1. General introduction
**TUBERCULOSIS**

The infectious pulmonary disease tuberculosis (TB) is caused by a group of closely related mycobacterial species known as the *Mycobacterium tuberculosis complex* (MTBC). These species have been coexisting for millennia in the human and animal population. The oldest known case of human TB has been detected in mummies of a woman and child from the Pre-Pottery Neolithic time period (≈9,000 kiloannum (ka)) in Atlit-Yam, Israel (1). In animals, the earliest proven TB case even dates back as early as the Upper Paleolithic period, where an extinct ancient Bison (≈17,000 ka) from Natural Trap Cave in Wyoming, USA was positively tested on bovine TB (2). Further dating the origin of human TB is a difficult process and under a lot of debate (3). However, new approaches rely on DNA sequencing of the global diversity of human-adapted MTBC strains, and coupling this MTBC phylogeny with the human phylogeny of mitochondrial genomes, representing the main human mitochondrial haplogroups. Interestingly, geographically located haplogroups showed evidence for co-divergence with the most frequent MTBC lineages in these areas. Important branching points that coincided within both phylogenies (MTBC and human) suggested that MTBC must have an African origin. Furthermore, it is highly likely that human TB dates back as early as the early out-of-Africa migration events in human evolution approximately 70,000 years ago (4). Throughout human evolution, TB has been an important human pathogen with the potential to have devastating effects on the human population. These effects become clearly noticeable, when during the Early modern Period, one-fourth of the European population was killed by TB. This high mortality rate disease is commonly known as ‘the white plague’ (5). It was not until late in the 19th century that Robert Koch identified *Mycobacterium tuberculosis* as the causative agent of TB. This important milestone in TB research, which was awarded with the Nobel Prize in 1905, paved the road for mycobacterial research that allowed development of anti-TB drugs and detection of tubercle bacilli. From the 1940s to 1960s, the discoveries of several anti-TB drugs rapidly followed one another. These drugs, known as the first-line TB-drugs, consisted of: streptomycin (1943), isoniazid (1952), pyrazinamide (1952), rifampicin (1957) and ethambutol (1961). These drugs showed a remarkable potency as anti-TB treatment. The identification of these compounds, and also the combination of these compounds, initiated the modern era of effective anti-TB chemotherapy (6). The first-line drugs, combined with the availability of the Bacillus Calmette-Guérin (BCG) vaccine (1921) formed the first prevention and treatment option that was available to the large public (7). Both strategies dramatically decreased the incidence and death rates of the TB epidemic in the Western/developed countries. However, due to the absence of treatment efficacy data and clinical outcome data and due to inconsistent use of anti-TB chemotherapies, the number of antibiotic resistant strains steadily increased over time. Added to this, the emergence of the HIV-1 epidemic contributed significantly to the resurgence of TB in the ‘90s. Over the recent decades, the multi-drug resistance coupled with high TB susceptible human sub-population in synergy with HIV showed that the TB epidemic is challenging to control (8). Today, despite tremendous efforts to combat TB, *M. tuberculosis* is still the leading bacterial
cause of death known to mankind killing 1.5 million people annually (9). Estimates suggest that one-third of the human population is either infected or has previously been exposed to the bacillus. These disturbing numbers illustrate the persistent, widespread nature of TB and the devastating effects on mankind. Clearly, there is an urgent need to develop novel intervention strategies, chemotherapies and immunotherapies to combat TB.

**MYCOBACTERIAL SPECIES**

The genus *Mycobacterium* belongs to the Actinobacteria and has a cell envelope containing unique long C\textsubscript{60-90}-chain mycolic acids. Mycobacteria are G+C rich (62-70%), irregular rod shaped (diameter X length; 0.2-0.5 μM X 1.5-4.0μM) (10). Mycobacteria are also called acid-fast organisms, which can be visualized with staining methods like Ziehl-Neelsen. Because mycobacteria belong to the Actinobacteria, they have been mistakenly labelled as Gram- positives. However, we now know that the structure of the mycobacterial cell envelope resembles that of Gram- negatives, i.e., bacteria with an inner and outer membrane layer. Therefore, mycobacteria can not be really classified according to the Gram scheme.

Mycobacteria are commonly found in nature and its ecology is very diverse, occupying a variety of habitats including water, soil and food-sources and also including different host organisms. Over the years, more than 150 different mycobacterial species have been classified using phylogenomic studies such as 16S rRNA gene typing (figure 1) and whole genome sequencing (11). Based on their genotype-phenotype properties, different groups of mycobacteria have been classified. For instance, differences in growth-rates led to a division into fast and slow-growing mycobacteria. This division runs parallel with the phylogenetic tree based on genomic sequences, indicating that the observed differences in growth speed are probably based on a fundamental genetic differences between these species. In addition, a differentiation between free living (saprophytic) and host-adapted species has been made and disease causing properties led to a division of tuberculous and non-tuberculous mycobacteria (NTM)(11) (figure 1). The nontuberculous group consists of pathogenic species not belonging to the MTBC and not *M. leprae*, the causative agent leprosy (figure 1). Although the closely related pathogens *M. marinum* and *M. avium* cause tuberculosis in fish and birds, respectively, they are not classified as members of the tuberculous group. However, many mycobacterial species in the NTM group are pathogens or opportunistic pathogens in immuno-compromised patients, for instance in HIV patients, and can cause diseases or tuberculosis-like diseases. Known members are *M. ulcerans*, the causative agents of Buruli ulcers and members of the *Mycobacterium avium* complex (MAC), *M. abscessus*. and *M. kansasii* (11) (figure 1).

**MYCOBACTERIAL CELL ENVELOPE**

The architecture of the mycobacterial cell envelope resembles that of Gram- negative bacteria (14, 15). Mycobacteria possess an extraordinary hydrophobic cell envelope constructed of three main constituents (figure 2). The plasma membrane, the innermost layer, resembles the classical bacterial membrane and consist of normal phosphatidyl-inositol–based (PI) lipids including several mycobacterial-specific (glyco)lipids and lipo-proteins (16). The phospho- and mycobacterial-specific lipid composition can vary between mycobacterial species. On top of the inner-membrane is a complex and unique, covalently linked, mycolyl-arabinogalactan-peptidogalactan cell-wall core (mAGP-complex), which is essential for bacterial survival (figure 2). The mAGP comprises a peptidoglycan (PG) layer that is covalently linked via a phosphoryl-N-acetylglucosaminyl-rhamnose (L-Rhap-D-GlcNAc) linker unit to the galactan domain (figure 2, in green) of arabinogalactan (AG) polysaccharide core. In turn, the terminal arabinosyl residues (figure 2, in blue), at the non-reducing ends, are esterified to the α-alkyl-β-hydroxy long chain fatty acid mycolic acids, the hallmark of the mycobacterial cell envelope.
These long chain mycolic acids (MAs) have been proposed to form the inner leaflet of the outer membrane (figure 2). The MA structure is based on a C3 carboxylic acid, 3-hydroxypropionic acid, where the fatty acid chain and meroc-mycolic (meromycolic) chain is positioned on CO₂ and CO₃, respectively. Between mycobacterial species, the meroc-chain and fatty acid chain can vary in length from C₄₂-C₆₂ and C₂₂-C₂₆, (20). Furthermore, the meromycolic chains are defined by chemical groups (keto, methoxy, epoxy, wax ester and hydroxy) and the chiral stereo-center cis/trans configuration on the C double bond, which diversify the mycolic acids further. Again, variations in subtypes can differ between species (21, 22). The outer-leaflet of the outer-membrane is formed by a variety of non-covalently attached glycolipids (figure 2), including acyl-glycerols, free fatty acids, (poly)ketides, trehalose esters (PDIM, cord-factor TDM), PI-based lipoglycans (lipomannan (LM) and lipoarabinomannan (LAM)) and (lipo)proteins. Many of these lipids are again unique for mycobacteria and some closely related genera. Together, the inner and outer leaflet of the outer membrane form an asymmetrical bilayer also referred to as the ‘mycomembrane’ (figure 2). Interestingly, the mycomembrane is only slightly thicker than the inner-membrane, which suggests that MAs are in a folded configuration in the inner leaflet. The outermost layer of the mycobacterial cell envelope is the capsule, a loosely attached polysaccharide-rich layer that also contains a minority of proteins and lipids (15, 23-25) (figure 2). In M. tuberculosis, the capsule mainly consists of three polysaccharides: capsular α-glucan, a high molecular weight glyco-polymer that structurally resembles intracellular glycogen and represents up to 80% of its capsular content, D-arabino-D-mannan (AM) and D-mannan (23, 26, 27) (discussed below). The amount and composition of the capsular constituents and entities of the mycomembrane and or mAGP complex vary between or even within mycobacterial species. Interestingly, the capsule is the outermost layer and its constituents are the first to interact with the host. Therefore, these layers are expected to be important in host-pathogen interactions in pathogenic mycobacteria. In this regard, capsular polysaccharides α-glucan and AM of M. tuberculosis are associated with immunomodulation and virulence (23, 24, 27, 28). Furthermore, several fragments from both the mycobacterial mAGP complex and the outer leaflet of the mycomembrane are well known immunomodulatory factors. Therefore, it is highly likely that the dynamic interplay between all these compounds and host immunity is crucial for adaptation and long-term survival of the tubercle bacilli in its host.

**BACTERIAL CAPSULES**

Many bacterial species from diverse genera produce an extracellular layer, rich in polysaccharides that is known as the capsule- or glycocalyx layer. The capsule covers bacteria and serves as a protective barrier to shield off bacterial membrane surfaces, which makes them more resistant to harsh environmental conditions and membrane attack complexes. The capsule consists mainly of high molecular weight exopolysaccharides, which can be covalently attached to the wall in Gram- positives bacteria or modified with a lipid moiety in Gram- negatives. Most capsules are negatively charged and highly hydrated with water (up to 95%) with a capsular layer thickness between 100-400nm in *Escherichia coli* (29). The capsular polysaccharides have remarkable structural and compositional diversity as illustrated by *E. coli* that can produce a variety of glycopolymers (250 +) capsule polysaccharides (30, 31). Several well-known human pathogens rely on (variable) capsules, including *Streptococcus pneumoniae*, *Neisseria meningitidis*, *Bacillus anthracis*, *Haemophilus influenzae*, *Salmonella enterica*, *E.coli* and *Klebsiella pneumoniae* (32). Intensive research over the years indicated that the capsule of pathogenic bacteria is important or even crucial for bacterial persistence in its host. These sugar structures are implicated in blocking damaging of the cell membrane by the complement-mediated membrane attack complex or antimicrobial peptides, or shielding bacterial cell structures, thereby blocking recognition by antibodies and scavenging.
receptors of phagocytes (33-36). Finally, capsules can modulate the adaptive immune system. In *Klebsiella*, the capsule modulates the ability of B-cells, to the production of antibodies to homogenous and heterologous antigens (37). Additionally, a special class of bacterial capsular polysaccharides is the so-called zwiterionic polysaccharides (ZPSs), which interact at several stages with the adaptive immunity. These ZPSs activate CD4+ T-cells in a MHC-II dependent manner, induce CD8+ T-cells by CD4+ T-cell TCR crosslinking and modulate diverse pro and inflammatory cytokines, which indicates that ZPSs are potent adaptive immune modulators (reviewed in (38)). All together, capsule polysaccharides have important biological functions for encapsulated bacterial organisms and are usually essential for bacterial pathogenesis. For this reason, capsule polysaccharides are widely used as vaccines. Over the years, a variety of mono- or poly-valent capsule-based vaccines were developed against *Streptococcus pneumoniae*, *Neisseria meningitidis*, *Haemophilus influenzae* (39, 40) and *Campylobacter jejuni* (41). These capsule polysaccharide based vaccines have been proven to be a very successful strategy in controlling devastating diseases associated with encapsulated bacteria. Recently, for *M. tuberculosis*, one of the most challenging human bacterial pathogens, a polysaccharide-based capsule layer is identified (15, 26). Therefore, when the molecular and chemical insights of these capsular polysaccharides have been revealed, perhaps this knowledge could be used in controlling TB.

THE MYCOBACTERIAL CAPSULE

The presence of a capsule layer in Mycobacteria was suggested many decades ago. The first signs of the mycobacterial capsular layer came from the 50s, when Chapman and colleagues infected cell lines with *Mycobacterium leprae* and detected a space between the phagosomal membrane and the bacterial cell wall, referred to as capsular space (42). Not much later, Hanks reported an unstainable halo surrounding pathogenic mycobacteria (42). These capsule polysaccharide based vaccines have proven to be a very successful strategy in controlling devastating diseases associated with encapsulated bacteria. Recently, for *M. tuberculosis*, one of the most challenging human bacterial pathogens, a polysaccharide-based capsule layer is identified (15, 26). Over the next decades, advances in new electron microscopy techniques as well as sample preparation methods allowed better visualization of the ETZ in *Mycobacterium leprae in vivo* using electron microscopy (43). Over the next decades, advances in new electron microscopy techniques as well as sample preparation methods allowed better visualization of the ETZ (44-47). Further pioneering work by the group of Daffé on the ETZ revealed that immuno-gold labeling with anti-arabinan stained the outermost ETZ layer of *M. tuberculosis* (26). Both this immuno-gold labeling and the presence of the ETZ was massively reduced when bacterial cells were mechanically pre-treated with glass beads, which suggested that this zone seems to be a loosely associated capsule layer. However, only recently, the definite proof of the existence of a mycobacterial capsule was established by our group and showed direct visualization of the mycobacterial capsular layer in its (near) native state by Cryo-EM; the layer proved to be labile to detergent action and contained α-glucan, phosphatidyl-inositol mannosides (PIMs), lipoarabinomannan (LAM), arabinomannan (AM) and secreted proteins (15). Over the years, several groups investigated the capsule material with different methods, such as cytometry and immuno-cytochemistry, and gained various insights into the molecular composition of this layer. Different reports have revealed that both mycobacteria groups NTM as well as tuberculosis group, are composed of mainly polysaccharides and protein with a minority of lipids (26, 48-51). However, the ratio of the three compound classes varies according to the mycobacterial species, for instance, in *M. smegmatis* the capsule is rich in proteins while the major component in *M. tuberculosis* are polysaccharides (26, 48). The capsular polysaccharide of *M. tuberculosis* consists primarily of three oligosaccharides, namely α-glucan, D-arabinomannan and D-mannan (figure 3). In addition small traces of xylose are present, possibly in other, yet, unidentified polysaccharides (23). (1) The homopolysaccharide α-glucan is the main capsular sugar representing up to 80% of its content and is a high molecular weight (>100 kDa) polymer of linear α-D-1-4 linked glucose units with core substitutions at position 6 (α-1-6 linked), every 4 to 6 residues by an α-D-1-4 linked oligoglucoside, which vary in length (one to nine glucose units; figure 3B). This α-glucan polymer resembles cellular glycoprotein in structure and chemical composition but is smaller in size (±100.000 Da). However, the optical rotatory scattering values of capsular glucan ([α]_238 ±211) versus glycogen ([α]_238 ±167) is lower, which suggests minor differences in the absolute configuration of the two polymers (28). Furthermore, the molecular weight, as determined by analytical ultracentrifugation, is slightly higher for capsular glucan (13.0 x 10^6 Da versus 7.5 x 10^4 Da for intracellular mycobacterial glycoprotein) suggesting a more compact 3D-structure (52). (2) Hetero-polysaccharide AM (A) has a molecular weight of ±13kDa and its structure resembles the glycoprotein part of LAM in *M. tuberculosis*. This saccharide consists of arabinose and mannoside residues whereby an α-1-6 linked mannosyl-chain (lower part AM) of usually 20-30 mannoside residues long, can be core substituted with α-1-2 linked mannosyl units at some positions (figure 3A). Furthermore, this mannoside chain is substituted with one α-1-2 linked arabinosyl-mannosyl tetra antenna (figure 3A) that at the non-reducing end is substituted with mannoside residues (“caps”) in *M. tuberculosis* as well as other slow growing mycobacteria (23). (3) Homopolysaccharide mannan, is identical to LM without the lipid moiety and is composed of the α-1-6 mannosyl-chain of AM part (figure 3A). The phosphatidyl-inositol lipid anchors of LM and LAM are similar and bear at position R1 and R2 respectively, tuberculostearic acid and palmitic acid (figure 3A).

CAPSULAR α-GLUCAN; BIOSYNTHESIS ROUTES AND BEYOND

In *M. tuberculosis* the major capsular polysaccharide is α-glucan representing up to 80% of its content and is therefore, an important constituent of the capsules’ barrier in regard to physical as well as chemical properties and suspected to be an important virulence factor. Capsular α-glucan is likely to be the first component to interact with the host and associated with host immune modulation, including enhanced phagocytosis. Capsular α-glucan, is a linear polysaccharide that resembles cellular glycoprotein and therefore, probably, shares common members of the biosynthesis machinery of intracellular glycoprotein. Glycoprotein biosynthesis in *E. coli* and other bacterial species is well-researched (reviewed in (54)). Like *E. coli*, mycobacteria have several genes coding for common known glycoprotein biosynthesis enzymes. Indeed, in mycobacteria, the classical glycoprotein Gli enzymes prove to be involved in biosynthesis of both cellular glycoprotein and or capsular α-glucan polymers. Different reports on *M. tuberculosis* show...
Capsular α-glucan is a large glucopolymer of about $13.0 \times 10^6$ Da in size. Figure is adapted from (53).

The α-glucan biosynthesis enzymes, GlgB, GlgE and GlgA in combination with Rv3032, are essential to the synthesis of the α-glucans. Genetic deletion of $Rv3032$ in $M. tuberculosis$ dramatically decreases the glycogen levels (56). Finally, the biosynthesis of the α-glucan polymers occurs via the OtsA/B-TreS-GlgE pathway as all these members of the synthesis machinery are present in mycobacteria. The biosynthesis route occurs in four steps process; synthesis of trehalose (OtsA/B pathway), conversion of trehalose to maltose (TreS), generation of activated sugar maltose-P intermediate (Pep-2) and biosynthesis of the α-glucans by glucosyltransferase GlgE from the maltose-P intermediate (58-62). Based on the literature on glycogen/capsular α-glucan biosynthesis prior this study and on potential α-glucans biosynthesis enzymes encoded by the mycobacterial genome, three biosynthesis routes of intra- and capsular α-glucans are possible (depicted in figure 4). Until today, α-glucan biosynthesis, regulation, localization as well as to which degree each of these biosynthesis routes contributes to the α-glucan polymer production, is largely unknown. The α-glucans MGLPs and cellular glycogen are cytosolic glucopolymers and, very likely, produced intracellularly. In addition, most biosynthesis enzymes (Glg enzymes and TreS/Pep2) are predicted to be cytosolic proteins supporting the intracellular biosynthesis hypothesis. However, this does not exclude an extracellular biosynthesis route of capsular α-glucan. If the glucopolymers are synthesized in the cytosol, capsular α-glucan needs to be transported to the capsule by a specialized transport system. Translocation of such a large, high molecular weight and hydrophilic polymer over the hydrophobic cell envelope of mycobacteria to its final destination, the capsule, is extremely challenging. In Gram-negative bacteria, several capsular polysaccharide transport systems have been identified where capsular polysaccharide biosynthesis machinery and transport is (in)directly coupled (29). It is possible that a comparable capsular polysaccharide transport system exists in $M. tuberculosis$. However, the capsular α-glucan transport system in mycobacteria is still not identified and remains enigmatic. Although our basic understanding in α-glucans biosynthesis is emerging, research into the biological function of this polysaccharide is extremely complex. The α-glucan biosynthesis enzymes, GlgB, GlgE and GlgA in combination with $Rv3032$, are essential to $M. tuberculosis$ (28, 60, 63). Although construction of capsular α-glucan negative mutants is essential to study its biological function in vivo, these mutants are not available yet. With this awareness, only a limited number of studies have addressed the role of capsular α-glucan. The mycobacterial α-glucan capsule probably interacts at different levels with host immunity. Gagliardi and colleagues have shown that CD1 expression in dendritic cells (DCs) is blocked by capsular α-glucan and α-glucan modulates cytokine production by suppressing interleukin IL-12 and stimulation IL-10 production in a CD80 dependent way, suggesting involvement of capsular α-glucan in shifting toward Th2 anti-inflammatory immune response (64). In addition, capsular α-glucan hampers DC function dependent on its interaction to the C-type lectin DC-SIGN resulting in induction of IL-10 (65). Capsule components of $M. tuberculosis$ have an anti-phagocytic effect on macrophages (66). Along similar lines,
capsular α-glucan interacts with complement receptor CR3, which mediates non-opsonic binding to \textit{M. tuberculosis} (67) and suggests interaction with this receptor may promote intracellular survival by reducing respiratory burst and production of IL-12 (68). Furthermore, genetic deletion of glycosyltransferase \textit{GlgA} in \textit{M. tuberculosis} impairs virulence in mice, suggesting an important role of α-glucan \textit{in vivo} (28). Finally, and of notice, capsular α-glucan is host glycogen-like and therefore could be involved in evading parts of the immune system by molecular mimicry, as proposed by Daffé and colleagues (69). All together, these data suggest a significant role of capsular α-glucan in pathogenesis and virulence of \textit{M. tuberculosis} and warrants further research on this important capsule component.

SCOPE THESIS AND CHAPTER CONTENT

Mycobacterial capsular α-glucan is the major capsule polysaccharide in \textit{M. tuberculosis} and is expected to be important for pathogenesis and virulence. Although, over the years, major research efforts have focused on this capsular polysaccharide, some major questions remain unanswered. This is partially due to the currently available technical tools to quantify capsular glucan. Usually, capsular α-glucan detection was analyzed by conventional, time consuming, techniques like HPLC and gas chromatography, which does not allow an unbiased, genome wide screening to identify (un)known genes involved in capsular α-glucan production, regulation or transport. The aim of the research outlined in this thesis was to investigate the major capsular polysaccharide α-glucan and to elucidate the biosynthesis routes, transport system and the biological function of this important capsular component of \textit{M. tuberculosis}. In chapter 2, we describe newly developed, antibody-based techniques to overcome a major limitation in the mycobacterial capsular polysaccharide research field. We portray the epitope-fine characterization of an α-glucan monoclonal antibody and different, fast capsular α-glucan detection and quantitation assays. With this tool in hand, we set-up a genome-wide screening methodology by coupling the Mab-based detection assays with large scale transposon mutagenesis in different mycobacterial species. From this screen several interesting mutants were identified and investigated in detail, described in later chapters. This chapter forms the heart of this thesis. In chapter 3, we describe and fully elucidate the biosynthesis routes of intra and extracellular α-glucans in mycobacteria and reveal, for the first time, the importance of the capsular α-glucan in \textit{M. tuberculosis} virulence. This work shows how mycobacterial α-glucans are synthesized as well as a genetic engineered capsular α-glucan negative mutant that allows \textit{in vivo} study into the biological function of this capsule polysaccharide. In chapter 4, we reveal the in-depth analysis of a group of mutants in the ABC transport locus \textit{pst}, which were picked-up in the α-glucan antibody based transposon screen. Transporter \textit{Pst} is important in the acquisition of inorganic phosphate and these mutants massively overproduce capsular α-glucan. In this study, we reveal that inorganic phosphate stress is an important trigger to regulate the capsule α-glucan levels in mycobacteria and focus on the possible stress response pathways involved in inorganic phosphate mediated regulation of capsular α-glucan. By studying the stringent response initiator \textit{Relm} and transcription factor \textit{SigE}, a mediator of a complex regulatory network, under stress conditions, we were able to identify the pathway of capsular α-glucan regulation under \textit{Pi}- stress and present the \textit{in vivo} relevance of \textit{Pi}-stress in mycobacteria. This chapter gives a good insight into how mycobacterial capsular α-glucan is influenced under stress conditions. In chapter 5, a detailed study is performed on a fascinating transposon mutant identified in \textit{M. marinum}. This mutant had a transposon inserted in a gene orthologous to \textit{M. tuberculosis} \textit{H37Rv rv2179c}, of previously unknown function. A mutant in \textit{M. smegmatis} was affected in capsular α-glucan production and one in \textit{M. marinum} severely attenuated in causing a bacterial infection \textit{in vivo} in the zebrafish. We describe that the gene \textit{rv2179c} codes for a protein of the superfamily DEDDh exonucleases, which is capable of hydrolyzing specifically polyA RNA nucleotides, which suggests that this enzyme can hydrolyze PolyA tails of mRNA transcripts. Additionally, we give an in-depth structural analysis of protein function, including its binding cavity, enzyme mode of action and nucleotide specificity. Due to its polyA specific RNA editing activity, we annotated the enzyme as RNase AS (Adenosine Specific). This chapter shows how powerful our unbiasedly genome wide mutant screening of capsular α-glucan is in investigating “the unknown”. In chapter 6, the studies presented in this thesis are evaluated in a general discussion and in chapter 7, the overall summary of this work is presented including the addenda.