially replace $Mg^{2+}$ while $Zn^{2+}$ was inhibitory. Divalent cations seem to play an essential role in the activity of kinases, namely in the stabilization of the negatively charged phosphate donor groups (Niehues et al, 2003). The temperature profile for activity of the recombinant Mak from *M. bovis* BCG was similar to that of the native Mak from *A. missouriensis*, with maximum near 60ºC. The *S. coelicolor* recombinant Mak had maximal activity at about 45ºC. All three enzymes had optimal activity between pH 7.0 and 9.0 (Jarling et al, 2004; Niehues et al, 2003), which is also the optimum pH range for the activity of other kinases (Niehues et al, 2003). The Mak from *M. bovis* BCG had no detectable aminoglycoside phosphotransferase activity, which is in agreement with the data obtained with the *S. coelicolor* Mak (Jarling et al, 2004). It is likely that many of the putative aminoglycoside phosphotransferases with significant amino acid identity with maltokinases (>40%) are incorrectly annotated and may instead have maltokinase activity.

The frequent association of the *mak* gene with the trehalose synthase (*treS*) gene, either as a bicystronic unit or as a fused gene, strongly suggested a synergistic action and a common biochemical pathway. While such a pathway has been previously proposed (Jarling et al, 2004), only recent research (Elbein et al, 2010; Kalscheuer et al, 2010) corroborated this hypothesis by demonstrating that *Mycobacterium smegmatis* and *M. tuberculosis* convert excess trehalose into glycogen and α-glucan through the sequential action of TreS, Mak and of a maltosyltransferase (GlgE) (Fig. 4.6). The latter enzyme was shown to use maltose-1-phosphate as the substrate for the transfer of maltose to glycogen and α-glucan but it's also capable of catalyzing the reverse reaction *in vitro* (Kalscheuer et al, 2010). The genetic organization of some organisms including *A. missouriensis* further show the relation of maltose-1-phosphate

and glycogen, since a glycogen-branching enzyme gene (*glgB*) is located immediately downstream the contiguous *treS* and *mak* genes. In *Rubrobacter xylanophilus*, a bacterium that constitutively accumulates high levels of trehalose, a *glgB* gene is also located immediately downstream a fused *treS-mak* bifunctional gene and a similar function may be anticipated. Interestingly, the *M. smegmatis* TreS was also found to possess amylase activity and capable to catalyze the synthesis of trehalose from glycogen, with maltose as intermediate (Pan et al, 2008). Curiously, *Mycobacterium leprae*, in which glycogen has not been reported, lacks both *treS* and *mak* homologues (Berg et al, 2007). The *glgE* homologue is also absent from this organism showing that this pathway is totally inactive in *M. leprae*. However the probable reason for the missing glycogen may be explained by the absence of other genes involved in glycogen synthesis like *glgB* or *glgP*, which are pseudogenes in *M. leprae*. On the other hand, both *Mycobacterium gilvum* and *Mycobacterium vanbaalenii* possess a *mak* and *glgE* but lack a *treS* gene. Unfortunately no data on trehalose, maltose or glycogen synthesis is available for these organisms. Nevertheless, among the mycobacterial enzymes involved in this metabolic circuit (trehalose $\rightarrow$ maltose $\rightarrow$ maltose-1-phosphate $\rightarrow$ glycogen $\rightarrow$ maltose $\rightarrow$ trehalose) (Fig. 6), the maltokinase was the only one lacking a biochemical study (De Smet et al, 2000; Elbein et al, 2010; Pan et al, 2004).

Mycobacteria possess three different pathways for the synthesis of trehalose (De Smet et al, 2000; Murphy et al, 2005). This disaccharide, constitutively present in the cytoplasm of mycobacterial cells, is an essential structural component of cell walls and its levels increase during osmotic and thermal stress (De Smet et al, 2000; Takayama et al, 2005; Woodruff et al, 2004). Although the TreS has much higher affinity for maltose than for trehalose (Pan et al, 2004), in most cases it seems to be involved in trehalose degradation (Murphy et al, 2005) and has only been implicated in trehalose biosynthesis in *M. smegmatis* (Woodruff et al, 2004). *M. smegmatis* possesses an essential trehalase that converts trehalose into glucose (Fig. 4.6) (Carroll

et al, 2007). Elbein and coworkers argue that it is energetically favorable to convert trehalose to glycogen by the recently discovered TreS-Mak-maltosyltransferase pathway (Fig. 4.6) (Elbein et al, 2010). However, normal ATP intracellular concentrations within the 1 to 10 mM range (Veech et al, 1979) are inhibitory of the maltosyltransferase (GlgE) activity (Elbein et al, 2010) but not of the Mak activity. These authors also propose that the role of this pathway is to convert excess trehalose into glycogen during stress conditions. Therefore, under non-stressing conditions, in which the ATP levels rise, maltose-1-phosphate cannot be the substrate for glycogen, which may be synthesized via ADP-glucose (Elbein et al, 2010).

It was shown that maltose-1-phosphate is also used to synthesize the α-glucan (Kalscheuer et al, 2010), but since mycobacteria contain other polysaccharides that contain an α-(1→4)-chain, namely the methylglucose lipopolysaccharide (MGLP), it is tempting to speculate that maltose-1-phosphate may also be used as donor of maltose units for those macromolecules, under specific conditions (Fig. 4.6) (Kalscheuer et al, 2010). This idea is toughened by the perceived redundancy of the Rv3032 pathway, a gene known to be involved in the synthesis of MGLP (Stadthagen et al, 2007), and of the TreS → Mak → GlgE pathway (Kalscheuer et al, 2010). Simultaneous inhibition of both pathways is lethal and further confirms their overlap and a possible involvement of Mak in MGLP synthesis (Kalscheuer et al, 2010). Based on the confirmed overlap of both pathways and on the ability of GlgE in transferring maltose-1-phosphate to maltotetraose and bigger α-(1→4)-glucose polymers (Fig 4.7), strongly suggest an involvement of the Mak and GlgE in the extension of MGLP.

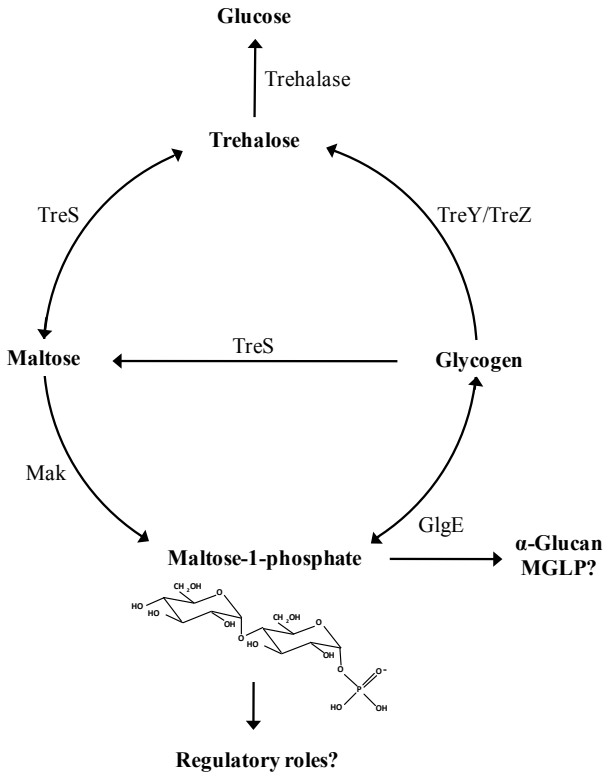



**Figure: 4.6:** Metabolic circuit between trehalose, maltose, glycogen and other α-glucans.

The constitutive expression of Mak in *A. missouriensis* (Niehues et al, 2003) and the suggested essentiality of the *M. tuberculosis mak* gene (Rv0127) may also indicate other alternative roles for maltose-1-phosphate in mycobacterial physiology. This is also supported by the fact that only the *mak* and *glgE* (Rv1327c) genes but not the *treS* gene (Rv0126), were considered essential for growth. Furthermore the non essentiality of the TreS → Mak → GlgE pathway, shown by TreS and GlgE simultaneous inactivation, also points to other possible roles beyond the α-glucans synthesis, since maltose-1-phosphate was still found in small amounts in *M. tuberculosis* TreS/GlgE double mutants, and therefore synthesized independently of this pathway (Kalscheuer

et al, 2010). Finally, the presence of a putative promoter within the terminal part of the *treS* gene indicating independent regulation of *mak* expression, once again suggests an alternative role for the Mak and maltose-1-phosphate in mycobacteria. Since trehalose-6-phosphate is an important signaling molecule in yeast and plant metabolism (Blazquez et al, 1993; Paul et al, 2008), maltose-1-phosphate could, hypothetically, play a similar role in mycobacteria. In fact, the involvement of maltose-1-phosphate in the regulation of sugar metabolism in *E. coli* has been proposed earlier but not confirmed experimentally (Decker et al, 1999).

Only now, the role(s) of maltose-1-phosphate in mycobacteria are starting to be revealed with many questions still open. The biochemical properties of the Mak from *M. bovis* BCG represent a significant step towards the elucidation of its role in mycobacterial metabolism. Furthermore, the fast and simple method for high-level production and purification of bioactive Mak from *M. bovis* BCG, a major bottleneck in mycobacterial research, will contribute to any structural studies that will precede any serious attempts to find inhibitors for this essential enzyme.

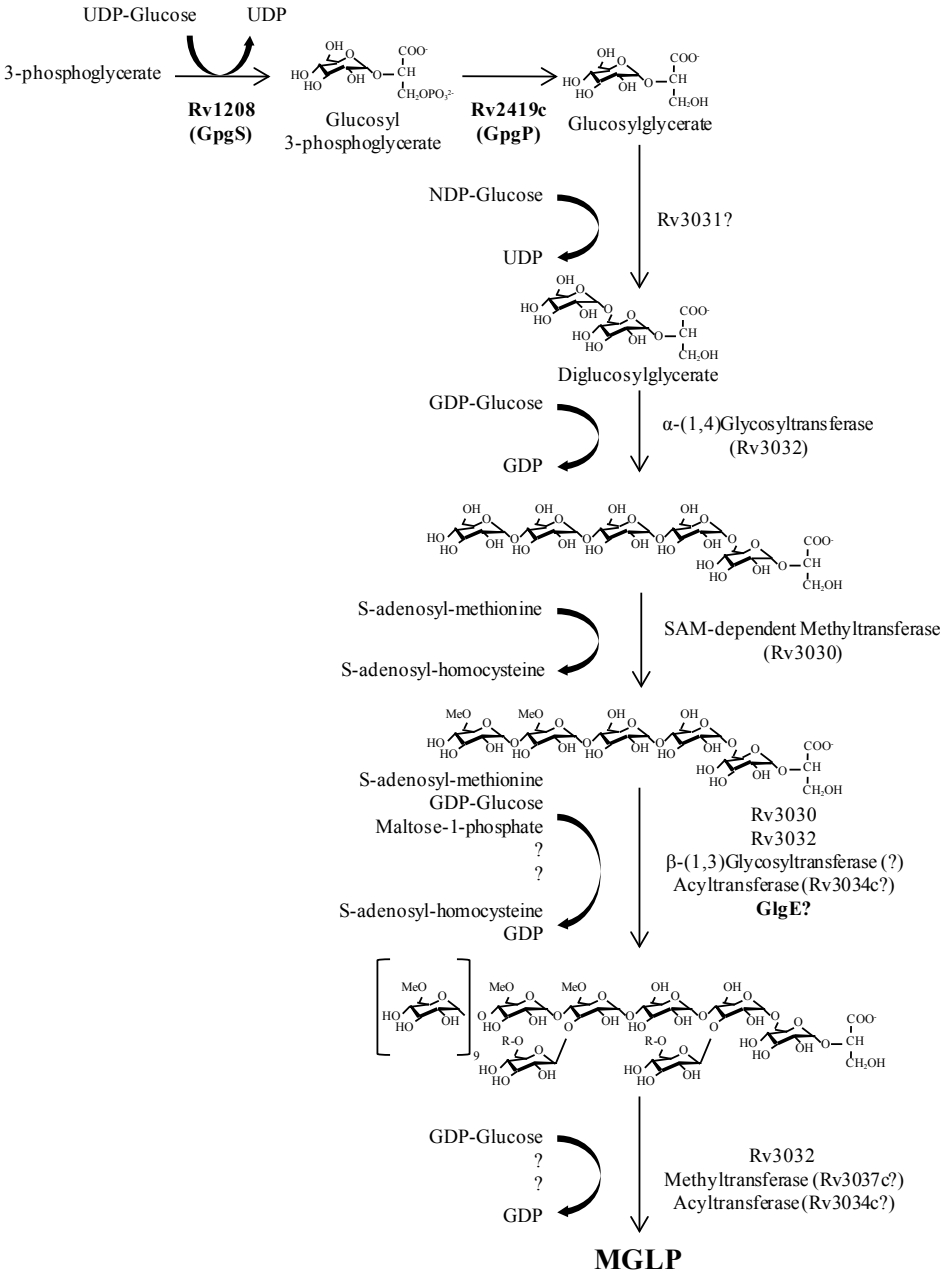



**Figure 4.7:** Proposed pathway for the synthesis of MGLP based on this thesis results and available knowledge. Partially adapted from (Jackson & Brennan, 2009)

## Acknowledgements and work contributions

Pedro Lamosa designed the NMR experiments and acquired the data. This work was supported by Fundação para a Ciência e a Tecnologia, FEDER, through start-up funding by Center for Neuroscience and Cell Biology (Associate Laboratory) to Nuno Empadinhas and Project PTDC/BIA-BCM/112459/2009.

# Chapter 5

## Concluding remarks

## GpgS and GpgP

Mycobacteria synthesize two polymethylated polysaccharides (PMPS), the methylmannose polysaccharide (MMP) and the methylglucose lipopolysaccharide (MGLP) (Jackson & Brennan, 2009). These polysaccharides form stable 1:1 complexes with fatty acids, modulating fatty acid synthesis and are virtually capable of forming these complexes with all the long chain fatty acids in the cytoplasm (Banis et al, 1977; Ilton et al, 1971; Yabusaki et al, 1979). MGLP has been isolated also from the extracellular medium in a culture of *M. smegmatis* suggesting that these structures are somehow extruded (Saier & Ballou, 1968). Much has been speculated about the PMPS role in transporting the fatty acid chains to the cell exterior for mycolic acid synthesis. However, the physiological evidence for this assumption remains to be experimentally confirmed.

Interestingly, only the MGLP has been found in the slow growing pathogenic mycobacteria (Stadthagen et al, 2007) and some of the genes involved in its biosynthesis have been considered essential for *Mycobacterium tuberculosis* growth (Sassetti et al, 2003), making this biosynthetic pathway an interesting target for study and possible drug development.

The MGLP structure was extensively studied and its biosynthesis has recently received some attention. However, several of the genes and enzymes involved in the biosynthetic pathways remained unknown.

The experience of our group with the synthesis of glucosylglycerate and the fact that MGLP contains glucosylglycerate at the reducing end prompted us to study the initial steps of the biosynthesis of this polysaccharide.

At the start of this work there was no knowledge about any of the genes responsible for MGLP biosynthesis and only some of the biosynthetic steps had been proposed (Kamisango et al, 1987). However, a few months laetr, a cluster of genes in *M. tuberculosis* was confirmed to be involved in MGLP synthesis. An α-(1→4)-

glucosyltransferase (Rv3032) and an essential SAM-dependent 6-*O*-methyltransferase (Rv3030) were experimentally confirmed to be involved in MGLP synthesis (Stadthagen et al, 2007). Another α-(1→4)-glucosyltransferase (Rv1212c) was also reported to be capable to functionally replace the Rv3032 (Stadthagen et al, 2007).

The present work reveals the two initial steps of MGLP synthesis and identifies the genes and enzymes responsible for those steps. We confirmed that the synthesis of the polysaccharide starts with the synthesis of glucosyl-3-phosphoglycerate (GPG), reaction catalyzed by the GpgS (Rv1208) (Empadinhas et al, 2008), which is followed by the subsequent dephosphorylation of GPG to glucosylglycerate, catalyzed by a GpgP (Rv2419c) (Mendes et al, unpublished). Both genes are present in all sequenced mycobacteria (Mendes et al, unpublished; Empadinhas et al, 2008) but while the GpgS gene was considered essential for *M. tuberculosis* growth, that was not the case for the GpgP gene (Sassetti et al, 2003). This suggests that one or more of the other GpgP homologues found in the *M. tuberculosis* genome or even an unspecific phosphatase could potentially exhibit GpgP activity and replace this enzyme's function. Nevertheless, this activity was not detected during purification of the native GpgP from *M. hassiacum* and *M. vanbaalenii* cell extracts and, in case it exists, it may hypothetically be expressed under specific circumstances and at lower rate. Interestingly, unlike all the other previously characterized GpgPs/MpgPs that belong to the HAD-like hydrolase superfamily, the mycobacterial GpgP belongs to the histidine phosphatase superfamily. The mycobacterial GpgP was annotated as a putative phosphoglycerate mutase (PGM). However, we have shown that this enzyme lacks such activity. While several other *M. tuberculosis* proteins shared some level of similarity to the *M. tuberculosis* GpgP and were also annotated as putative PGMs, only one (Rv0489) seems to be a likely candidate for that role (Muller et al, 2005). The widespread mis-annotation issues in the histidine phosphatase superfamily seem to be caused by the mutase-dominated superfamily record, while the phosphatase activity is actually the dominant (Rigden, 2008). The identification of Rv2419c as a

GpgP will have an impact of many of the available bacterial genomes and significantly contributes to expanding the knowledge and broadening the activities on the histidine phosphatase superfamily.

While the genes *gpgS* and *gpgP* are often arranged in an operon-like structure in the organisms that harbour them (Costa et al, 2007; Costa et al, 2006) this is not the case for mycobacterial species and related organisms harboring the histidine phosphatase superfamily GpgP. The existence of two types of GpgPs probably reflects a situation of two non-homologous isofunctional enzymes that had a different ancestor. Structural studies on both types of GpgPs will provide further knowledge about the evolution of both enzymes and will bring a definitive answer on whether this is truly a case of convergent evolution.

The widespread occurrence of MGLP in pathogenic mycobacteria may suggest a vital role in mycobacterial physiology. It is therefore likely that their inability to properly assemble this structure could hamper lipid synthesis and consequently cell wall assembly leading to cell death. This hypothesis is further reinforced by the proposed essentiality of some of the genes involved in MGLP synthesis (Sassetti et al, 2003). The data gathered in this work not only unraveled the two first steps in the synthesis of MGLP, but also expanded the knowledge on the enzyme families that were studied.

## Maltokinase

*M. bovis* BCG was the first organism where maltose-1-phosphate was identified (Narumi & Tsumita, 1967) but an enzyme for the synthesis of this compound (maltokinase, Mak) was only later identified in another actinobacteria, *A. missouriensis* (Drepper et al, 1996). Interestingly, the *mak* gene (Rv0127) was considered essential for the growth of *M. tuberculosis* H37Rv (Sassetti et al, 2003). The phosphate donor flexibility we found in the mycobacterial Mak, unlike the other

characterized Maks, may reflect an utter necessity for maltose-1-phosphate, reinforcing the essentiality of the *mak* gene in *M. tuberculosis* (Sassetti et al, 2003). Recently, the accumulation of maltose-1-phosphate was found to be toxic and to trigger pleiotropic stress responses that lead to the death of *M. tuberculosis* (Kalscheuer et al, 2010).

The frequent association of the *mak* gene with the trehalose synthase (*treS*) gene, throughout the bacterial world, strongly indicates a common biochemical pathway. It was later found that *M. smegmatis* and *M. tuberculosis* convert excess trehalose into glycogen through the action of TreS, Mak and of a maltosyltransferase (GlgE) (Elbein et al, 2010; Kalscheuer et al, 2010). The GlgE is capable of using maltose-1-phosphate as the donor of maltose blocks not only to glycogen but also to other α-(1→4)-glucose chains with 4 or more glucose units (Kalscheuer et al, 2010). Given that mycobacteria contain other α-(1→4)-glucans and polysaccharides like the MGLP, this activity hints that maltose-1-phosphate could also be used as donor of maltose units for those macromolecules (Kalscheuer et al, 2010; Mendes et al, 2010). Moreover the overlap of the TreS → Mak → GlgE and Rv3032 pathways further suggests a role of maltose-1-phosphate in MGLP synthesis (Kalscheuer et al, 2010).

Maltose-1-phosphate has been proposed to be involved in the regulation of sugar metabolism in *E. coli* but this has not been experimentally confirmed (Decker et al, 1999). On the other hand, trehalose-6-phosphate, a molecule similar to maltose-1-phosphate, was found to be an important signaling molecule in yeast and plant metabolism (Blazquez et al, 1993; Paul et al, 2008). Genetic hints yet to be experimentaly confirmed suggest independent expression of the *treS* and *mak* genes in *M. tuberculosis*, *M. bovis* and *M. marinum* (Mendes et al, 2010). Furthermore, the non essentiality of the TreS → Mak → GlgE pathway, shown by TreS inactivation, also points to other possible roles beyond the α-glucans synthesis (Kalscheuer et al, 2010; Mendes et al, 2010). Collectively, this suggests a possible role of maltose-1-phosphate

as a regulator/signaling molecule. Nevertheless, additional studies will have to be performed to confirm this hypothesis.

Several questions remain open and the role(s) of maltose-1-phosphate in mycobacterial metabolism only now begins to be unraveled. While further studies are required to answer these questions, the biochemical properties of the maltokinase from *Mycobacterium bovis* BCG, the first of this bacterial group to be characterized, provide an important step towards the clarification of its role in mycobacterial metabolism.

In summary, the findings in this thesis significantly contribute to our understanding of the molecular players operating in *M. tuberculosis* metabolism through the identification of new functions in the genome of this pathogen, and further provide an important basis to help us devise new strategies to combat this insidious pathogen and, ultimately, tuberculosis.

# References

Ames BN (1966). Assay of inorganic phosphate, total phosphate and phosphatases. *Methods Enzymol* **8:** 115–118

Azuma I, Takeda K, Yamamura Y, Yanagihara Y, Mifuchi I (1976). Occurrence of 3-O-methylmannose in the polysaccharide of *Leptospira biflexa* Urawa. *J Bacteriol* **128**(1)**:** 492-494

Balganesh TS, Alzari PM, Cole ST (2008). Rising standards for tuberculosis drug development. *Trends Pharmacol Sci* **29**(11)**:** 576-581

Balows A (1992). The Prokaryotes: a handbook on the biology of bacteria: ecophysiology, isolation, identification, applications, New York: Springer-Verlag.

Banis RJ, Peterson DO, Bloch K (1977). *Mycobacterium smegmatis* fatty acid synthetase. Polysaccharide stimulation of the rate-limiting step. *J Biol Chem* **252**(16)**:** 5740-5744

Barry CE, Crick DC, McNeil MR (2007). Targeting the formation of the cell wall core of *Mycobacterium tuberculosis*. *Infect Disord Drug Targets* **7**(2)**:** 182-202

Belanger AE, Besra GS, Ford ME, Mikusova K, Belisle JT, Brennan PJ, Inamine JM (1996). The embAB genes of *Mycobacterium avium* encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. *Proc Natl Acad Sci U S A* **93**(21)**:** 11919-11924

Belanger AE, Hatfull GF (1999). Exponential-phase glycogen recycling is essential for growth of *Mycobacterium smegmatis*. *J Bacteriol* **181**(21)**:** 6670-6678

Belisle JT, Mahaffey SB, Hill PJ (2009). Isolation of mycobacterium species genomic DNA. *Methods Mol Biol* **465:** 1-12

Berg S, Kaur D, Jackson M, Brennan PJ (2007). The glycosyltransferases of *Mycobacterium tuberculosis* - roles in the synthesis of arabinogalactan, lipoarabinomannan, and other glycoconjugates. *Glycobiology* **17**(6)**:** 35-56R

Bergmeyer HU (1974). Methods of enzymatic analysis, 2d English edn, Weinheim New York: Verlag Chemie; Academic Press.

Besra GS, Brennan PJ (1997). The mycobacterial cell wall: biosynthesis of arabinogalactan and lipoarabinomannan. *Biochem Soc Trans* **25**(3)**:** 845-850

Bhamidi S, Scherman MS, Rithner CD, Prenni JE, Chatterjee D, Khoo KH, McNeil MR (2008). The identification and location of succinyl residues and the characterization of the interior arabinan region allow for a model of the complete primary structure of *Mycobacterium tuberculosis* mycolyl arabinogalactan. *J Biol Chem* **283**(19)**:** 12992-13000

Blazquez MA, Lagunas R, Gancedo C, Gancedo JM (1993). Trehalose-6-phosphate, a new regulator of yeast glycolysis that inhibits hexokinases. *FEBS Lett* **329**(1-2)**:** 51-54

Bloch K (1977). Control mechanisms for fatty acid synthesis in *Mycobacterium smegmatis*. *Adv Enzymol Relat Areas Mol Biol* **45:** 1-84

Bloch K, Vance D (1977). Control mechanisms in the synthesis of saturated fatty acids. *Annu Rev Biochem* **46:** 263-298

Bradford MM (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72:** 248-254

Brennan P, Ballou CE (1967). Biosynthesis of mannophosphoinositides by *Mycobacterium phlei*. The family of dimannophosphoinositides. *J Biol Chem* **242**(13)**:** 3046-3056

Brennan PJ (2003). Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* **83**(1-3)**:** 91-97

Cardoso FS, Castro RF, Borges N, Santos H (2007). Biochemical and genetic characterization of the pathways for trehalose metabolism in *Propionibacterium freudenreichii*, and their role in stress response. *Microbiology* **153**(Pt 1)**:** 270-280

Carroll JD, Pastuszak I, Edavana VK, Pan YT, Elbein AD (2007). A novel trehalase from *Mycobacterium smegmatis* - purification, properties, requirements. *FEBS J* **274**(7)**:** 1701-1714

Chapman GB, Hanks JH, Wallace JH (1959). An electron microscope study of the disposition and fine structure of *Mycobacterium lepraemurium* in mouse spleen. *J Bacteriol* **77**(2)**:** 205-211

Costa J, Empadinhas N, da Costa MS (2007). Glucosylglycerate biosynthesis in the deepest lineage of the Bacteria: characterization of the thermophilic proteins GpgS and GpgP from *Persephonella marina*. *J Bacteriol* **189**(5)**:** 1648-1654

Costa J, Empadinhas N, Goncalves L, Lamosa P, Santos H, da Costa MS (2006). Characterization of the biosynthetic pathway of glucosylglycerate in the archaeon *Methanococcoides burtonii*. *J Bacteriol* **188**(3)**:** 1022-1030

Cywes C, Godenir NL, Hoppe HC, Scholle RR, Steyn LM, Kirsch RE, Ehlers MR (1996). Nonopsonic binding of *Mycobacterium tuberculosis* to human complement receptor type 3 expressed in Chinese hamster ovary cells. *Infect Immun* **64**(12)**:** 5373-5383

Daffé M, Draper P (1998). The envelope layers of mycobacteria with reference to their pathogenicity. *Adv Microb Physiol* **39:** 131-203

Daffé M, Etienne G (1999). The capsule of *Mycobacterium tuberculosis* and its implications for pathogenicity. *Tuber Lung Dis* **79**(3)**:** 153-169

Daffé M, Reyrat J-M (2008). The mycobacterial cell envelope, Washington, DC: ASM Press.

De Smet KA, Weston A, Brown IN, Young DB, Robertson BD (2000). Three pathways for trehalose biosynthesis in mycobacteria. *Microbiology* **146 ( Pt 1):** 199-208

Decker K, Gerhardt F, Boos W (1999). The role of the trehalose system in regulating the maltose regulon of *Escherichia coli*. *Mol Microbiol* **32**(4)**:** 777-788

Dinadayala P, Lemassu A, Granovski P, Cerantola S, Winter N, Daffe M (2004). Revisiting the structure of the anti-neoplastic glucans of *Mycobacterium bovis* Bacille Calmette-Guerin. Structural analysis of the extracellular and boiling water extract-derived glucans of the vaccine substrains. *J Biol Chem* **279**(13)**:** 12369-12378

Drepper A, Peitzmann R, Pape H (1996). Maltokinase (ATP:maltose 1-phosphotransferase) from *Actinoplanes sp.*: demonstration of enzyme activity and characterization of the reaction product. *FEBS Lett* **388**(2-3)**:** 177-179

Dworkin M, Falkow S (2006). The prokaryotes: a handbook on the biology of bacteria, 3rd edn. New York; [London]: Springer.

Dye C (2009). Doomsday postponed? Preventing and reversing epidemics of drug-resistant tuberculosis. *Nat Rev Microbiol* **7**(1)**:** 81-87

Elbein AD, Pastuszak I, Tackett AJ, Wilson T, Pan YT (2010). Last Step in the Conversion of Trehalose to Glycogen: a mycobacterial enzyme that transfers maltose from maltose 1-phosphate to glycogen. *Journal of Biological Chemistry* **285**(13)**:** 9803-9812

Empadinhas N, Albuquerque L, Henne A, Santos H, da Costa MS (2003). The bacterium *Thermus thermophilus*, like hyperthermophilic archaea, uses a two-step pathway for the synthesis of mannosylglycerate. *Appl Environ Microbiol* **69**(6)**:** 3272-3279

Empadinhas N, Albuquerque L, Mendes V, Macedo-Ribeiro S, da Costa MS (2008). Identification of the mycobacterial glucosyl-3-phosphoglycerate synthase. *FEMS Microbiol Lett* **280**(2)**:** 195-202

Empadinhas N, Marugg JD, Borges N, Santos H, da Costa MS (2001). Pathway for the synthesis of mannosylglycerate in the hyperthermophilic archaeon *Pyrococcus horikoshii*. Biochemical and genetic characterization of key enzymes. *J Biol Chem* **276**(47)**:** 43580-43588

Empadinhas N, Mendes V, Simoes C, Santos MS, Mingote A, Lamosa P, Santos H, Costa MS (2007). Organic solutes in *Rubrobacter xylanophilus*: the first example of di-myo-inositol-phosphate in a thermophile. *Extremophiles* **11**(5)**:** 667-673

Empadinhas N, Pereira PJ, Albuquerque L, Costa J, Sa-Moura B, Marques AT, Macedo-Ribeiro S, da Costa MS (2011). Functional and structural characterization of a novel mannosyl-3-phosphoglycerate synthase from *Rubrobacter xylanophilus* reveals its dual substrate specificity. *Mol Microbiol* **79**(1)**:** 76-93

Ferguson JA, Ballou CE (1970). Biosynthesis of a mycobacterial lipopolysaccharide. Properties of the polysaccharide methyltransferase. *J Biol Chem* **245**(16)**:** 4213-4223

Fernandes C, Empadinhas N, da Costa MS (2007). Single-step pathway for synthesis of glucosylglycerate in *Persephonella marina*. *J Bacteriol* **189**(11)**:** 4014-4019

Fernandes C, Mendes V, Costa J, Empadinhas N, Jorge C, Lamosa P, Santos H, da Costa MS (2010). Two Alternative Pathways for the Synthesis of the Rare Compatible Solute Mannosylglucosylglycerate in *Petrotoga mobilis*. *Journal of Bacteriology* **192**(6)**:** 1624-1633

Flick PK, Bloch K (1974). Vitro alterations of the product distribution of the fatty synthetase from *Mycobacterium phlei*. *J Biol Chem* **249**(4)**:** 1031-1036

Flick PK, Bloch K (1975). Reversible inhibition of the fatty acid synthetase complex from *Mycobacterium smegmatis* by palmitoyl-coenzyme A. *J Biol Chem* **250**(9)**:** 3348-3351

Forsberg LS, Dell A, Walton DJ, Ballou CE (1982). Revised structure for the 6-O-methylglucose polysaccharide of *Mycobacterium smegmatis*. *J Biol Chem* **257**(7)**:** 3555-3563

Foster JM, Davis PJ, Raverdy S, Sibley MH, Raleigh EA, Kumar S, Carlow CK (2010). Evolution of bacterial phosphoglycerate mutases: non-homologous isofunctional enzymes undergoing gene losses, gains and lateral transfers. *PLoS ONE* **5**(10)**:** e13576

Fothergill-Gilmore LA, Watson HC (1989). The phosphoglycerate mutases. *Adv Enzymol Relat Areas Mol Biol* **62:** 227-313

Frehel C, Ryter A, Rastogi N, David H (1986). The electron-transparent zone in phagocytized *Mycobacterium avium* and other mycobacteria: formation, persistence and role in bacterial survival. *Ann Inst Pasteur Microbiol* **137B**(3)**:** 239-257

Friedberg I (2006). Automated protein function prediction--the genomic challenge. *Brief Bioinform* **7**(3)**:** 225-242

Gagliardi MC, Lemassu A, Teloni R, Mariotti S, Sargentini V, Pardini M, Daffe M, Nisini R (2007). Cell wall-associated alpha-glucan is instrumental for *Mycobacterium tuberculosis* to block CD1 molecule expression and disable the function of dendritic cell derived from infected monocyte. *Cell Microbiol* **9**(8)**:** 2081-2092

Gao LY, Guo S, McLaughlin B, Morisaki H, Engel JN, Brown EJ (2004). A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion. *Mol Microbiol* **53**(6)**:** 1677-1693

Geurtsen J, Chedammi S, Mesters J, Cot M, Driessen NN, Sambou T, Kakutani R, Ummels R, Maaskant J, Takata H, Baba O, Terashima T, Bovin N, Vandenbroucke-Grauls CM, Nigou J, Puzo G, Lemassu A, Daffe M, Appelmelk BJ (2009). Identification of mycobacterial alpha-glucan as a novel ligand for DC-SIGN: involvement of mycobacterial capsular polysaccharides in host immune modulation. *J Immunol* **183**(8)**:** 5221-5231

Goude R, Renaud S, Bonnassie S, Bernard T, Blanco C (2004). Glutamine, glutamate, and alpha-glucosylglycerate are the major osmotic solutes accumulated by *Erwinia chrysanthemi* strain 3937. *Appl Environ Microbiol* **70**(11)**:** 6535-6541

Gray GR, Ballou CE (1971). Isolation and characterization of a polysaccharide containing 3-O-methyl-D-mannose from *Mycobacterium phlei*. *J Biol Chem* **246**(22)**:** 6835-6842

Gray GR, Ballou CE (1972). The 6-O-methylglucose-containing lipopolysaccharides of *Mycobacterium phlei*. Locations of the acyl groups. *J Biol Chem* **247**(24)**:** 8129-8135

Grellert E, Ballou CE (1972). Biosynthesis of a mycobacterial lipopolysaccharide. Evidence for an acylpolysaccharide methyltransferase. *J Biol Chem* **247**(10)**:** 3236-3241

Harris LS, Gray GR (1977). Acetylated methylmannose polysaccharide of *Streptomyces*. *J Biol Chem* **252**(8)**:** 2470-2477

Hett EC, Rubin EJ (2008). Bacterial growth and cell division: a mycobacterial perspective. *Microbiol Mol Biol Rev* **72**(1)**:** 126-156, table of contents

Hindsgaul O, Ballou CE (1984). Affinity purification of mycobacterial polymethyl polysaccharides and a study of polysaccharide-lipid interactions by 1H NMR. *Biochemistry* **23**(3)**:** 577-584

Hoffmann C, Leis A, Niederweis M, Plitzko JM, Engelhardt H (2008). Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc Natl Acad Sci U S A* **105**(10)**:** 3963-3967

Hunter RL, Olsen MR, Jagannath C, Actor JK (2006). Multiple roles of cord factor in the pathogenesis of primary, secondary, and cavitary tuberculosis, including a revised description of the pathology of secondary disease. *Ann Clin Lab Sci* **36**(4)**:** 371-386

Ilton M, Jevans AW, McCarthy ED, Vance D, White HB, 3rd, Bloch K (1971). Fatty acid synthetase activity in *Mycobacterium phlei*: regulation by polysaccharides. *Proc Natl Acad Sci U S A* **68**(1)**:** 87-91

Iturriaga G, Suarez R, Nova-Franco B (2009). Trehalose metabolism: from osmoprotection to signaling. *Int J Mol Sci* **10**(9)**:** 3793-3810

Jackson M, Brennan PJ (2009). Polymethylated polysaccharides from *Mycobacterium* species revisited. *J Biol Chem* **284**(4)**:** 1949-1953

Jarling M, Cauvet T, Grundmeier M, Kuhnert K, Pape H (2004). Isolation of mak1 from *Actinoplanes missouriensis* and evidence that Pep2 from *Streptomyces coelicolor* is a maltokinase. *J Basic Microbiol* **44**(5)**:** 360-373

Jedrzejas MJ (2000). Structure, function, and evolution of phosphoglycerate mutases: comparison with fructose-2,6-bisphosphatase, acid phosphatase, and alkaline phosphatase. *Prog Biophys Mol Biol* **73**(2-4)**:** 263-287

Jorge CD, Lamosa P, Santos H (2007). Alpha-D-mannopyranosyl-(1-->2)-alpha-D-glucopyranosyl-(1-->2)-glycerate in the thermophilic bacterium *Petrotoga miotherma* - structure, cellular content and function. *FEBS J* **274**(12)**:** 3120-3127

Kalscheuer R, Syson K, Veeraraghavan U, Weinrick B, Biermann KE, Liu Z, Sacchettini JC, Besra G, Bornemann S, Jacobs WR (2010). Self-poisoning of *Mycobacterium tuberculosis* by targeting GlgE in an alpha-glucan pathway. *Nature Chemical Biology* **6**(5)**:** 376-384

Kamisango K, Dell A, Ballou CE (1987). Biosynthesis of the mycobacterial O-methylglucose lipopolysaccharide. Characterization of putative intermediates in the initiation, elongation, and termination reactions. *J Biol Chem* **262**(10)**:** 4580-4586

Kaur D, Guerin ME, Skovierova H, Brennan PJ, Jackson M (2009a). Chapter 2: Biogenesis of the cell wall and other glycoconjugates of *Mycobacterium tuberculosis*. *Adv Appl Microbiol* **69:** 23-78

Kaur D, Pham H, Larrouy-Maumus G, Riviere M, Vissa V, Guerin ME, Puzo G, Brennan PJ, Jackson M (2009b). Initiation of methylglucose lipopolysaccharide biosynthesis in mycobacteria. *PLoS ONE* **4**(5)**:** e5447

Kawaguchi A, Bloch K (1974). Inhibition of glucose 6-phosphate dehydrogenase by palmitoyl coenzyme A. *J Biol Chem* **249**(18)**:** 5793-5800

Kawaguchi A, Bloch K (1976). Inhibition of glutamate dehydrogenase and malate dehydrogenases by palmitoyl coenzyme A. *J Biol Chem* **251**(5)**:** 1406-1412

Klahn S, Steglich C, Hess WR, Hagemann M (2010). Glucosylglycerate: a secondary compatible solute common to marine cyanobacteria from nitrogen-poor environments. *Environmental Microbiology* **12**(1)**:** 83-94

Knoche H, Esders TW, Koths K, Bloch K (1973). Palmityl coenzyme A inhibition of fatty acid synthesis. Relief by bovine serum albumin and mycobacterial polysaccharides. *J Biol Chem* **248**(7)**:** 2317-2322

Kojima N, Uchikawa K, Araki Y, Ito E (1985). A common linkage saccharide unit between teichoic acids and peptidoglycan in cell walls of *Bacillus coagulans*. *J Biochem* **97**(4)**:** 1085-1092

Kubica GP, Wayne LG (1984). The Mycobacteria: a sourcebook, New York: Dekker.

Lee YC (1966). Isolation and characterization of lipopolysaccharides containing 6-O-methyl-D-glucose from *Mycobacterium* species. *J Biol Chem* **241**(8)**:** 1899-1908

Lee YC, Ballou CE (1964). 6-O-Methyl-D-Glucose from Mycobacteria. *J Biol Chem* **239**(10)**:** PC3602-3603

Lemassu A, Daffé M (1994). Structural features of the exocellular polysaccharides of *Mycobacterium tuberculosis*. *Biochem J* **297 ( Pt 2):** 351-357

Lornitzo FA, Goldman DS (1968). Intracellular localization of a 6-O-methyl-D-glucose containing soluble polysaccharide from *Mycobacterium tuberculosis*. *Biochim Biophys Acta* **158**(3)**:** 329-335

Machida Y, Bloch K (1973). Complex formation between mycobacterial polysaccharides and fatty acyl-CoA derivatives. *Proc Natl Acad Sci U S A* **70**(4)**:** 1146-1148

Maitra SK, Ballou CE (1977). Heterogeneity and refined structtures of 3-O-methyl-D-mannose polysaccharides from *Mycobacterium smegmatis*. *J Biol Chem* **252**(8)**:** 2459-2469

Makarov V, Manina G, Mikusova K, Mollmann U, Ryabova O, Saint-Joanis B, Dhar N, Pasca MR, Buroni S, Lucarelli AP, Milano A, De Rossi E, Belanova M, Bobovska A, Dianiskova P, Kordulakova J, Sala C, Fullam E, Schneider P, McKinney JD, Brodin P, Christophe T, Waddell S, Butcher P, Albrethsen J, Rosenkrands I, Brosch R, Nandi V, Bharath S, Gaonkar S, Shandil RK, Balasubramanian V, Balganesh T, Tyagi S, Grosset J, Riccardi G, Cole ST (2009). Benzothiazinones kill *Mycobacterium tuberculosis* by blocking arabinan synthesis. *Science* **324**(5928)**:** 801-804

Maloney DH, Ballou CE (1980). Polymethylpolysaccharide synthesis in an ethionine-resistant mutant of *Mycobacterium smegmatis*. *J Bacteriol* **141**(3)**:** 1217-1221

Matsuhashi M (1966). [Biosynthesis in the bacterial cell wall]. *Tanpakushitsu Kakusan Koso* **11**(10)**:** 875-886

Mendes V, Maranha A, Lamosa P, da Costa MS, Empadinhas N (2010). Biochemical characterization of the maltokinase from *Mycobacterium bovis* BCG. *BMC Biochem* **11:** 21

Mita M, Yasumasu I (1981). Inhibition of dihydrofolate reductase by palmitoyl coenzyme A. *Int J Biochem* **13**(2)**:** 229-232

Muller P, Sawaya MR, Pashkov I, Chan S, Nguyen C, Wu Y, Perry LJ, Eisenberg D (2005). The 1.70 angstroms X-ray crystal structure of *Mycobacterium tuberculosis* phosphoglycerate mutase. *Acta Crystallogr D Biol Crystallogr* **61**(Pt 3)**:** 309-315

Murphy HN, Stewart GR, Mischenko VV, Apt AS, Harris R, McAlister MS, Driscoll PC, Young DB, Robertson BD (2005). The OtsAB pathway is essential for trehalose biosynthesis in *Mycobacterium tuberculosis*. *J Biol Chem* **280**(15)**:** 14524-14529

Narumi K, Tsumita T (1967). Identification of alpha,alpha-trehalose 6,6'-dimannosylphosphate and alpha-maltose 1-phosphate of Mycobacteria. *J Biol Chem* **242**(9)**:** 2233-2239

Niehues B, Jossek R, Kramer U, Koch A, Jarling M, Schroder W, Pape H (2003). Isolation and characterization of maltokinase (ATP : maltose 1-phosphotransferase) from *Actinoplanes missouriensis*. *Archives of Microbiology* **180**(4)**:** 233-239

Nimmich W (1970). Occurrence of 3-O-methylmannose in lipopolysaccharides of *Klebsiella* and *Escherichia coli*. *Biochim Biophys Acta* **215**(1)**:** 189-191

Nobre A, Alarico S, Fernandes C, Empadinhas N, da Costa MS (2008). A unique combination of genetic systems for the synthesis of trehalose in *Rubrobacter xylanophilus*: properties of a rare actinobacterial TreT. *J Bacteriol*

Ortalo-Magne A, Dupont MA, Lemassu A, Andersen AB, Gounon P, Daffe M (1995). Molecular composition of the outermost capsular material of the tubercle bacillus. *Microbiology* **141 ( Pt 7):** 1609-1620

Pan YT, Carroll JD, Asano N, Pastuszak I, Edavana VK, Elbein AD (2008). Trehalose synthase converts glycogen to trehalose. *FEBS J* **275**(13)**:** 3408-3420

Pan YT, Koroth Edavana V, Jourdian WJ, Edmondson R, Carroll JD, Pastuszak I, Elbein AD (2004). Trehalose synthase of *Mycobacterium smegmatis*: purification, cloning, expression, and properties of the enzyme. *Eur J Biochem* **271**(21)**:** 4259-4269

Paul MJ, Primavesi LF, Jhurreea D, Zhang Y (2008). Trehalose metabolism and signaling. *Annu Rev Plant Biol* **59:** 417-441

Paul TR, Beveridge TJ (1994). Preservation of surface lipids and determination of ultrastructure of *Mycobacterium kansasii* by freeze-substitution. *Infect Immun* **62**(5)**:** 1542-1550

Pavelka MS, Jr. (2007). Another brick in the wall. *Trends Microbiol* **15**(4)**:** 147-149

Pereira PJ, Empadinhas N, Albuquerque L, Sa-Moura B, da Costa MS, Macedo-Ribeiro S (2008). *Mycobacterium tuberculosis* glucosyl-3-phosphoglycerate synthase: structure of a key enzyme in methylglucose lipopolysaccharide biosynthesis. *PLoS ONE* **3**(11)**:** e3748

Peterson DO, Bloch K (1977). *Mycobacterium smegmatis* fatty acid synthetase. Long chain transacylase chain length specificity. *J Biol Chem* **252**(16)**:** 5735-5739

Pitarque S, Larrouy-Maumus G, Payre B, Jackson M, Puzo G, Nigou J (2008). The immunomodulatory lipoglycans, lipoarabinomannan and lipomannan, are exposed at the mycobacterial cell surface. *Tuberculosis (Edinb)* **88**(6)**:** 560-565

Pommier MT, Michel G (1986). Isolation and characterization of an O-methylglucose-containing lipopolysaccharide produced by *Nocardia otitidis-caviarum*. *J Gen Microbiol* **132**(9)**:** 2433-2441

Portevin D, de Sousa-D'Auria C, Montrozier H, Houssin C, Stella A, Laneelle MA, Bardou F, Guilhot C, Daffe M (2005). The acyl-AMP ligase FadD32 and AccD4-containing acyl-CoA carboxylase are required for the synthesis of mycolic acids and essential for mycobacterial growth: identification of the carboxylation product and determination of the acyl-CoA carboxylase components. *J Biol Chem* **280**(10)**:** 8862-8874

Postma PW, Lengeler JW, Jacobson GR (1993). Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev* **57**(3)**:** 543-594

Pym AS, Brodin P, Brosch R, Huerre M, Cole ST (2002). Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Mol Microbiol* **46**(3)**:** 709-717

Qu Q, Lee SJ, Boos W (2004). TreT, a novel trehalose glycosyltransferring synthase of the hyperthermophilic archaeon *Thermococcus litoralis*. *J Biol Chem* **279**(46)**:** 47890-47897

Rafidinarivo E, Laneelle MA, Montrozier H, Valero-Guillen P, Astola J, Luquin M, Prome JC, Daffe M (2008). Trafficking pathways of mycolic acids: structures, origin, mechanism of formation and storage form of mycobacteric acids. *J Lipid Res*

Rainey FA, Ward-Rainey N, Kroppenstedt RM, Stackebrandt E (1996). The genus Nocardiopsis represents a phylogenetically coherent taxon and a distinct actinomycete lineage: proposal of Nocardiopsaceae fam. nov. *Int J Syst Bacteriol* **46**(4)**:** 1088-1092

Raynaud C, Etienne G, Peyron P, Laneelle MA, Daffe M (1998). Extracellular enzyme activities potentially involved in the pathogenicity of *Mycobacterium tuberculosis*. *Microbiology* **144 ( Pt 2):** 577-587

Rigden DJ (2008). The histidine phosphatase superfamily: structure and function. *Biochem J* **409**(2)**:** 333-348

Riley LW (2006). Of mice, men, and elephants: *Mycobacterium tuberculosis* cell envelope lipids and pathogenesis. *J Clin Invest* **116**(6)**:** 1475-1478

Robertson DE, Lai MC, Gunsalus RP, Roberts MF (1992). Composition, Variation, and Dynamics of Major Osmotic Solutes in *Methanohalophilus* Strain FDF1. *Appl Environ Microbiol* **58**(8)**:** 2438-2443

Saier MH, Jr., Ballou CE (1968). The 6-O-methylglucose-containing lipopolysaccharide of *Mycobacterium phlei*. Identification of D-glyceric acid and 3-O-methyl-D-glucose in the polysaccharide. *J Biol Chem* **243**(5)**:** 992-1005

Sambou T, Dinadayala P, Stadthagen G, Barilone N, Bordat Y, Constant P, Levillain F, Neyrolles O, Gicquel B, Lemassu A, Daffe M, Jackson M (2008). Capsular glucan and intracellular glycogen of *Mycobacterium tuberculosis*: biosynthesis and impact on the persistence in mice. *Mol Microbiol* **70**(3)**:** 762-774

Sani M, Houben EN, Geurtsen J, Pierson J, de Punder K, van Zon M, Wever B, Piersma SR, Jimenez CR, Daffe M, Appelmelk BJ, Bitter W, van der Wel N, Peters PJ (2010). Direct visualization by cryo-EM of the mycobacterial capsular layer: a labile structure containing ESX-1-secreted proteins. *PLoS Pathog* **6**(3)**:** e1000794

Sassetti CM, Boyd DH, Rubin EJ (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**(1)**:** 77-84

Sassetti CM, Rubin EJ (2003). Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A* **100**(22)**:** 12989-12994

Schwebach JR, Glatman-Freedman A, Gunther-Cummins L, Dai Z, Robbins JB, Schneerson R, Casadevall A (2002). Glucan is a component of the *Mycobacterium tuberculosis* surface that is expressed in vitro and in vivo. *Infect Immun* **70**(5)**:** 2566-2575

Shaka AJ, Barker PB, Freeman R (1985). Computer-Optimized Decoupling Scheme for Wideband Applications and Low-Level Operation. *Journal of Magnetic Resonance* **64**(3)**:** 547-552

Smith WL, Ballou CE (1973). The 6-O-methylglucose-containing lipopolysaccharides of *Mycobacterium phlei*. Locations of the neutral and acidic acyl groups. *J Biol Chem* **248**(20)**:** 7118-7125

Stadthagen G, Sambou T, Guerin M, Barilone N, Boudou F, Kordulakova J, Charles P, Alzari PM, Lemassu A, Daffe M, Puzo G, Gicquel B, Riviere M, Jackson M (2007). Genetic basis for the biosynthesis of methylglucose lipopolysaccharides in *Mycobacterium tuberculosis*. *J Biol Chem* **282**(37)**:** 27270-27276

Stanley SA, Raghavan S, Hwang WW, Cox JS (2003). Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. *Proc Natl Acad Sci U S A* **100**(22)**:** 13001-13006

Stokes RW, Norris-Jones R, Brooks DE, Beveridge TJ, Doxsee D, Thorson LM (2004). The glycan-rich outer layer of the cell wall of *Mycobacterium tuberculosis* acts as an antiphagocytic capsule limiting the association of the bacterium with macrophages. *Infect Immun* **72**(10)**:** 5676-5686

Takayama K, Wang C, Besra GS (2005). Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin Microbiol Rev* **18**(1)**:** 81-101

Tuffal G, Albigot R, Riviere M, Puzo G (1998). Newly found 2-N-acetyl-2,6-dideoxy-beta-glucopyranose containing methyl glucose polysaccharides in *M.bovis* BCG: revised structure of the mycobacterial methyl glucose lipopolysaccharides. *Glycobiology* **8**(7)**:** 675-684

Tuffal G, Ponthus C, Picard C, Riviere M, Puzo G (1995). Structural elucidation of novel methylglucose-containing polysaccharides from *Mycobacterium xenopi*. *Eur J Biochem* **233**(1)**:** 377-383

Tung KK, Ballou CE (1973). Biosynthesis of a mycobacterial lipopolysaccharide. Properties of the polysaccharide: acyl coenzyme A acyltransferase reaction. *J Biol Chem* **248**(20)**:** 7126-7133

Tzvetkov M, Klopprogge C, Zelder O, Liebl W (2003). Genetic dissection of trehalose biosynthesis in *Corynebacterium glutamicum*: inactivation of trehalose production leads to impaired growth and an altered cell wall lipid composition. *Microbiology* **149**(Pt 7)**:** 1659-1673

van der Wel N, Hava D, Houben D, Fluitsma D, van Zon M, Pierson J, Brenner M, Peters PJ (2007). *M. tuberculosis* and *M. leprae* translocate from the phagolysosome to the cytosol in myeloid cells. *Cell* **129**(7)**:** 1287-1298

Veech RL, Lawson JW, Cornell NW, Krebs HA (1979). Cytosolic phosphorylation potential. *J Biol Chem* **254**(14)**:** 6538-6547

Vollmer W, Holtje JV (2004). The architecture of the murein (peptidoglycan) in gram-negative bacteria: vertical scaffold or horizontal layer(s)? *J Bacteriol* **186**(18)**:** 5978-5987

Watkins HA, Baker EN (2006). Structural and functional analysis of Rv3214 from *Mycobacterium tuberculosis*, a protein with conflicting functional annotations, leads to its characterization as a phosphatase. *J Bacteriol* **188**(10)**:** 3589-3599

Watkins HA, Baker EN (2008). Cloning, expression, purification and preliminary crystallographic analysis of the RNase HI domain of the *Mycobacterium tuberculosis* protein Rv2228c as a maltose-binding protein fusion. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**(Pt 8)**:** 746-749

Weisman LS, Ballou CE (1984a). Biosynthesis of the mycobacterial methylmannose polysaccharide. Identification of a 3-O-methyltransferase. *J Biol Chem* **259**(6)**:** 3464-3469

Weisman LS, Ballou CE (1984b). Biosynthesis of the mycobacterial methylmannose polysaccharide. Identification of an alpha 1-->4-mannosyltransferase. *J Biol Chem* **259**(6)**:** 3457-3463

Wititsuwannakul D, Kim KH (1977). Mechanism of palmityl coenzyme A inhibition of liver glycogen synthase. *J Biol Chem* **252**(21)**:** 7812-7817

Wolf A, Kramer R, Morbach S (2003). Three pathways for trehalose metabolism in *Corynebacterium glutamicum* ATCC13032 and their significance in response to osmotic stress. *Mol Microbiol* **49**(4)**:** 1119-1134

Woodruff PJ, Carlson BL, Siridechadilok B, Pratt MR, Senaratne RH, Mougous JD, Riley LW, Williams SJ, Bertozzi CR (2004). Trehalose is required for growth of *Mycobacterium smegmatis*. *J Biol Chem* **279**(28)**:** 28835-28843

Yabusaki KK, Ballou CE (1978). Interaction of mycobacterial polymethylpolysaccharides with paranaric acid and palmitoyl-coenzyme A: structural specificity and monomeric dissociation constants. *Proc Natl Acad Sci U S A* **75**(2)**:** 691-695

Yabusaki KK, Ballou CE (1979). Effect of polymethylpolysaccharides on the hydrolysis of palmitoyl coenzyme A by a thioesterase from *Mycobacterium smegmatis*. *J Biol Chem* **254**(24)**:** 12314-12317

Yabusaki KK, Cohen RE, Ballou CE (1979). Conformational changes associated with complex formation between a mycobacterial polymethylpolysaccharide and palmitic acid. *J Biol Chem* **254**(15)**:** 7282-7286

Yamada H, Cohen RE, Ballou CE (1979). Characterization of 3-O-methyl-D-mannose polysaccharide precursors in *Mycobacterium smegmatis*. *J Biol Chem* **254**(6)**:** 1972-1979

Zuber B, Chami M, Houssin C, Dubochet J, Griffiths G, Daffe M (2008). Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *J Bacteriol* **190**(16)**:** 5672-5680