



UNIVERSITY OF  
BIRMINGHAM

# **Biophysical and biochemical characterization of proteins involved in $\alpha$ -glucan biosynthesis in mycobacteria**

Ali Asghar Kermani

Supervisors: Dr. Klaus Fütterer and Professor Gurdyal Besra

Doctoral Thesis

Submitted to the School of Biosciences, University of Birmingham

UNIVERSITY OF  
BIRMINGHAM

**University of Birmingham Research Archive**

**e-theses repository**

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

## Abbreviations

AG.....	arabinogalactan
AUC.....	analytical ultracentrifugation
BACTH.....	bacterial adenylate cyclase two-hybrid system
BSA.....	bovine serum albumin
CAZy.....	Carbohydrate-Active enZYmes database
CR3.....	receptor type 3
DC.....	dendritic cell
DC-SIGN.....	dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin
DGG.....	diglucosylglycerate
DOT.....	directed observed therapy
GlcNAc.....	<i>N</i> -acetyl- $\alpha$ -D -glucosamine
IL-12.....	interleukin-12
IPTG.....	isopropyl $\beta$ -D -1-thiogalactopyranoside
ITC.....	isothermal titration calorimetry
LAM.....	lipoarabinomannan
M1-P.....	maltose 1-phosphate
mAGP.....	mycolyl-arabinogalactan-peptidoglycan
MDR.....	multi-drug resistant
MGLP.....	methylglucosyl lipopolysaccharides
MHC.....	major histocompatibility complex
MR.....	molecular replacement
Msm.....	<i>Mycobacterium smegmatis</i>
Mtb.....	<i>Mycobacterium tuberculosis</i>
NCS.....	non-crystallographic symmetry
NF-kB.....	nuclear factor kappa-light-chain-enhancer of activated B cells
PBS.....	phosphate buffer saline
PEP.....	phosphoenolpyruvate
PI3P.....	phosphatidylinositol 3-phosphate

PMPS..... polymethylated polysaccharides  
 PMSF.....phenyl methyl sulfonyl fluoride  
 RMSD.....root mean square deviation  
 RNI.....reactive nitrogen intermediates  
 ROI.....reactive oxygen intermediates  
 SAXS.....small angle X-ray scattering  
 SEC.....size exclusion chromatography  
 STD NMR.....saturation transfer difference NMR  
 TACO..... tryptophan aspartate-containing coat protein  
 TB.....tuberculosis  
 TDR.....totally drug-resistant  
 TLC.....thin layer chromatography  
 TLR..... toll-like receptor  
 TNF.....tumor necrosis factor  
 TPP..... trehalose phosphate phosphatase  
 TPS..... trehalose phosphate synthase  
 XDR..... extensively drug resistant

## Abstract

A capsule composed of mainly  $\alpha$ -glucan forms the most outermost layer of cell envelope of *Mycobacterium tuberculosis*, the causing agent of tuberculosis. Capsule acts as an inert barrier to prevent the diffusion of macromolecules towards inner parts of the envelope. Also, there is growing evidence that capsular glucans mediate the adhesion and penetration of bacilli into the host cells. In GlgE pathway, one of the three pathways involved in  $\alpha$ -glucan synthesis, trehalose is converted to glucan through the activity of four enzymes, trehalose synthase, maltokinase Pep2, maltosyltransferase GlgE and branching enzyme GlgB. The first two enzymes catalyze the conversion of trehalose to maltose, and phosphorylating maltose, respectively. It has shown that *M. tuberculosis* TreS and Pep2 form an octameric complex together. During this study we solved the structure of the complex at 3.6 Å in *M. smegmatis*. The structure reveals two pairs of Pep2 monomers bind to the opposite sides of the diamond shaped TreS tetramer in a 4 + 4 complex. However, studying the stoichiometry of the complex using other techniques such as: isothermal titration calorimetry (ITC), analytical ultracentrifugation (AUC) and X-ray scattering indicates a 4 + 2 complex formation in the solution. Moreover, our data using size exclusion chromatography would suggest that this complex formation is pH dependent and favors complex formation at more acidic pHs. This study is a step forward towards better understanding of the capsule synthesis in *M. tuberculosis* and highlights the role of complex formation between enzymes as an effective strategy to control the metabolic pathways in mycobacteria.

## DECLARATION

The work presented in this thesis was carried out in the School of Biosciences at the University of Birmingham, U.K., B15 2TT during the period October 2012 to September 2015.

The work in this thesis is original except where acknowledged by references. No part of the work is being, or has been submitted for a degree, diploma or any other qualification at any other University.

## Acknowledgments

My especial thanks go to Darwin Trust of Edinburgh and Professor Kenneth Murray for funding my Doctoral studies, without which it was only a dream would not come true.

I would like to thank my supervisors Dr. Klaus Fütterer and Professor Del Besra for their support and encouragement throughout my PhD project. They taught me how to handle challenging scientific experiments and also, and more importantly, critical thinking and planning the experiments in such a fashion that they will provide me with the answer of scientific questions.

I am indebted to Rosemary Parslow and Nick Cotton for their support in using analytical ultracentrifugation and isothermal titration calorimetry, respectively. And to Rana Roy because of his advice, training and friendship during my PhD project.

I also would like to thank Del's lab people in training me with necessary techniques and 8<sup>th</sup> floor and 7<sup>th</sup> people for their friendships and supports, especially Dr. Mike Tomlinson.

I am also grateful to my parents who have always supported me with my dreams, emotionally and financially. To my brother and my sister who were always available to listen to my problems and challenges.

## Table of Contents

<b>1</b>	<b>Introduction .....</b>	<b>9</b>
1.1	History of tuberculosis .....	1
1.2	Tuberculosis in modern era .....	3
1.3	The genus of <i>Mycobacterium</i> .....	4
1.4	Pathology and immunology of <i>M. tuberculosis</i> infection.....	6
1.4.1	<i>Phagocytosis of M. tuberculosis by macrophages</i> .....	6
1.4.2	<i>Development of TB granuloma</i> .....	7
1.4.3	<i>Survival inside macrophages</i> .....	9
1.4.4	<i>Beyond phagocytes</i> .....	13
1.5	Current treatment of TB and rise of MDR and XDR strains .....	13
1.6	The mycobacterial cell envelope .....	18
1.6.1	<i>Peptidoglycan</i> .....	18
1.6.2	<i>Arabinogalactan</i> .....	21
1.6.3	<i>Mycolic acids</i> .....	23
1.6.4	<i>Capsular layer</i> .....	25
1.7	GlgE pathway enzymes.....	38
1.7.1	<i>Trehalose synthase</i> .....	38
1.7.2	<i>Maltokinase Pep2</i> .....	46
1.7.3	<i>Maltosyltransferase GlgE</i> .....	50
1.7.4	<i>Branching enzyme GlgB</i> .....	55
1.8	Rv3032 pathway.....	57
1.9	Aims and objectives .....	61
<b>2</b>	<b>Materials and methods .....</b>	<b>63</b>
2.1	Cloning of <i>Mtb</i> and <i>MsmglgE</i> and <i>glgB</i> .....	64
2.2	Testing the expression and solubility of recombinant proteins.....	68
2.3	Protein expression and purification .....	70
2.4	Size exclusion chromatography .....	72
2.5	Checking the nature of <i>MsmTreS</i> and <i>MsmPep2</i> degradation with western blot .....	73
2.6	M1P purification from <i>MsmΔglgE</i> .....	74
2.7	Activity assay .....	75



2.7.1	Enzymatic activity of Pep2 ( $\Delta 70$ ) .....	77
2.8	Crystallization and data collection .....	78
2.9	Checking the expression and solubility of C-terminal his-tagged <i>MsmPep2</i> .....	79
2.10	Overexpression of Pep2 in <i>M. smegmatis</i> .....	79
2.11	Limited proteolysis of <i>MsmPep2</i> .....	80
2.12	Thermal shift assay .....	81
2.13	Small Angle X-ray Scattering.....	82
2.14	Analytical ultracentrifugation .....	83
2.15	Bacterial two-hybrid screen .....	84
2.16	$\beta$ -galactosidase activity assay.....	86
2.17	Isothermal titration calorimetry assay.....	87
	<b>Results .....</b>	<b>88</b>
<b>3</b>	<b>Crystallization of Pep2 .....</b>	<b>89</b>
3.1	Crystallization of maltokinase .....	90
3.2	Expression and purification of <i>MtbPep2</i> and <i>MsmPep2</i> .....	90
3.3	Verifying activity of <i>MsmPep2</i> .....	92
3.4	X-ray diffraction of <i>MsmPep2</i> .....	95
3.5	Expression and purification of C-terminally tagged <i>MsmPep2</i> .....	95
3.6	Limited proteolysis of <i>MsmPep2</i> .....	98
3.7	Increasing pH promotes <i>MsmPep2</i> stabilization.....	100
3.8	Summary .....	103
<b>4</b>	<b>Characterization of the <i>Mycobacterium smegmatis</i> TreS:Pep2 complex .....</b>	<b>104</b>
4.1	Crystallization of <i>MsmTreS:Pep2</i> complex.....	105
4.2	Protein expression and purification of <i>MsmTreS</i> and <i>MsmPep2</i> .....	105
4.3	Verifying activity of <i>MsmTreS</i> .....	107
4.4	<i>MsmTreS</i> forms a complex with <i>MsmPep2</i> .....	107
4.5	<i>MsmTreS:Pep2</i> complex formation is pH dependent .....	108
4.6	Assessing binding affinity of the complex by isothermal titration calorimetry .....	112
4.7	Complex formation increases the activity of <i>MsmPep2</i> in acidic pH .....	117
4.8	<i>MsmTreS</i> forms a complex with inactive mutant <i>MsmPep2</i> -K145A.....	120
4.9	Crystallization of the <i>MsmTreS:Pep2</i> complex.....	123
4.10	X-ray crystal structure of the <i>MsmTreS:Pep2</i> complex .....	126

4.11	The structure of <i>MsmPep2</i> .....	130
4.12	Truncating <i>MsmPep2</i> .....	132
4.13	TreS and Pep2 active sites.....	136
4.14	The TreS-Pep2 contact interface.....	136
4.15	Small Angle X-ray Scattering.....	139
4.16	Stoichiometry of the <i>MsmTreS:Pep2</i> complex.....	141
4.17	Summary .....	143
<b>5</b>	<b>Crystallization and interaction studies of <i>Mycobacterium smegmatis</i> GlgE with Pep2 and GlgB .....</b>	<b>144</b>
5.1	Crystallization of <i>MtbGlgE</i> .....	145
5.2	Protein expression and purification of <i>MtbGlgE</i> and <i>MsmGlgB</i> .....	145
5.3	Expression and purification of <i>MsmGlgE</i> .....	148
5.4	Crystallization of <i>MtbGlgE</i> .....	150
5.5	Probing stability of GlgE using thermal shift assay .....	150
5.6	Testing the interaction of <i>MsmGlgE</i> with <i>MsmPep2</i> and GlgB using size exclusion chromatography.....	153
5.7	Bacterial Two-hybrid system.....	156
5.8	Summary .....	159
<b>6</b>	<b>Crystallization and interaction studies of Rv3031 and Rv3032.....</b>	<b>160</b>
6.1	Crystallization of Rv3031 and Rv3032 .....	161
6.2	Expression and purification of Rv3031 .....	161
6.3	Expression and purification of Rv3032.....	166
6.4	Domain structure of Rv3031 and Rv3032 .....	168
6.5	Summary .....	168
<b>7</b>	<b>Discussion .....</b>	<b>170</b>
<b>8</b>	<b>References .....</b>	<b>183</b>
<b>9</b>	<b>Supplements .....</b>	<b>193</b>

## LIST OF FIGURES

FIGURE 1-1. <i>M. TUBERCULOSIS</i> LIFE CYCLE. . . . .	4
FIGURE 1-2. THE LIFE CYCLE OF <i>M. TUBERCULOSIS</i> . . . . .	8
FIGURE 1-3. INTRACELLULAR LIFESTYLE OF <i>MYCOBACTERIUM TUBERCULOSIS</i> .. . . .	11
FIGURE 1-4. THE <i>MYCOBACTERIUM TUBERCULOSIS</i> CELL ENVELOPE. . . . .	19
FIGURE 1-5. STRUCTURE OF <i>M. TUBERCULOSIS</i> PEPTIDOGLYCAN BEFORE CROSS-LINKING. . . . .	20
FIGURE 1-6. CHEMICAL STRUCTURE OF ARABINOGALACTAN (AG) FROM <i>M. TUBERCULOSIS</i> AND ROLE OF AFTA, AFTB, AFTC, EMBA AND EMBB IN ITS SYNTHESIS. . . . .	22
FIGURE 1-7. CHEMICAL STRUCTURE OF MYCOLIC ACIDS FROM <i>M. TUBERCULOSIS</i> . . . . .	24
FIGURE 1-8. VISUALIZATION OF CAPSULE IN NATIVE-STATE USING ELECTRON MICROSCOPE. . . . .	26
FIGURE 1-9. SCHEMATIC OF THE SPATIAL ORGANIZATION OF MYCOBACTERIAL CELL ENVELOPE LAYERS AND LOCALIZATION OF THE ESX-1 SECRETION SYSTEM. . . . .	31
FIGURE 1-11. THE MOST WIDELY DEMONSTRATED PATHWAY RESPONSIBLE FOR TREHALOSE BIOSYNTHESIS IN MOST ORGANISMS. . . . .	36
FIGURE 1-12. SCHEMATIC MECHANISMS OF GLYCOSIDE HYDROLASES. . . . .	40
FIGURE 1-13. THE CRYSTAL STRUCTURE OF TREHALOSE SYNTHASE (TRES) FROM <i>M. TUBERCULOSIS</i> . . . . .	42
FIGURE 1-14. POSITION OF ACTIVE SITE AND TOPOLOGY OF SOLVENT CHANNEL FORMED BETWEEN CHAINS A AND B OF <i>MTBTRES</i> . . . . .	43
FIGURE 1-15. THE CATALYTIC MECHANISM OF TRES. . . . .	44
FIGURE 1-16. THE CRYSTAL STRUCTURE OF <i>MTBPEP2</i> . . . . .	47
FIGURE 1-17. MALTOSE AND NUCLEOTIDE BINDING SITES OF <i>MTBPEP2</i> . . . . .	49
FIGURE 1-18. THE CRYSTAL STRUCTURE OF <i>STREPTOMYCES COELICOLOR</i> GLGE. . . . .	53

FIGURE 1-19. CO-CRYSTALLIZATION OF MALTOSE AND GLGE .....	54
FIGURE 1-20. DOUBLE DISPLACEMENT MECHANISM OF GLGE.....	54
FIGURE 1-21. THE CRYSTAL STRUCTURE OF <i>M. TUBERCULOSIS</i> GLGB.. .....	56
FIGURE 1-22. THE SUBSTRATE-BINDING MODEL FOR <i>MYCOBACTERIUM TUBERCULOSIS</i> GLGB .....	56
FIGURE 1-23. STRUCTURES OF MYCOBACTERIAL POLYMETHYLATED POLYSACCHARIDES (PMPS).....	58
FIGURE 1-24. PROPOSED PATHWAY FOR THE BIOSYNTHESIS OF MGLP IN <i>M. TUBERCULOSIS</i> . .....	60
FIGURE 2-1. MEASURING PHOSPHORYLATION OF MALTOSE TO MALTOSE 1-PHOSPHATE USING AN ENZYMATICALLY COUPLED REACTION. ....	76
FIGURE 3-2. THE SIZE EXCLUSION CHROMATOGRAM OF <i>MtB</i> PEP2 AND <i>Msm</i> PEP2 FROM A SEPHACRYL S200 COLUMN.. .....	93
FIGURE 3-3. THIN LAYER CHROMATOGRAPHY OF REACTION MIXTURES PROBING THE ACTIVITY OF <i>Msm</i> PEP2. ....	94
FIGURE 3-4. THE DIFFRACTION PATTERN OF <i>Msm</i> PEP2 CRYSTALS GROWN IN 2.1 M MALIC ACID pH 7.0 AND FROZEN IN LIQUID NITROGEN.. .....	96
FIGURE 3-5. PURIFICATION OF RECOMBINANT C-TERMINALLY TAGGED <i>Msm</i> PEP2 (A) AND PEP2 OVEREXPRESSED IN <i>M. SMEGMATIS</i> (B).....	97
FIGURE 3-6. SDS-PAGE ANALYSIS OF <i>Msm</i> PEP2.....	99
FIGURE 4-2. THIN LAYER CHROMATOGRAPHY ANALYSIS OF REACTION PRODUCTS PROBING THE ACTIVITY OF <i>Msm</i> TRES.....	109
FIGURE 4-3. EXPLORING COMPLEX FORMATION BETWEEN <i>Msm</i> TRES AND <i>Msm</i> PEP2 IN A RANGE OF MOLAR RATIOS USING SIZE-EXCLUSION CHROMATOGRAPHY. ....	110
FIGURE 4-4. EXPLORING COMPLEX FORMATION BETWEEN <i>Msm</i> TRES AND <i>Msm</i> PEP2 UNDER DIFFERENT pH VALUES USING SEPHACRYL S300-HR SIZE EXCLUSION RESIN.....	111
FIGURE 4-5. SCHEMATIC REPRESENTATION OF THE ESSENTIAL COMPONENTS OF AN ITC SYSTEM.. .....	114

FIGURE 4-6. ISOTHERMAL TITRATION CALORIMETRY (ITC) GRAPHS OF TITRATION OF <i>MsmTRES</i> WITH <i>MsmPEP2</i> ..	115
FIGURE 4-7. MICHAELIS-MENTON KINETICS OF <i>MsmPEP2</i> ACTIVITY UNDER DIFFERENT PH CONDITIONS, IN ABSENCE AND PRESENCE OF <i>MsmTRES</i> , AND MEASURED BY AN ENZYME-COUPLED ASSAY .....	118
FIGURE 4-8. MEASURING THE ENZYMATIC ACTIVITY OF TRES USING AN ENZYMATICALLY COUPLED REACTION BETWEEN TREHALOSE TO GLUCOSE CONVERSION.....	121
FIGURE 4-9. EXPLORING COMPLEX FORMATION BETWEEN <i>MsmTRES</i> AND MUTANT PEP2-K145A.....	122
FIGURE 4-10. SDS-PAGE ANALYSIS OF CRYSTALS HARVESTED FROM <i>MsmTRES</i> :PEP2 CRYSTAL TRAYS.....	125
FIGURE 4-11. THE SUPERIMPOSITION OF THE TRES TETRAMER OF COMPLEX 1 (CHAINS A – D, IN HUES OF BLUE) TO THE TRES TETRAMER OF THE COMPLEX 2 (CHAINS E- H, IN GRAY).....	128
FIGURE 4-12. CRYSTAL STRUCTURE OF THE <i>MsmTRES</i> :PEP2 COMPLEX.....	129
FIGURE 4-13. STRUCTURE OF <i>MsmPEP2</i> .....	131
FIGURE 4-14. THE SIZE EXCLUSION PROFILE OF <i>MsmPEP2</i> ( $\Delta 10$ - $\Delta 70$ ) AND EVIDENCE OF COMPLEX FORMATION BETWEEN MUTANT PEP2 ( $\Delta 70$ ) AND <i>MsmTRES</i> .....	134
FIGURE 4-15. MICHAELIS-MENTON KINETICS OF <i>MsmPEP2</i> WILD TYPE AND <i>MsmPEP2</i> ( $\Delta 70$ ) ACTIVITY, MEASURED BY AN ENZYME-COUPLED ASSAY. ....	135
FIGURE 4-16. CONTACT SURFACE BETWEEN TRES AND PEP2.....	137
FIGURE 4-17. SMALL-ANGLE X-RAY SCATTERING (SAXS) ANALYSIS OF THE <i>MsmTRES</i> :PEP2 COMPLEX. ....	140
FIGURE 4-18. SEDIMENTATION EQUILIBRIUM ANALYSIS OF THE <i>M. SMEGMATIS</i> TRES:PEP2 COMPLEX.....	142
FIGURE 5-7. B-GALACTOSIDASE ACTIVITY OF OVERNIGHT CULTURES PREPARED FROM BTH101 COMPETENT CELLS .....	158
FIGURE 6-1. SDS-PAGE ANALYSIS OF RECOMBINANT RV3031 OVEREXPRESSION AND SOLUBILITY.....	162
FIGURE 6-2. SDS-PAGE ANALYSIS OF RECOMBINANT RV3031 OVEREXPRESSION AND SOLUBILITY IN THE PRESENCE OF CHAPERONES. ....	164

FIGURE 6-3. PURIFICATION OF RECOMBINANT Rv3031 FROM THREE LITERS OF BACTERIAL CULTURES. ....	165
FIGURE 6-4. PURIFICATION OF RECOMBINANT Rv3032. ....	167

## List of Tables

TABLE 1-1. CLASSIFICATION OF ANTI-TB DRUGS.....	15
TABLE 1-2. APPROVED DRUGS TO TREAT TUBERCULOSIS. ....	16
TABLE 1-3. SOME OF THE MAIN CAPSULAR PROTEINS OF <i>M. TUBERCULOSIS</i> .....	32
TABLE 2-1. SEQUENCE OF PRIMERS FOR CLONING.....	66
TABLE 2-2. THERMOCYCLER SETTINGS FOR PCR AMPLIFICATIONS. ....	68
TABLE 2-3. LIST OF COMPETENT CELLS/PLASMIDS USED DURING TRANSFORMATION WITH APPROPRIATE ANTIBIOTICS. .....	69
TABLE 2-4. LIST OF THE BUFFERS AND THEIR pH VALUES USED IN THE SCREEN .....	82
TABLE 2-5. THE CONCENTRATIONS OF <i>MsMTreS</i> , <i>TreS</i> - <i>Pep2</i> COMPLEX IN 1:1 AND 2:1 MOLAR RATIO BEFORE LOADING ON ANALYTICAL ULTRACENTRIFUGE. ....	84
TABLE 4-2. THERMODYNAMIC PARAMETERS OF <i>MsMTreS</i> : <i>Pep2</i> INTERACTION DETERMINED USING ISOTHERMAL TITRATION CALORIMETRY.....	116
TABLE 4-3. MICHAELIS-MENTON PARAMETERS OF <i>MsMPep2</i> -CATALYZED PHOSPHORYLATION OF MALTOSE IN ABSENCE AND PRESENCE OF <i>MsMTreS</i> .....	119
TABLE 4-4. LIST OF CONDITIONS GENERATING CRYSTALS OF <i>MsMTreS</i> : <i>Pep2</i> COMPLEX.....	124
TABLE 4-6. LIST OF <i>TreS</i> AND <i>Pep2</i> RESIDUES IN BINDING INTERFACE. ....	138
TABLE S9-1. ALL APPROVED DRUGS TO TREAT TUBERCULOSIS.....	195
TABLE S9-2. LIST OF NEW ANTI-TB DRUGS. ....	197
TABLE S9-3. LIST OF RECENTLY IDENTIFIED TARGETS IN <i>M. TUBERCULOSIS</i> .....	198
TABLE S9-4. X-RAY DIFFRACTION DATA AND REFINEMENT STATISTICS.....	199

# **1 Introduction**



## 1.1 History of tuberculosis

Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, has been the main cause of death from microbial infections in the history of human life (Daniel , 2006). Modern techniques in molecular biology and genome sequencing indicate that current members of mycobacterial family are originated from a common ancestor in East Africa around 3 million years ago (Daniel , 2006). However, the first evidences of tuberculosis go back to the discovery of a bone from Neolithic period (~ 5000 BC) in Heidelberg and spinal caries on the Egyptian mummies belong to 2400 BC (Herzog, 1998).

The oldest available text mentioning a chronic lung disease, most likely TB, is the Code of Hammurabi, a cuneiform script engraved on a stone pillar at 1948-1905 BC. Then, it appeared in Greek literature by Hippocrates (460-370 BC) as phthisis or consumption. He described the disease as a condition that usually leads to death and discouraged his students of visiting the patients in advanced stages of the disease, since the patient is doomed to death it could damage the credibility of the physician. Although, he was able to observe TB-induced nodular lesions for the first time in the tissues of pigs, cattle and sheep but he didn't find any analogous lesions in humans because human autopsy was not performed in that time (Basel, 1998). Aristotle (384-322 BC) described "scrofula" on the skin of phthisic pigs and believed that the disease is contagious not hereditary, which was the general opinion in that time (Basel, 1998). However, it was Roman physician Caelius Aurelianus in the 5<sup>th</sup> century AD, who described the symptoms of the disease for the first time as latent fever, strong cough with thin purulent sputum, breathing difficulties, flushed cheeks and ashen skin on the rest of the body. In spite of repellent appearance,

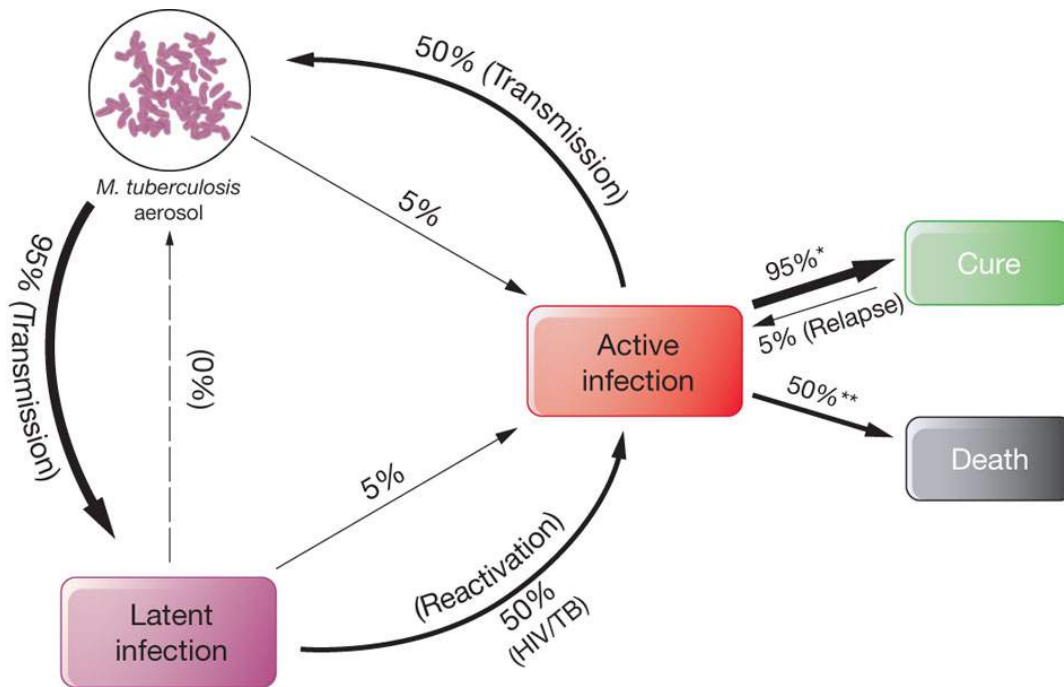
the patient usually shows stunning physical and mental activities. Clarissimus Galen, the physician of Roman Emperor Marcus Aurelius in 174, also re-emphasized on the contagious nature of consumption and therefore banned personal contact with phthisis patients (Basel, 1998).

TB became prevalent throughout Europe during 18<sup>th</sup> and 19<sup>th</sup> centuries displacing leprosy as a dominant infectious disease. The rate of death in metropolitan cities in Europe was 1,000 to 100,000 death per year (Daniel , 2006). In spite of high number of patients suffering from tuberculosis, there wasn't an agreement on the nature of tuberculosis between practitioners and scientists. While in Northern Europe it was considered as hereditary, in Southern Europe they would treat the disease as an infections condition. In 1790, Benjamin Marten considered the disease as infectious and assigned it to "some certain species of animalcula", the term which used to be used for amazing very small living creatures. This term was used for the first time by Anton van Leeuwenhoek in 1676 when he observed bacteria and spermatozoa under his microscope. However, Jean-Antoine Villemin a century later was able to demonstrate the contagious nature of tuberculosis for the first time by studying the rabbits inoculated with a "small amount of purulent liquid from an tuberculosis cavity" isolated from an active TB patient. The major progress towards the bacteriology of tuberculosis was achieved by Robert Koch (1834-1910) when he demonstrated a specific microbe is the main cause of tuberculosis (Basel, 1998). He visualized the bacilli for the first time using special staining techniques within the infected tissue and eventually called them tubercle bacilli, or *Mycobacterium tuberculosis*. Using Tyndall's method of sterilizing, he was, also, able to culture the

bacteria as a thin and delicate layer on the media and by inoculating the cultured cells into laboratory animals, demonstrated the transmission of tuberculosis. Finally he was awarded Nobel Prize due to his efforts on the bacteriology of *Mycobacterium tuberculosis* in 1905.

## **1.2 Tuberculosis in modern era**

It is estimated that one third of world population (about 2 billion people) are infected with tuberculosis (Koul et al., 2011). However, in most cases the infection remains dormant and only in 5% of patients will develop clinical disease (Figure 1-1). In case of normal TB, 95% of patients recover upon treatment with a cocktail of four different antibiotics, isoniazid, rifampicin, ethambutol and pyrazinamide for a duration of 6 month. But if it remains untreated it can be fatal in 50% of patients, currently leading to the death of around 1.5 million people per year (Koul et al., 2011). In 2012, 8630 cases of TB were reported in United Kingdom, out of those 432 patients died (Public Health England, 2014).



**Figure 1-1.** *M. tuberculosis* life cycle. For more details see section 1.2. Adopted from (Koul et al., 2011), with permission.

### 1.3 The genus of *Mycobacterium*

The genus of mycobacterium is composed of more than 50 species, which most of them are non-pathogenic bacteria related closely to soil bacteria *Streptomyces* and *Actinomyces* (Cosma, Sherman and Ramakrishnan, 2003). The main pathogenic species of this genus are *M. tuberculosis*, *M. leprae* and *M. ulcerans* causing tuberculosis, leprosy and Buruli ulcers, respectively. The key factor in the success of these pathogens is their ability to survive inside the host macrophages, despite the antimicrobial nature of these cells (Cosma, Sherman and Ramakrishnan, 2003). The host arrests the growth of mycobacterium inside complex structures called granuloma but fails to eradicate the bacterial cells. Mycobacterial cells can survive inside the granuloma for several years in

latent or dormant form and activate as soon as the immune system of the patient weakens (Cosma, Sherman and Ramakrishnan, 2003 and Russell, 2007).

### ***M. tuberculosis* complex**

*M. tuberculosis* complex is composed of *M. tuberculosis*, *M. africanum*, *M. bovis*, *M. canetti* and *M. microti*, that share more than 99% nucleotide identity at some genes but are different in biochemistry, morphology, host range and produce different disease patterns in the infected animals (Cosma, Sherman and Ramakrishnan, 2003). Although, *M. tuberculosis* is considered as main cause of tuberculosis, with 1.5 million deaths annually, *M. africanum* infects more people in some parts of Africa than *M. tuberculosis*. *M. bovis* can infect both humans and cattle and was the main cause of tuberculosis before pasteurization of milk. *M. canetti* infections are rare and *M. microti* only infects voles. *M. tuberculosis*, *M. africanum* and *M. bovis* produce similar diseases in humans. Attenuated forms of *M. bovis* BCG has been used as vaccines (Cosma, Sherman and Ramakrishnan, 2003). *M. tuberculosis* is an acid-fast, rod shaped bacillus with an unique cell envelope (Kaufmann, 2001). This fact that stained bacteria can not be de-stained by an acid, acid fast, is used as a laboratory test to determine if the sample of tissue or blood is infected with *M. tuberculosis*. Bacilli are slow growing with a replication time of about 20 hours. This slow rate of growth represents a key reason for chronic nature of infection, complication in microbial diagnosis and long-term drug treatment (Kaufmann, 2001).

## **1.4 Pathology and immunology of *M. tuberculosis* infection**

TB is an air-borne disease, which transmits by inhalation of a few droplets containing 1-3 bacilli secreted from respiratory tract of an infected person with active TB (Figure 1-2). Expelled droplets can remain for several hours and only a single bacterium is sufficient for pathogenesis (Russell, 2007). The lung is the main entry port and site of infection but TB can be found in any tissue, known as extrapulmonary TB. The chance of developing TB outside the lungs is less than 10% (Kaufmann, 2001).

### **1.4.1 Phagocytosis of *M. tuberculosis* by macrophages**

Contrary to obligate extracellular pathogens, such as *Neisseria* species, which developed multiple mechanisms to avoid entry into phagocytes, *M. tuberculosis* is quite eager to gain entry to macrophages (Ernst, 1998). However, internalization by different cell surface receptors can change the fate of bacilli inside the macrophage. There are different receptors on the surface of macrophages, such as dendritic cell-specific intercellular adhesion molecule (ICAM)-3-grabbing non-integrin (DC-SIGN), mannose receptor (MR) and Fc receptors that are involved in the uptake of *M. tuberculosis* (Pieters, 2008). Internalizing the bacterium via Fc receptors, for instance, stimulates the inflammatory response inside the macrophage, while ingestion of mycobacterial cells through complement receptor type 3 (CR3) is dependent on the presence of cholesterol (Pieters, 2008 & Gatfield and Pieters, 2000). Cholesterol is essential for both recruiting and retaining TACO and mycobacterial uptake by host cells. TACO, also known as coronin 1, is a peripherally associated protein with membrane but without any transmembrane

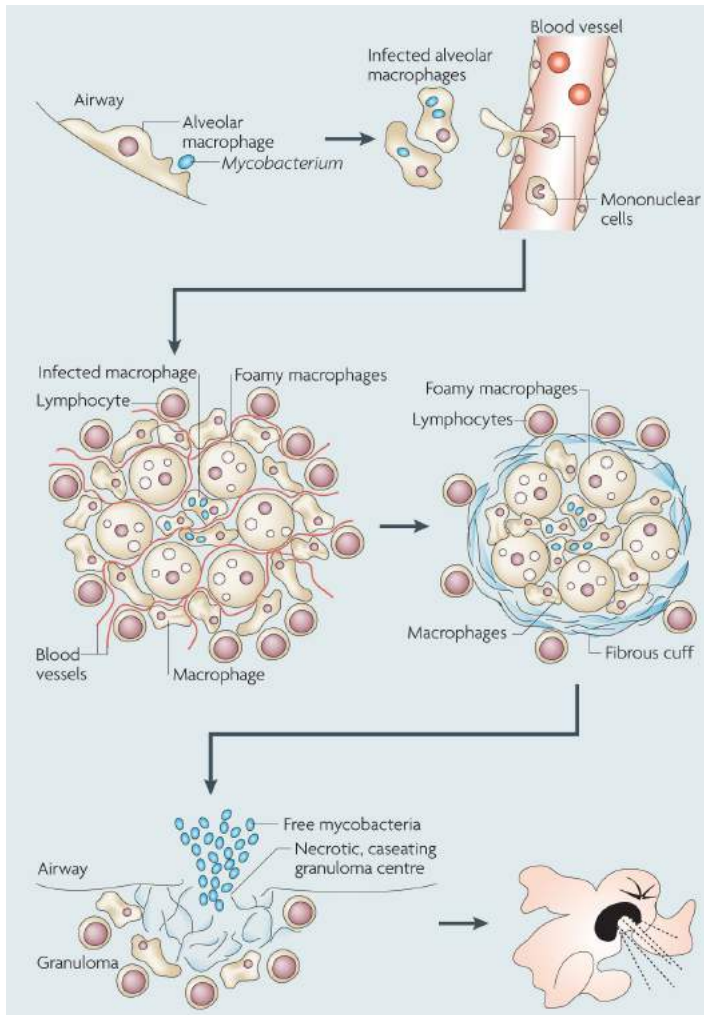
domains. Moreover, mycobacterium can consume cholesterol to survive inside the host cells (see 1.4.3). Therefore, internalization by some surface receptors is more beneficial to *M. tuberculosis* in long-term survival inside the host cells (Sambou et al., 2008; Pieters, 2008).

#### **1.4.2 Development of TB granuloma**

Following internalization by macrophages, bacilli exploit different strategies to ensure long-term survival inside host cells. In vitro studies of interactions between macrophages and *M. tuberculosis* bacilli have shown that the macrophages produce signaling factors, called cytokines, when ligands to the Toll-like receptor (TLR) molecules on the surface of bacteria interact with host TLRs. TLRs are transmembrane proteins composed of repeated leucine-rich motifs in the extracellular domain and a cytoplasmic domain, which activates the signaling pathways including NF- $\kappa$ B. A very well studied example of TLR agonists is 19 kDa lipoprotein from *M. tuberculosis*, which induces the production of interleukin-12 (IL-12) by human macrophages, upon interaction with host TLRs (Brightbill et al., 1999). IL-12 is essential for mediating immune responses and generating T-helper type 1-lymphocyte responses.

Tumor necrosis factor (TNF)- $\alpha$  is the main cytokine secreted by activated macrophages and involved in stimulating inflammatory response and therefore, recruiting other mononuclear cells to the site of infection (Russell, 2007). Recruited cells including foamy macrophages, mononuclear phagocytes and lymphocytes, as well as extracellular matrix

will surround the infected macrophages and form a very complex structure known as tubercle or granuloma (Figure 1-2). This structure is obvious in X-ray images of the lungs



**Figure 1-2.** The life cycle of *M. tuberculosis*. In lungs alveolar macrophages engulf the bacilli and by producing cytokines activate T-lymphocytes and recruit other mononuclear cells from neighboring blood vessels. Recruited cells will surround the infected macrophages and form a very complex structure known as tubercle or granuloma. In patients with attenuated immune system mycobacterium activates, breaks granuloma and enters the airways, a process know as caseation. Adopted from Russell, 2007, with permission.



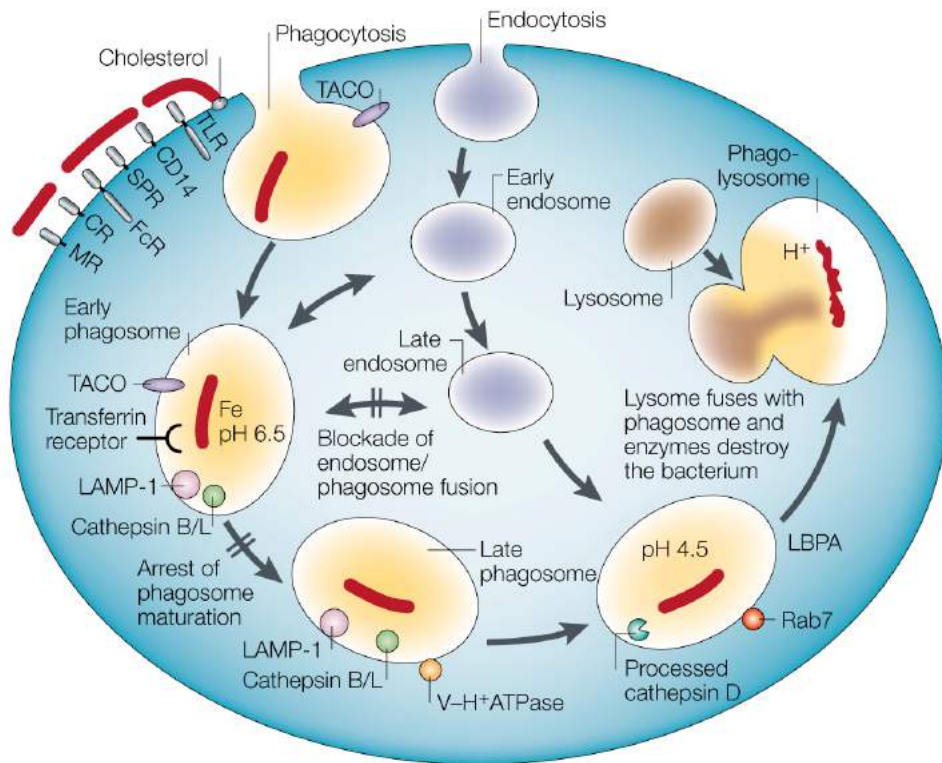
of infected person. Depending on the stage of disease, granuloma may contain both actively dividing and dormant bacilli. In fact granuloma represents a balance between *M. tuberculosis* and immune system until the immune system becomes weak, mainly due to old age, poor nutrition or co-infection with HIV, and infection turns into a life-threatening condition. In this case the cell walls of granuloma break down and thousands of viable, infectious bacilli are released into the air, a process known as caseation (Figure 1-2) (Russell, 2007 and Kaufmann, 2001).

#### **1.4.3 Survival inside macrophages**

*M. tuberculosis* survives inside the macrophage by halting phagosome maturation and therefore, preventing phagosome fusion with pre-formed lysosomes (Russell, 2007) (Figure 1-3). The first mechanism to prevent phagosome maturation is through recruiting tryptophan aspartate-containing coat protein (TACO) to the phagosome. Gatfield and Pieters (2000) showed that TACO binds to the phagosomal membrane only in the presence of cholesterol. On the other hand, cholesterol is essential for mycobacterial uptake by macrophages. Cholesterol depleted macrophages show substantial reduction in the mycobacterial uptake by about 85 to 90% (Gatfield and Pieters, 2000). Just prior to phagocytosis, the concentration of cholesterol at the site of bacterial entry increases thus; bacilli will be packed in cholesterol rich phagosomes. Cholesterol acts as a binding site for mycobacterial cells and promotes the interaction between bacilli and receptors on the surface of host cells (Gatfield and Pieters, 2000). *M. tuberculosis* responds to internalization by activating calcium flux pumps (Pieters, 2008). These pumps are dependent on TACO and following activation stimulate the production of phosphatase

calcineurin that is essential for preventing phagosome-lysosome fusion. The exact role of calcineurin in blocking lysosomal delivery of mycobacteria is not well understood but activated calcineurin is known to be involved in several downstream signalling pathways including the transfer of transcription factors to the nucleus and dephosphorylation of proteins involved in endocytosis (Pieters, 2008).

The second approach developed by *M. tuberculosis* to prevent phagosome-lysosome fusion is by preventing phosphatidylinositol 3-phosphate (PI3P) accumulation on phagosomal membranes (Pieters, 2008). Phosphatidylinositol 3-kinase (PI3 kinase) phosphorylates PI on position 3 (of inositol) to generate PI3P on the early endosomal and phagosomal membranes. PI3P acts as a docking site for several proteins involved in the fusion of phagosomes into lysosomes. *M. tuberculosis* uses two different strategies to block this trafficking event regulated by PI3P. The mycobacterial cell wall component lipoarabinomannan (LAM) interferes with the activity of PI3 kinase hVP34 and subsequently, prevents PI3P production on phagosomal membranes and phagosomal maturation into lysosomes. Additionally, *M. tuberculosis* prevents PI3P accumulation on phagosomal membranes via secreting SapM phosphatase into the host cell cytosol. SapM hydrolyzes PI3P accumulated on phagosomal membranes (Pieters, 2008).



**Figure 1-3.** Intracellular lifestyle of *Mycobacterium tuberculosis*. Following engulfing by alveolar macrophages, bacilli will end up in phagosome that has a very harsh environment. In order to survive, mycobacterium needs to arrest the phagosome maturation and subsequent fusion with preformed lysosomes. See section 1.4.3 for more details. However, some phagosomes mature and fuse with preformed lysosomes. Phagosomal maturation more likely happens in activated macrophages, producing IFN- $\gamma$  interferon. Adopted from Kaufmann, 2001, with permission.

PknG is also, another protein secreted by *M. tuberculosis* that is involved in preventing phagosome-lysosome fusion. PknG is one of the several Ser/Thr kinases encoded by *M. tuberculosis* and one of two kinases (the second soluble kinase is PKnK) that is located in the cytosol and is not anchored into the membrane by a transmembrane helix (Pieters, 2008). PknG is necessary for growth in vivo and is highly conserved in the genome of all pathogenic mycobacteria (Pieters, 2008). The exact mechanism of PknG in blocking host trafficking machinery is not well known but probably as a kinase phosphorylates a host molecule that is involved in phagosome maturation into lysosomes (Pieters, 2008).

However, mycobacterium cannot always prevent phagosomal maturation and some phagosomes will succeed to mature and merge with endosomes prior to fusion with lysosomes. Phagosomal maturation more likely happens in activated macrophages, producing IFN- $\gamma$  interferon. IFN- $\gamma$  can stimulate anti-microbial mechanisms, notably reactive oxygen intermediates ROI and reactive nitrogen intermediates RNI, in macrophages (Kaufmann, 2001).

In vitro analysis of infected macrophages grown in culture indicates that the bacterium inside the phagosome is not isolated from the extracellular environment and can communicate with extracellular environment through fusion with recycling endosomes (Russell, 2007). For instance, mycobacterium obtains the iron, which is essential for intracellular survival, from endosomes using specialized iron-scavenging molecules (Kaufmann, 2001).

#### **1.4.4 Beyond phagocytes**

Mycobacterial components can increase the success of bacterium by manipulating the inflammatory response and causing persistence and subsequently late stage damage and transmission (Russell, 2007). It has been shown that cell wall components lipoarabinomannan and arabinomannan are released inside the infected macrophages and trafficked into the uninfected cells. Following the release of cell wall components into the *M. tuberculosis* vacuole, they accumulate in multilamellar vesicles. Multilamellar bodies come together in the multivesicular lysosomes known as major histocompatibility complex (MHC) class II-enriched compartments, which are released from infected macrophages into the extracellular environment through exocytosis. Exosomes are taken up by neighboring cells. If these lipids have biological activities, then their release can enhance the mycobacterium influence without instant prevention by host cells.

#### **1.5 Current treatment of TB and rise of MDR and XDR strains**

The current treatment for tuberculosis is based on using a cocktail of drugs, isoniazid, rifampicin, ethambutol and pyrazinamide, known as first line anti-TB drugs for duration of six months (Table 1-1 and 1-2). You can find a comprehensive list of anti-TB drugs with their targets in the Supplements section (Table S9-1). The regimen starts by using all four drugs everyday for two months and continues by using isoniazid and rifampicin for additional four months. This regimen has proved to be efficient in more than 95% of cases, especially in patients under directed observed therapy (DOT) and is prescribed to both pulmonary and extrapulmonary TB patients (Zumla, Nahid and Cole, 2013). However, there are some disadvantages attached to this regimen including: drug

intolerance and toxicities, causing interruptions in the treatment or changing the regimen; increasing the treatment duration in cases with reoccurrence TB and potential drug-drug interactions, especially in patients co-infected with HIV (Zumla, Nahid and Cole, 2013).

The concern about TB is on rise as some stains of *M. tuberculosis* have developed resistance against current anti-TB drugs (Koul et al., 2011). Strains resistant to isoniazid and rifampicin are known as multi-drug resistant (MDR) and those resistant to first line drugs as well as one or more fluoroquinolones (such as ciprofloxacin, levofloxacin, moxifloxacin and ofloxacin) and injectables (amikacin, capreomycin, kanamycin) are referred to as extensively drug resistant (XDR) (Dye, 2009). It is estimated that there are 650,000 new cases of MDR-TB annually (Zumla, Nahid and Cole, 2013). The treatment duration for MDR patients with no history of previous TB is about 20 month and for those with previous records of MDR-TB treatment is 28 month. The treatment includes a combination of four second-line drugs everyday and under DOT (Zumla, Nahid and Cole, 2013). Treating XDR-TB is more challenging than MDR, as it requires administering third-line anti-TB drugs, which are more expensive and usually have more side effects. The most recent categorization of TB includes a new generation of strains that are resistant to all first and second-line anti-TB drugs known as totally drug-resistant tuberculosis or TDR (Zumla, Nahid and Cole, 2013). Although the rate of TDR incidence is low, it illustrates the importance and urgency of developing new anti-TB drugs.

It has been shown, however, that incidence of MDR can be halted and reversed by good programmed management, including prompt diagnosis and appropriate treatment of drug-

resistant TB, as it has been illustrated in counties such as Estonia, Hong Kong and USA (Dye, 2009). Also, introducing new anti-TB drugs in recent years have increased our chance in combat against *M. tuberculosis* (Table S9-2). Developing new drugs that suppress new essential targets in *M. tuberculosis* can also help to overcome the TB resistance problem (Lamichhane, 2011). You can find list of recently identified targets in *M. tuberculosis* in supplements table S9-3.

**Table 1-1.** Classification of anti-TB drugs according to their efficacy in TB treatment. They are used to treat susceptible and resistant strains of tuberculosis. First line anti-TB drugs are used to treat drug-susceptible tuberculosis in a regimen consists of four drugs for duration of six month. Second line drugs are prescribed for drug-resistant TB and third line anti-TB drugs are not well defined in terms of efficacy and target.

Classification	Drugs
First line anti-TB drugs (oral)	Ethambutol, isoniazid, pyrazinamide, rifampicin/rifampin, rifambutin, rifapentine
Second line anti-TB drugs (injectables)	Amikacin, kanamycin, capreomycin, streptomycin, viomycin
Second line anti-TB drugs (oral and injectable fluoroquinolones)	Ciprofloxacin, gatifloxacin, levofloxacin, moxifloxacin and ofloxacin
Third-line anti-TB drugs	Amoxicillin plus clavulanate, clarithromycin, clofazimine, imipenem plus cilastatin and linezolid

**Table 1-2.** Approved drugs to treat tuberculosis. Most commonly used drugs in TB treatment. You can find comprehensive list of all approved anti-TB drugs in Appendix Table 1 (Brennan, Young and Robertson, 2008).

Drug name	Drug target/mechanism	Drug resistance mechanism	In vitro activity
Ethambutol (ETH)	Inhibits cell wall biosynthesis through inhibiting arabinosyl transferase	Mutations in <i>M. tuberculosis</i> <i>embA</i> or <i>embB</i>	Active against <i>M. tuberculosis</i> (H37Rv) with MIC 0.5 µg/ml
Isoniazid (INH)	A prodrug activated by KatG Inhibits fatty acid synthesis by inhibiting InhA	Mutation in <i>inhA</i> and activating enzyme KatG	MIC against <i>M. tuberculosis</i> (H37Rv) 0.025 µg/ml
Rifampicin	Inhibits transcription by binding to RNA polymerase β-subunit	Mutations in <i>rpoB</i> gene	
Pyrazinamide (PZA)	It is not well understood but experimental evidence suggests that PZA converts to pyrazinoic acid (POA) by pyrazinamidase (PZAase). POA can inhibit major processes within the cell, for instance fatty acid synthase and membrane transport, through acidifying the cell	Mutations in PZAase	MIC against <i>M. tuberculosis</i> (H37Rv) 6-50 µg/ml at pH 5.5



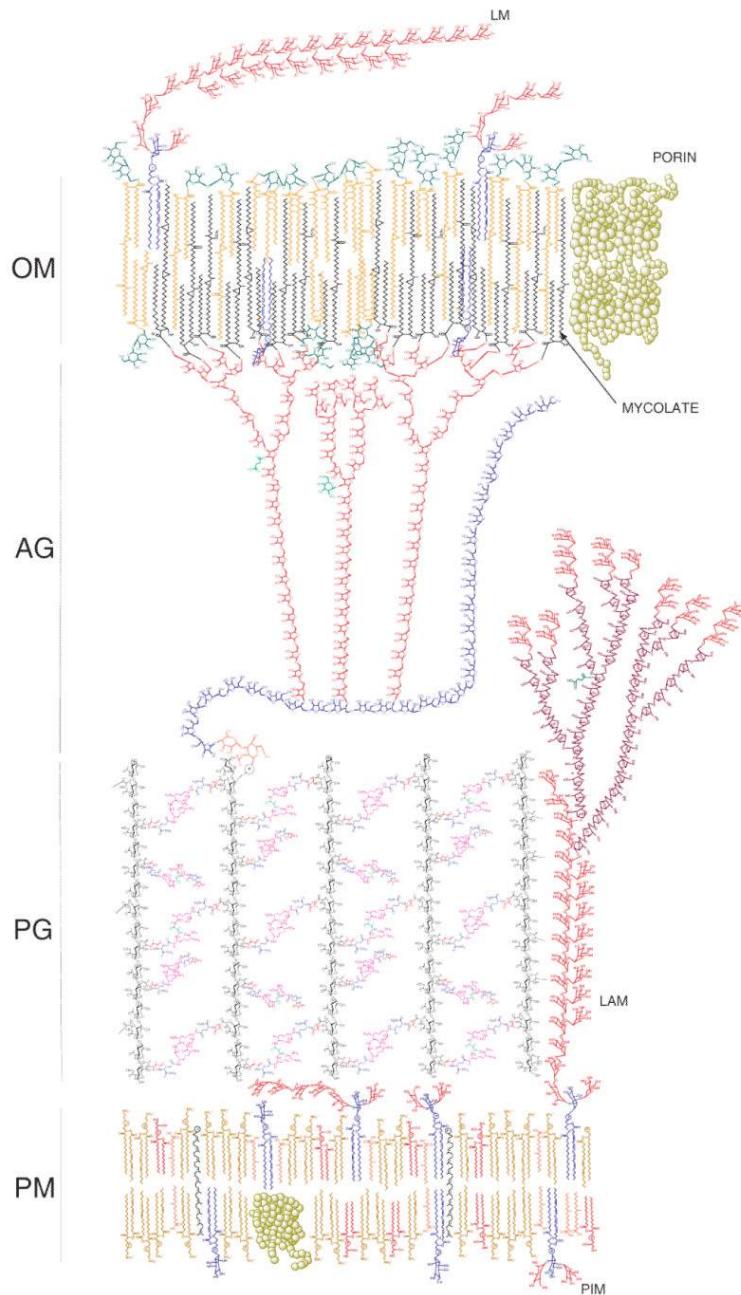
Continued from last page. Second line anti-TB drugs.

Drug name	Drug target/mechanism	Drug resistance mechanism	In vitro activity
Amikacin (AMI)	Inhibits protein synthesis by binding to the A site of 16S rRNA in the 30S ribosomal subunit	Ribosomal changes in the 16S rRNA	Bactericidal activity against all the drug-sensitive clinical isolates of <i>M. tuberculosis</i> at 2 µg/ml
Capreomycin (CAP)	Inhibits protein synthesis	Ribosomal changes in the 16S rRNA	Active against <i>M. tuberculosis</i> (MIC 2µg/ml) and <i>M. avium</i>
Cycloserine (CYS)	Inhibits peptidoglycan synthesis	Overexpression of the alanine racemase gene in <i>M. smegmatis</i>	Active against <i>M. tuberculosis</i> (H37Rv) with MIC 25 µg/ml
Ethionamide (ETA)	Inhibiting enoyl-ACP reductase encoded by <i>inhA</i> gene	Mutation in <i>inhA</i> gene	Active against <i>M. tuberculosis</i> (H37Rv) with MIC 0.25 µg/ml
Kanamycin (KAN)	Inhibits protein synthesis by binding to the A site of 16S rRNA in the 30S ribosomal subunit	Ribosomal changes in the 16S rRNA	MIC against <i>M. tuberculosis</i> (H37Rv) 2 µg/ml
Para-aminosalicylic acid (PAS)	Inhibits thymidylate synthase, ThyA	Mutations in <i>thyA</i> gene	MIC against <i>M. tuberculosis</i> (H37Rv) 0.3-1 µg/ml
Streptomycin (STR)	Similar to amikacin and kanamycin, inhibits protein synthesis through binding to the A site of 16S rRNA in the 30S ribosomal subunit	Ribosomal changes in the 16S rRNA and ribosomal protein S12	MIC against <i>M. tuberculosis</i> (H37Rv) 1 µg/ml

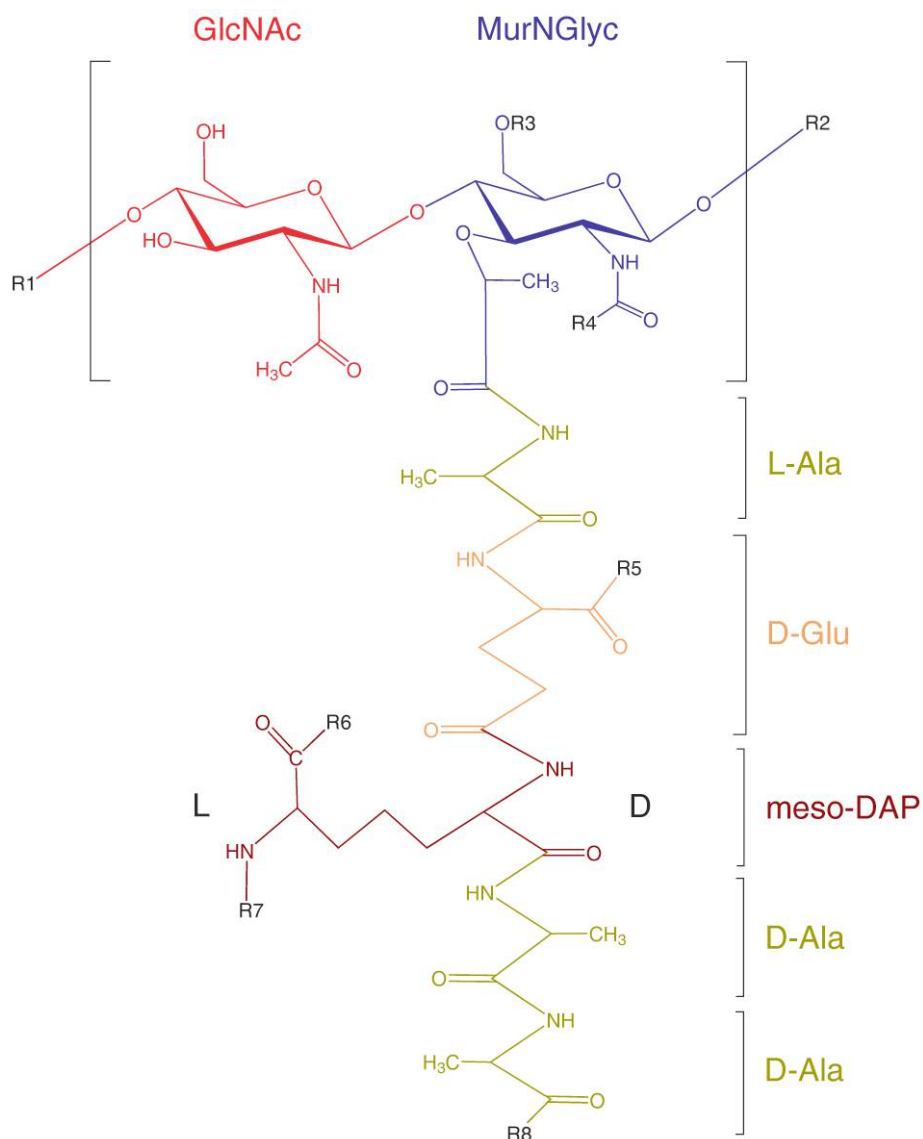
## 1.6 The mycobacterial cell envelope

### 1.6.1 Peptidoglycan

The unique cell envelope of *M. tuberculosis* composed of plasma membrane, cell wall and capsule protects the bacilli against the antibiotics and is involved in pathogenesis (Brennan, 2003) (Figure 1-4). The plasma membrane is a typical bacterial membrane but paramount to the mycobacterium survival (Daffe and Etienne , 1999). The cell wall core represents the hallmark of mycobacterial cell envelope and consists of mycolyl-arabinogalactan-peptidoglycan (mAGP) complex (Figure 1-4). Like peptidoglycan (PG) in Gram-negative bacteria, *M. tuberculosis* peptidoglycan is responsible for maintaining the cellular shape and osmotic stability and is composed of *N*-acetyl- $\alpha$ -D -glucosamine (GlcNAc) and a modified muramic acid (Mur) cross linked with L-alanyl- D -isoglutaminy-meso-diaminopimelyl- D -alanine (L-Ala- D -Glu-A<sub>2</sub>pm- D -Ala) with the Glu being further amidated (Figure 1-5). However, the main distinct feature of *M. tuberculosis* peptidoglycan, compare to other bacteria, is that some or all of the Mur residues are *N*-acetylated with glycolic acid (MurNGlyc) rather than *N*-acetylated (MurNAc). Peptide cross-links are composed of cross-links between two A<sub>2</sub>pm residues or between A<sub>2</sub>pm of MurNGlyc from one strand and D -Ala from MurNGlyc from a second strand (Brennan, 2003 and Kaur et al., 2009). A specific linker, a rhamnosyl residue attached to a *N*-acetylglucosaminosyl-1-phosphate residue, attaches peptidoglycan to the galactan domain of arabinogalactan (Kaur et al., 2009).



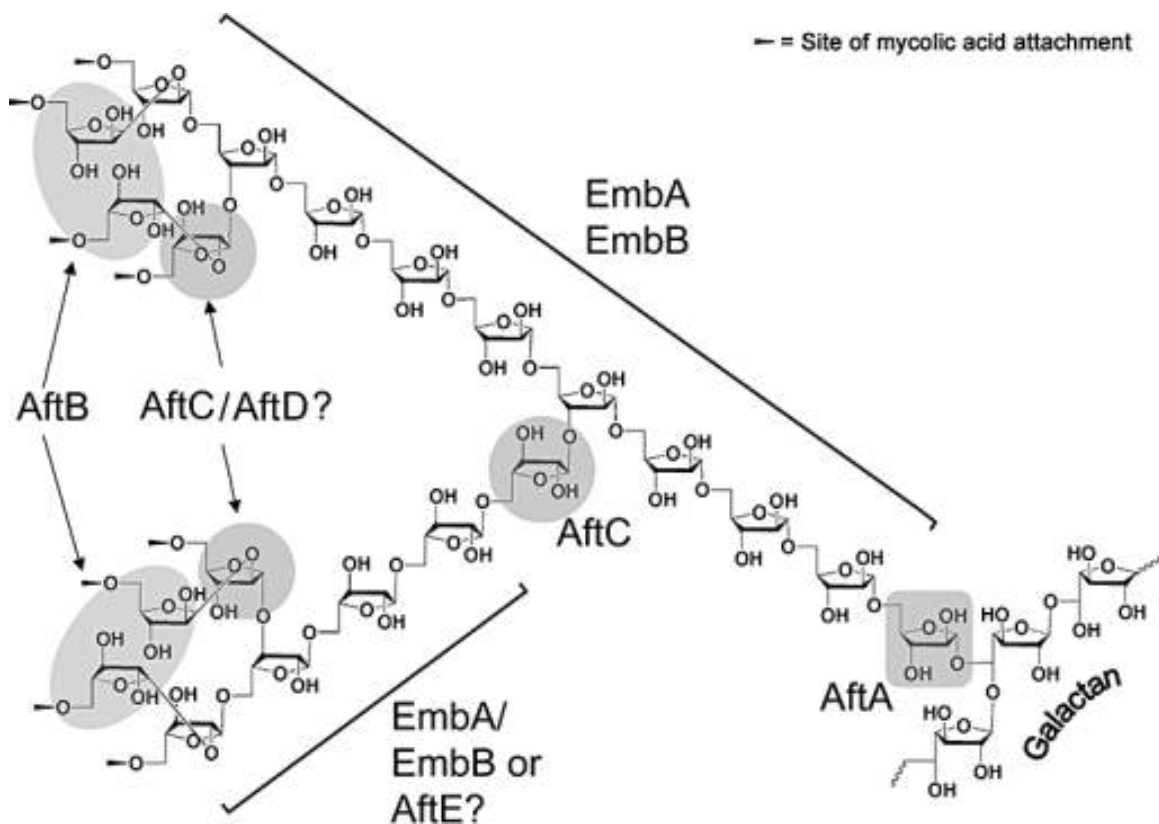
**Figure 1-4.** The *Mycobacterium tuberculosis* cell envelope. The envelope is composed of plasma membrane, cell wall core and capsule. Peptidoglycan, arabinogalactan and mycolic acids are the major constituents of the cell wall. Peptidoglycan contains *N*-acetyl muramic acid and *N*-acetyl glucosamine cross-linked by short peptides. Arabinogalactan is composed of arabinan chains attached to a linear galactan chain. A linker connects the peptidoglycan to the galactan part of arabinogalactan. The most outer layer of cell envelope is capsule that plays key roles in pathogenicity and resistance of mycobacterium to antibiotics. Adopted from (Kaur et al., 2009), with permission.



**Figure 1-5.** Structure of *M. tuberculosis* peptidoglycan before cross-linking. R1 and R2 represent *N*-glycolylmuramic acid and *N*-acetylglucosamine residues of adjacent monomers; R3, H or the linker unit of arabinogalactan; R4, H, CH<sub>3</sub> (*N*-acetyl) or CH<sub>2</sub>OH (*N*-glycolyl); R7, H or cross-link to last **D**-Ala or to the **D**-center of another meso-DAP residue; R5, R6, R8 OH, NH<sub>2</sub> or OCH<sub>3</sub>. Picture from (Kaur et al., 2009), with permission.

### 1.6.2 Arabinogalactan

The function of arabinogalactan (AG) is not well understood but apparently it produces a thick hydrophilic region between the PG and mycolic acid layer (Kaur et al., 2009). AG is composed of two domains, a galactan domain and one arabinan domain (Figure 1-4). The galactan domain is made of 30  $\beta$ -(1 $\rightarrow$ 5) and  $\beta$ -(1 $\rightarrow$ 6) linked galactofuranose (Gal $f$ ) residues attached to three arabinan chains at position 8, 10 and 12, each containing 31  $\beta$ -(1 $\rightarrow$ 2),  $\alpha$ -(1 $\rightarrow$ 5) and  $\alpha$ -(1 $\rightarrow$ 3) linked arabinofuranose (Ara $f$ ) residues (Tam and Lowary, 2009) (Figure 1-4 and 1-6). The molecular mechanism of AG biosynthesis has been well studied. Linker, galactan and arabinan domains of AG are synthesized on decaprenyl phosphate (Dec-P) and transferred to PG by an unknown ligase (Kaur et al., 2009). The first Gal $f$  residues are attached to the linker unit by galactosyltransferase Rv3782 (GlfT1) and the subsequent Gal $f$  residues are added by Rv3808c (GlfT2). Membrane-associated glycosyltransferases (GTs) are responsible for elongation of the arabinan domain of AG. So far seventeen GTs have been identified (Kaur et al., 2009). AftA (Rv3792) catalyzes the addition of first Ara $f$  residue to the galactan chain. AftB (Rv3805c) and AftC (Rv2673) catalyze the addition of  $\beta$ -(1 $\rightarrow$ 2) and  $\alpha$ -(1 $\rightarrow$ 3) linked Ara $f$  units to the non-reducing termini of the arabinan domain of AG, respectively. EmbA and EmbB act as  $\alpha$ -(1 $\rightarrow$ 5) arabinofuranosyl transferases that extend the arabinan chain after the first Ara $f$  residue (Figure 1-6).

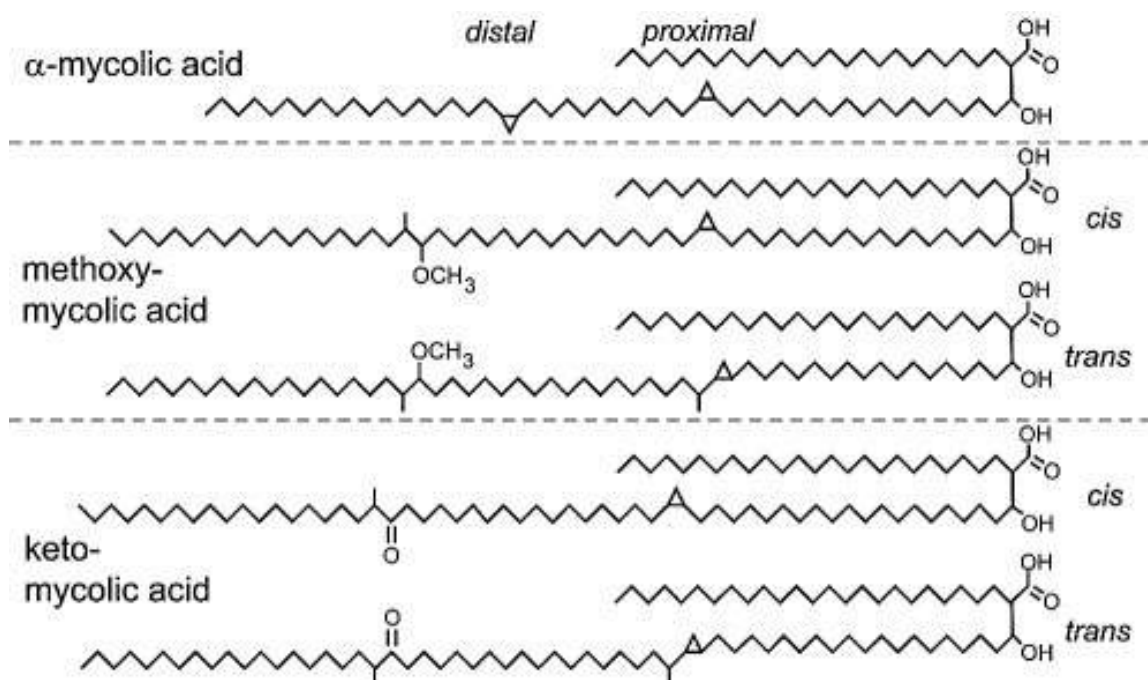


**Figure 1-6.** Chemical structure of arabinogalactan (AG) from *M. tuberculosis* and role of AftA, AftB, AftC, EmbA and EmbB in its synthesis. From Birch et al., 2008. For more information see section 1.6.2, with permission.

### 1.6.3 Mycolic acids

Mycolic acids composed of long chain ( $C_{70}$ - $C_{90}$ )  $\alpha$ -alkyl,  $\beta$ -hydroxy fatty acids, are a major component of the cell wall in *M. tuberculosis* (Takayama, Wang and Besra, 2005) (Figure 1-7). Mycolic acids are highly insoluble and along with free glycolipids called trehalose dimycolate (TDM), are responsible for impermeability of the mycobacterial cell envelope. They protect mycobacterium against antibiotics and the host immune system (Takayama, Wang and Besra, 2005 and Sani et al., 2010). The cyclopropane rings of mycolic acids protect bacillus against oxidative stress (Takayama, Wang and Besra, 2005).

Recent cryo-transmission electron microscopy (EM) data has shown that mycolic acids, referred to as mycomembrane, form a structure similar to the outer membrane of Gram-negative bacteria around the arabinogalactan-peptidoglycan layer (Sani et al., 2010). Mycolic acids can be divided into three groups  $\alpha$ , methoxy and keto mycolic acids (Figure 1-7). More than 70% of mycolic acids are  $\alpha$ -mycolic acids and only 10 to 15% of mycolic acids are methoxy and keto-mycolic acids.

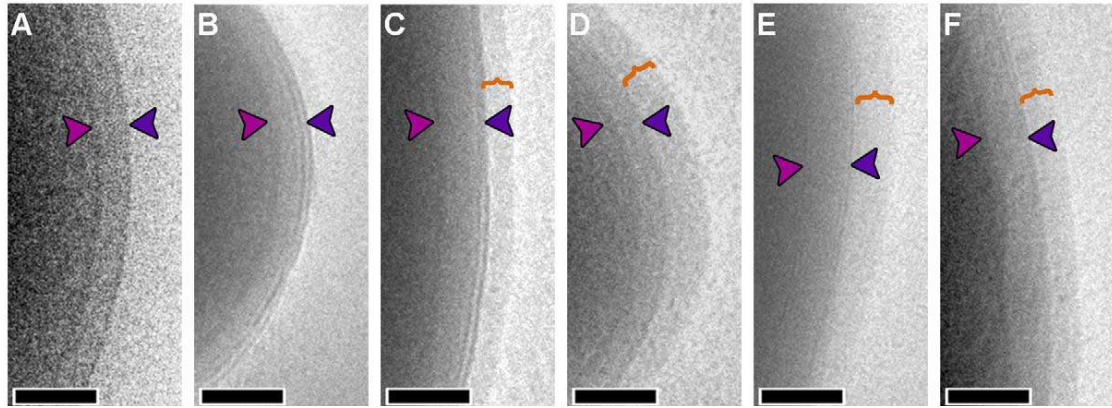


**Figure 1-7.** Chemical structure of mycolic acids from *M. tuberculosis*. For more information see section 1.6.3. Adopted from (Takayama, Wang and Besra, 2005). With permission.



#### 1.6.4 Capsular layer

The term capsule was used for the first time by Chapman and colleagues in 1958 (Chapman, Hanks and Wallace, 1959). Using ultrathin sectioning and electron microscope they observed a loose disposed membrane around the cell wall in *M. lepraemurium*. They proposed that tissue cells are likely responsible for producing capsule since this layer is in remote distance from bacterial cells and often two or three bacteria can be found within one membrane. (Chapman, Hanks and Wallace, 1959). Visualizing capsular layer and determining its nature was abandoned since Chapman study until recently when Sani and colleagues reasoned that growing the cells without detergent and agitation might be the key to detect the capsule in an intact form (Sani et al., 2010). They also used Gram-negative bacterium *Shigella flexneri* as a control (Figure 1-8A). *M. smegmatis* cells grown in the presence of Tween-80 and agitation (to prevent clumping) showed a cell envelope structure similar to *S. flexneri* (Figure 1-8B). But *M. smegmatis* and other mycobacterial species cells grown without perturbation showed an extra layer surrounding their mycomembrane (Figure 1-8 C-F).



**Figure 1-8.** Visualization of capsule in native-state using electron microscope. (A) Cryo-electron micrographs of intact *S. flexneri* cells, (B) *M. smegmatis* cells grown in the presence of detergent with agitation, which show a cell envelope similar to *S. flexneri*. (C) *M. smegmatis* cells grown without perturbation showing the presence of an extra layer (bracket) around cell wall. Observing capsule around the mycomembranes of *M. tuberculosis* (D), *M. marinum* (E) and *M. bovis* BCG (F). From Sani et al., 2010, with permission.

In order to support this observation that the extra layer is capsule, they immuno-labeled one of the mycomembrane proteins, OmpATb porin on *M. smegmatis* cells overexpressing this protein. Antibody efficiently labeled mycomembrane in both perturbed and non-perturbed conditions, but in latter condition, the capsular layer was distinct from the mycomembrane. This study revealed that both pathogenic and non-pathogenic strains contain a capsular layer around the cell wall and the capsule is distinct from the mycomembrane (Sani et al., 2010).

#### **1.6.4.1 Capsular functions**

In an earlier study by Cywes and colleagues, it was shown that capsular D-glucan mediates the direct (nonopsonic) binding of bacilli to the phagocytic receptor CR3 in the absence of serum (Cywes et al., 1997). In other words, capsular polysaccharides can act as ligands for direct binding to CR3 receptors. Binding to the CR3 lectin site leads to receptor activation and stimulation of phagocytosis (Cywes et al., 1997). Interestingly, phagocytosis via CR3 is considered as a safe portal of entry for *M. tuberculosis* because phagocytosis through this receptor leads to inhibition of respiratory burst and production of interleukin 12 (IL-12) (Ehlers and Daffe, 1998). Daffe and Etienne also demonstrated that capsule mediates the adhesion and penetration of bacilli into the host cells, since capsule removal significantly reduces the bacterial binding to the macrophages (Daffe and Etienne, 1999 and Sani et al., 2010).

Capsule removal by treating the *M. bovis* BCG culture with detergent significantly reduced the amounts of IL-12p40, IL-6 and TNF $\alpha$  (about 30%) indicating the role of capsule in diminishing the pro-inflammatory responses (Sani et al., 2010). Moreover, Sambou and his group showed that reduction in capsular glucan level, by disrupting  $\alpha$ -glucan biosynthesis enzyme GlgA, diminishes the ability of bacilli to persist in mice (Sambou et al., 2008). In a more recent study, Geurtsen found that the C-type lectin DC-SIGN acts as a cellular receptor for *M. tuberculosis*  $\alpha$ -glucan (Geurtsen, et al., 2009). This binding induces the production of cytokine IL10. IL10 production prevents dendritic cells (DCs) maturation (during maturation DCs convert from antigen capturing to antigen

presenting mode) and subsequently eliminates the antigen-presenting function of DCs and initiating the immune response (Geurtsen, et al., 2009).

Several distinct proteins have been detected in the capsular extracts of *M. marinum* including EspB (MM5457), EspF (MM5440), fatty acid synthase Fas (MM3962), iron-regulated elongation factor Ef-tu Tuf (MM1014), chaperonin GroEL1 (MM1126), KatG (MM2914), ESAT-6, CFP-10 and PPE68\_1 (Sani et al., 2010). The majority of these proteins are secreted by ESX-1, which is a type VII secretion system, known for its role in mycobacterial pathogenicity (Figure 1-9) (Stoop, Bitter and van der Sar, 2012).

In a recent study Weerd and colleagues investigated the epitope specificity of the anti-glycogen IgM antibody (Mab, IV58B6) against  $\alpha$ -glucan polymers and glycogen using saturation transfer difference NMR (STD NMR) technique (Weerd et al., 2015). They showed that Mab is able to bind glycogen and capsular  $\alpha$ -glucan and identifies the inner  $\alpha$ -(1-4)-linked glucose as epitope. Additionally, they used Mab to create a transposon-library based screening technique to determine putative genes involved in capsular  $\alpha$ -glucan biosynthesis. To achieve this, they cultured the transposon mutant libraries over nitrocellulose filters set on Middlebrook 7H10 agar plates. Then, they generated a back up by transferring the nitrocellulose filters on top of new filters (second filters) and cultured them on fresh Middlebrook 7H10 agar plates. In this way the capsular  $\alpha$ -glucan diffused to the second filters and first filters were kept aside as back up. The second filters were immunostained with Mab to detect the amount of capsular  $\alpha$ -glucan and compared with the level of capsular  $\alpha$ -glucan in wild type strain determined by the same

technique. Using this approach they identified 140 transposon mutants participating in biosynthesis of capsular  $\alpha$ -glucan including well-known enzymes in  $\alpha$ -glucan production, GlgB, GlgC and GlgP (Weerd et al., 2015).

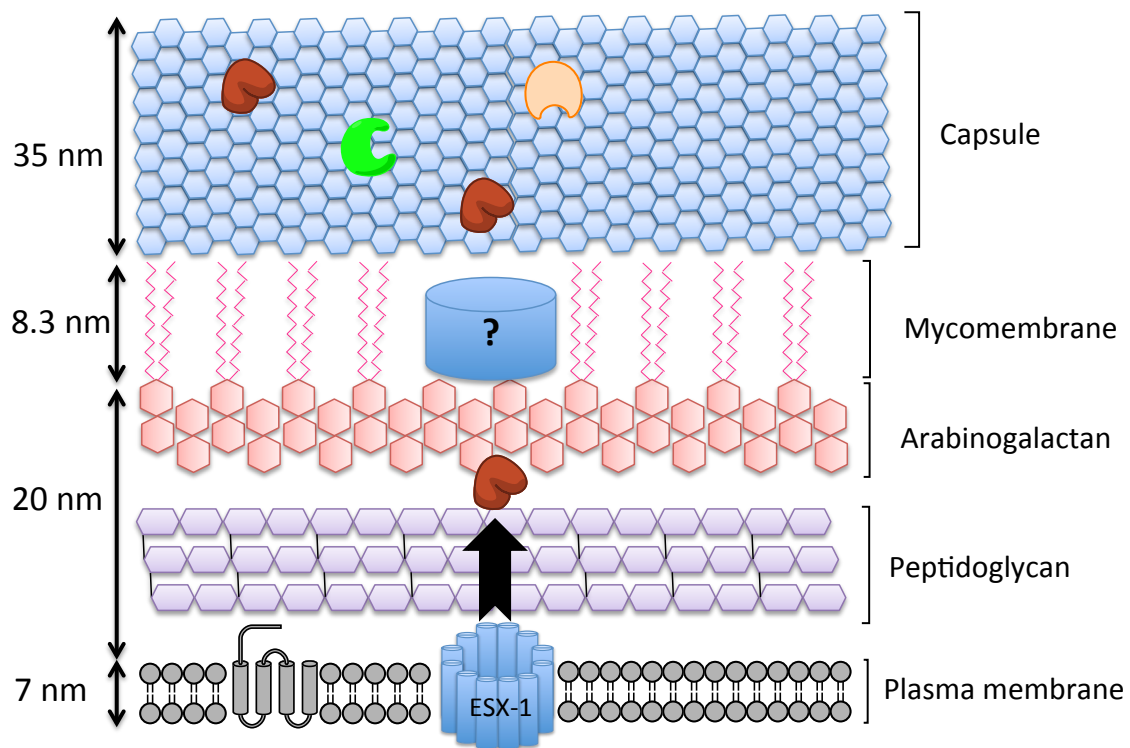
#### **1.6.4.2 Capsular composition**

Capsular layer is mainly composed of polysaccharides and proteins (about 97% of the total material) and small amount of lipids. The main polysaccharide is  $\alpha$ -D-glucan with a molecular weight of about 100 kDa. Glucan has a similar structure to intracellular glycogen,  $\rightarrow 4\text{-}\alpha\text{-D-Glcp-1}\rightarrow$  core substituted at some six positions with short chains, though, former has higher molecular weight and is more compact than the glycogen (Dinadayala, et al. 2008). Arabinomannan and mannan with 13 kDa and 4 kDa molecular weights, respectively, are also frequently found in the capsule. The mannan segment of arabinomannan and mannan consisted of  $\rightarrow 6\text{-}\alpha\text{-D-Manp-1}\rightarrow$  core substituted with a  $\alpha$ -D-mannosyl residue at some positions 2 (Daffe and Etienne , 1999).

The outermost capsule is a complex mixture of proteins including secreted proteins, exported to the extracellular medium, cell envelope-associated or cytoplasmic proteins. Since the secreted proteins are limited around the bacilli and in contact with the host membrane, it is very difficult to distinguish between secreted proteins and surface-associated proteins. Some of the main capsular proteins of *M. tuberculosis* are summarized in Table 1-3. Capsular proteins, such as  $\beta$ -lactamase Y49, are involved in the resistance of bacilli to  $\beta$ -lactam antibiotics, or other capsular proteins, especially catalase/peroxidase and superoxide dismutase protect the bacilli against the host's

microbicidal mechanisms by detoxification of reactive oxygen intermediates (Daffe and Etienne , 1999). Capsular lipids form only a small fraction of the capsule and located either on the cell surface, e. g. phospholipids, species specific lipids (dimycocerosate of phthiocerol) and type specific lipids of *M. tuberculosis* (PGL), or present in the inner parts of the capsule, such as trehalose dimycolates (Daffe and Etienne , 1999).

**$\alpha$ -D-glucan.** Glucan as the main constituent of the mycobacterium capsule is synthesized inside the cytoplasm and is transported to the cell surface by as yet unknown transporters. Several secretion systems and porins, such as ESX-1 and OmpATb have been identified in the mycobacterium cell wall that are involved in the transportation of macromolecules through cell wall, but their roles in the glucan translocation is not still determined (Figure 1-9). ESX-1 is a type VII secretion system located on the plasma membrane and responsible for secretion of potent T cell antigens such as ESAT-6 and CFP-10. In *M. marinum* a large amount of capsular proteins are in fact secreted by ESX-1 pathway (Sani et al., 2010). OmpATb is a porin present in the mycomembrane of *M. tuberculosis* and is essential for the passage of small hydrophilic molecules through the cell wall (Senaratne, et al. 1998).



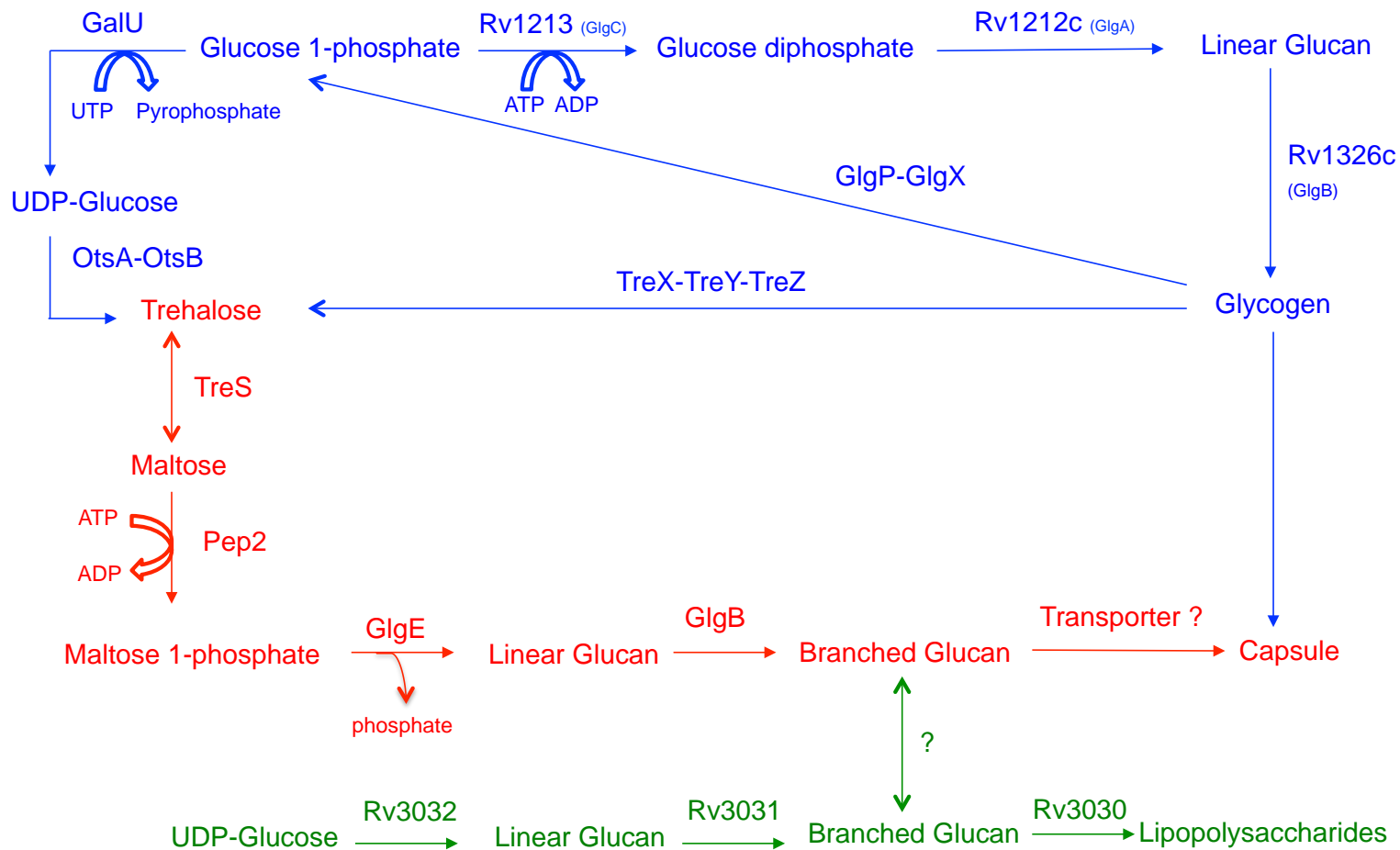
**Figure 1-9.** Schematic of the spatial organization of mycobacterial cell envelope layers and localization of the ESX-1 secretion system. The size and relative position of different layers of cell envelope including plasma membrane, arabinogalactan, mycolic acids and capsule has been depicted. The majority of proteins in the capsule are secreted by ESX-1 secretion system. Generated using ChemBioDraw.

**Table 1-3.** Some of the main capsular proteins of *M. tuberculosis* (Daffe and Etienne, 1999).

Protein (synonym)	Gene	Function
β-lactamase Y49 (Gen Bank 1370234)	<i>blaC</i>	β-lactam antibiotics resistance
FAP-TB	<i>apa</i>	Fibronectin-binding protein (antigen)
MPT53	<i>mpt53</i>	15 kDa antigen
MPT63	<i>mpt63</i>	16 kDa antigen
MPT64	<i>mpt64</i>	23.5 kDa antigen
Mycoloyl transferase A, B, C	<i>fbpA</i>	Cell wall synthesis
	<i>fbpB</i>	Fibronectin-binding protein
	<i>fbpC1</i>	T-cell antigens
Catalase/Peroxydase (MPT35)	<i>katG</i>	Survival in host cells, Reactive oxygen intermediates detoxification
CFP 29	<i>Cfp29</i>	T-cell antigen
ESAT-6	<i>Esat-6</i>	T-cell antigen
Heat Shock Protein Acr (16 kDa alpha-crystallin HSP)	<i>hspX</i>	Resistance to oxygen limitation
Nicotinamidase (pyrazinamidase)	<i>pncA</i>	Activation of pyrazinamide
Superoxide dismutase (MPT58)	<i>sodA</i>	Reactive oxygen intermediates detoxification
Thioredoxine (MPT46)	<i>trxA</i>	Reactive oxygen intermediates detoxification



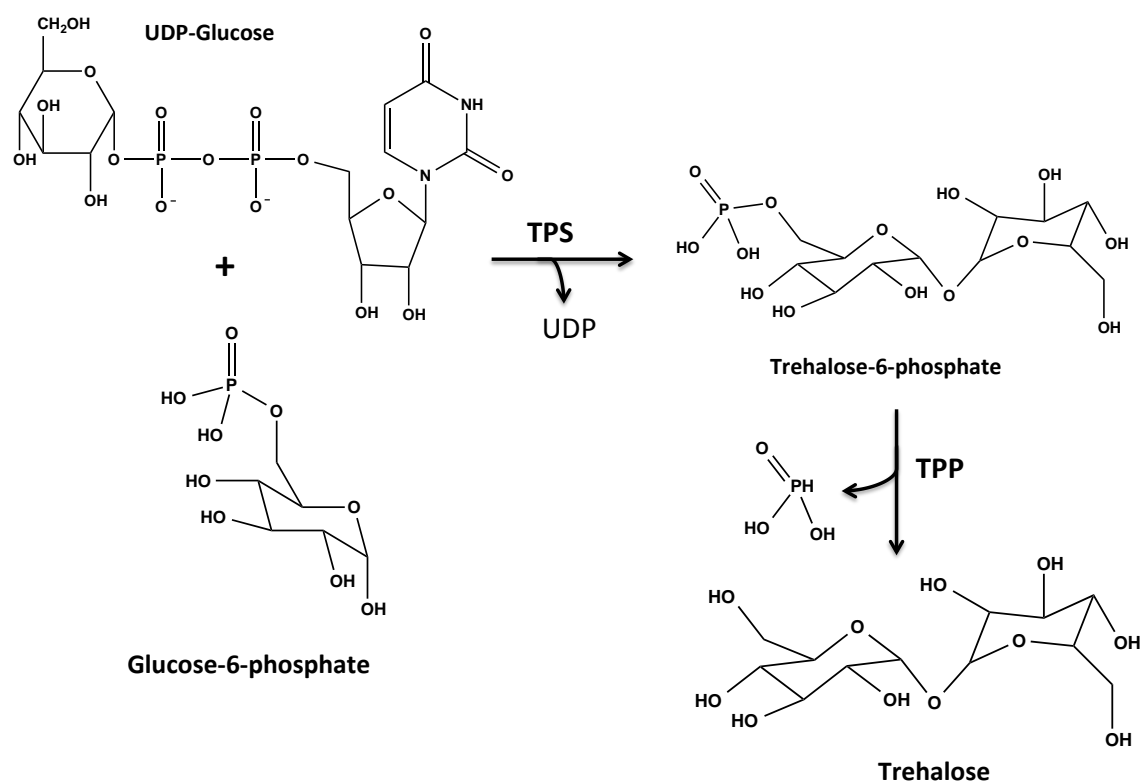
Mycobacteria produce three  $\alpha$ -glucans; cytosolic glycogen, capsular  $\alpha$ -glucan and methylglucose lipopolysaccharide in three interlinked pathways (Chandra, Chater and Bornemann, 2011) (Figure 1-10). The classical GlgC-GlgA pathway generates cytosolic glycogen in a series of reactions from glucose 1-phosphate. Diphosphoglucose pyrophosphorylase (GlgC) phosphorylates glucose 1-phosphate to glucose diphosphate. Subsequent polymerization of latter monosaccharide is mediated by glycogen synthase (GlgA) followed by introducing branches at some positions 6 by GlgB. The final glycogen is highly soluble and is used by bacterial cells as a carbon/energy storage molecule. In situations that glucose is needed two enzymes, glycogen phosphorylase (GlgP) and debranching enzyme (GlgX) breakdown glycogen to glucose 1-phosphate (Figure 1-10). Orthologous genes of GlgC-GlgA have been identified in *M. tuberculosis* (Sambou et al., 2008). In order to study the role of *glgC*, *glgA* and *treZ* genes in capsule biosynthesis and subsequently its impacts on persistence, Sambou and colleagues generated mutants of orthologous of these genes in *M. tuberculosis* and labeled them as H37Rv $\Delta$ *glgC*, H37Rv $\Delta$ *glgA* and H37Rv $\Delta$ *treZ*, respectively. They observed two fold decrease in glucan production in H37Rv $\Delta$ *glgC* and H37Rv $\Delta$ *glgA* but glucan production was not affected in H37Rv $\Delta$ *treZ*. Mice infected with H37Rv $\Delta$ *glgC* and H37Rv $\Delta$ *treZ* showed similar results to those mice infected with wild type H37Rv strain in terms of virulence and persistence. However, the ability of H37Rv $\Delta$ *glgA* mutants to persist in mice was significantly reduced indicating a direct role for the capsular glucan in persistence (Sambou et al., 2008).



**Figure 1-10.** Three interlinked pathways involved in the  $\alpha$ -glucan biosynthesis. Adapted from (Chandra, Chater and Bornemann, 2011).

The production of glycogen is highly connected to trehalose metabolism. Trehalose,  $\alpha$ ,  $\alpha$ -1, 1- glucosyl-glucose, is also another abundant source of carbon in bacterial cells that protects membrane and cellular proteins against environmental stresses such as cold, dehydration, heat and oxygen radicals. Trehalose, also, is involved in regulating glucose metabolism and is a component of various cell wall glycolipids in mycobacteria and other related organisms (Elbain et al., 2003 and Elbein, et al. 2010). Trehalose dimycolate (cord factor) composed of a trehalose core with the unusual fatty acid mycolic acid esterified at some position 6 is the most toxic lipid produced by *M. tuberculosis* and responsible for impermeability of the cell wall to different antibiotics (Zhang et al., 2011). There are two different pathways involved in the trehalose biosynthesis. The first pathway referred to as TPS/TPP or OtsA/OtsB uses UDP-glucose to make trehalose (Figure 1-11). Trehalose phosphate synthase (TPS, OtsA) initially transfers glucose from UDP-glucose to glucose 6-phosphate to make trehalose phosphate and releases UDP and then trehalose phosphate phosphatase (TPP, OtsB) removes the phosphate group to form trehalose and an inorganic phosphate (Elbain et al., 2003). Unlike to glycosyltransferases from many organisms that are very specific for the donor nucleotide-glucose, TPS from *M. smegmatis* and *M. tuberculosis* can utilize any type of ADP-glucose, UDP-glucose, TDP-glucose, CDP-glucose or GDP-glucose to make trehalose, suggesting the importance of trehalose biosynthesis for these organisms (Elbain et al., 2003). The second pathway also uses two enzymes to convert glycogen to trehalose. The first enzyme, maltooligosyl trehalose synthase (TreY), changes the  $\alpha$ - 1,4 linkage at the reducing end of glycogen to  $\alpha$ ,  $\alpha$ , 1, 1 linkage and the second enzyme, maltooligosyl trehalose trehalohydrolase (TreZ), cleaves the trehalose from glycogen chain (Elbain et

al., 2003 and Pan *et al.*, 2008). Breaking glycogen branches is mediated by debranching enzyme TreX (also known as GlgX).



**Figure 1-11.** The most widely demonstrated pathway responsible for trehalose biosynthesis in most organisms. During this pathway trehalose phosphate synthase (TPS or OtsA) transfers glucose from UDP-glucose to glucose-6-phosphate to form trehalose-6-phosphate and then the phosphate group is removed by trehalose phosphate phosphatase (TPP or OtsB) to make free trehalose (generated with ChemBioDraw).

Methylglucose lipo-polysaccharide is produced in a two-enzymatic pathway known as Rv3032. The first enzyme,  $\alpha$ -1,4 glucosyltransferase (Rv3032), turns over UDP-glucose or ADP-glucose to linear glucan and the second enzyme, Rv3031, introduces branches to the linear glucan. Methylglucose lipo-polysaccharides are composed of about 20 glucose residues attached by  $\alpha$ -1,4 links and 10% branching at positions 6. Although these molecules were discovered several decades ago, their function is not still clear but recent studies suggest that they are likely involved in fatty acid biosynthesis regulations within the cells (Chandra, Chater and Bornemann, 2011).

The fact that disruption of Rv3032 pathway has no effects on capsular glucan levels convinced Chandra and his colleagues to conclude that the newly discovered GlgE pathway should be responsible for capsular  $\alpha$ -glucan biosynthesis in mycobacteria (Chandra, Chater and Bornemann, 2011). However, Sambou and his group's efforts in studying the role of GlgC and GlgA from first pathway on capsular glucan production showed a complicated link between these two pathways. They observed that while in *glgC* knockout mutants both capsular glucan and glycogen production levels are reduced, in  $\Delta$ *glgA* mutants only capsular glucan production is affected. Moreover, latter mutants show lower persistence in mice cells, indicating the presence of a functional link between GlgA, capsular glucan production and immune system responses (Sambou et al., 2008).

The newly discovered pathway, GlgE pathway, involves a novel enzyme, maltosyltransferase GlgE, which transfers maltose from maltose 1-phosphate (M1P) to

maltooligosaccharides and genetically validated as a novel anti-tuberculosis target (Kalscheuer et al., 2010). The pathway also consists of three other enzymes as follows: trehalose synthase (TreS), maltokinase (Pep2) and GlgB. TreS catalyzes the reversible conversion of trehalose to maltose (MTase activity) and Pep2 converts the latter disaccharide into maltose 1-phosphate.

## **1.7 GlgE pathway enzymes**

### **1.7.1 Trehalose synthase**

By sequence similarity, trehalose synthase (TreS) is a member of glycoside hydrolase family GH13 that catalyzes the reversible interconversion of maltose and trehalose (Caner et al., 2013). However, it has been shown that TreS does not contribute significantly in the de novo production of trehalose in mycobacteria and mainly works in the direction of  $\alpha$ -maltose production, except in organisms that have access to the large amounts of cytosolic maltose from intracellular or extracellular sources (Miah, et al. 2013). Compare to other members of the GH13 family, which are monomeric, the crystal structure of TreS shows that this protein is a tetramer (Figure 1-13A) (Caner et al., 2013 & Roy et al., 2013). The members of this family have two conserved domains: a TIM barrel fold and a C-terminal  $\beta$ -sandwich domain (Figure 1-13B). TIM barrel, also known as triosephosphate isomerase, is composed of 8  $\beta$ -sheets surrounded by 8  $\alpha$ -helices. The TIM barrel core is decorated with a few loops that play different functional roles. For instance, the  $\beta$ 3-  $\beta$ 4 loop forms a calcium binding site,  $\beta$ 7-  $\beta$ 8 loop folds over the central  $\beta$ -barrel in chain A but is disordered in chain B and  $\beta$ 6-  $\beta$ 7 loop encompasses a helix-turn-helix motif (Figure 1-13B). The C-terminal domain consists of seven-stranded anti-

parallel  $\beta$ -sheets (Figure 1-13C). The function of this domain is not very clear but other members of the GH13 family harbor a carbohydrate-binding domain in this part of the enzyme (Caner et al., 2013).

Glycoside hydrolases are a group of enzymes that catalyze the hydrolysis of glycosidic bond between two or more carbohydrates using two different mechanisms, retaining or inverting (McCarter and Withers, 1994). Inverting glycoside hydrolases utilize a one step, single displacement mechanism, to change the anomeric configuration of substrate. During this process the general acid first attacks the hydrogen of water and in turn the oxygen, which now has a negative charge, attacks the anomeric carbon (Figure 1-12A). General acid aids this process by providing hydrogen during the transition state for glycosidic oxygen. On the other hand, retaining mechanism is a two step, double displacement process, in which the nucleophile amino acid attacks the anomeric center and forms a covalent glycosyl-enzyme intermediate (Figure 1-12B). At the same time, the second amino acid participates as an acid and protonates the glycosidic oxygen. In the second step, water attacks the glycosyl-enzyme intermediate and the second residue deprotonates the water molecule by acting this time as a base. As it will be explained later TreS catalyzes the interconversion of trehalose and maltose via a double displacement process (retaining) (Figure 1-15).

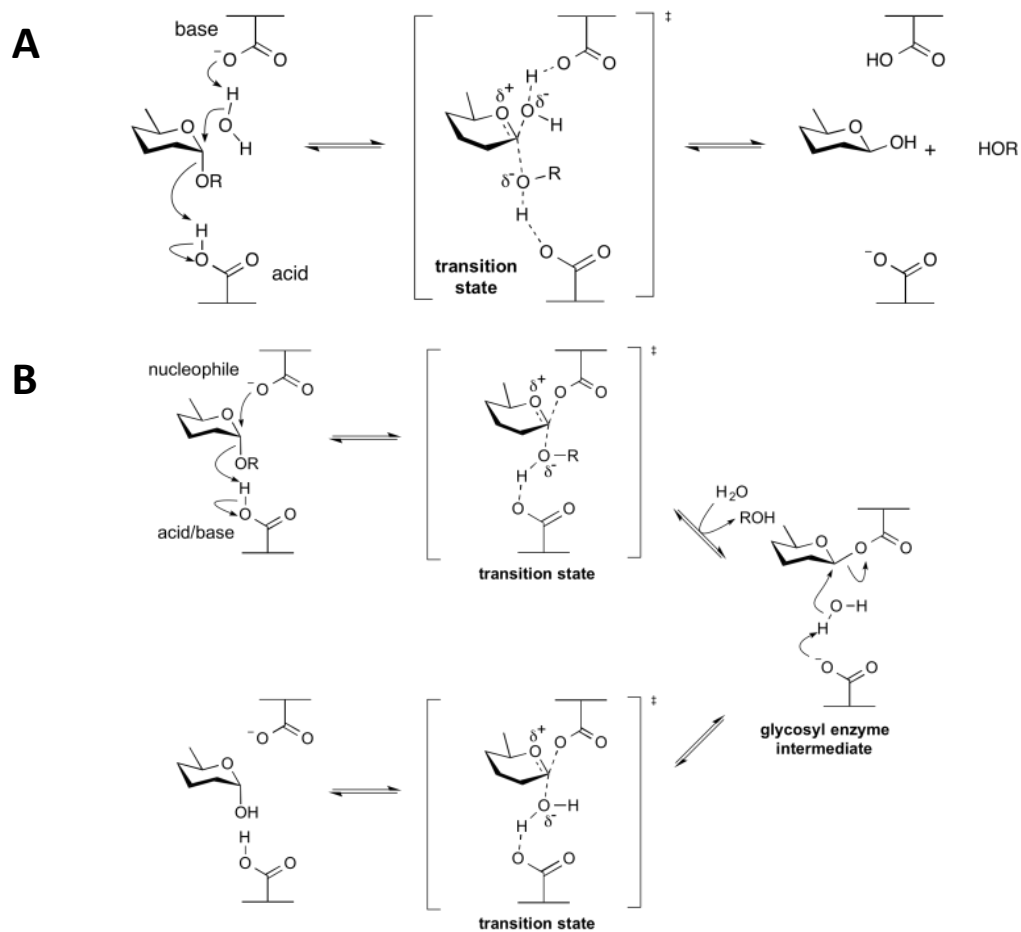
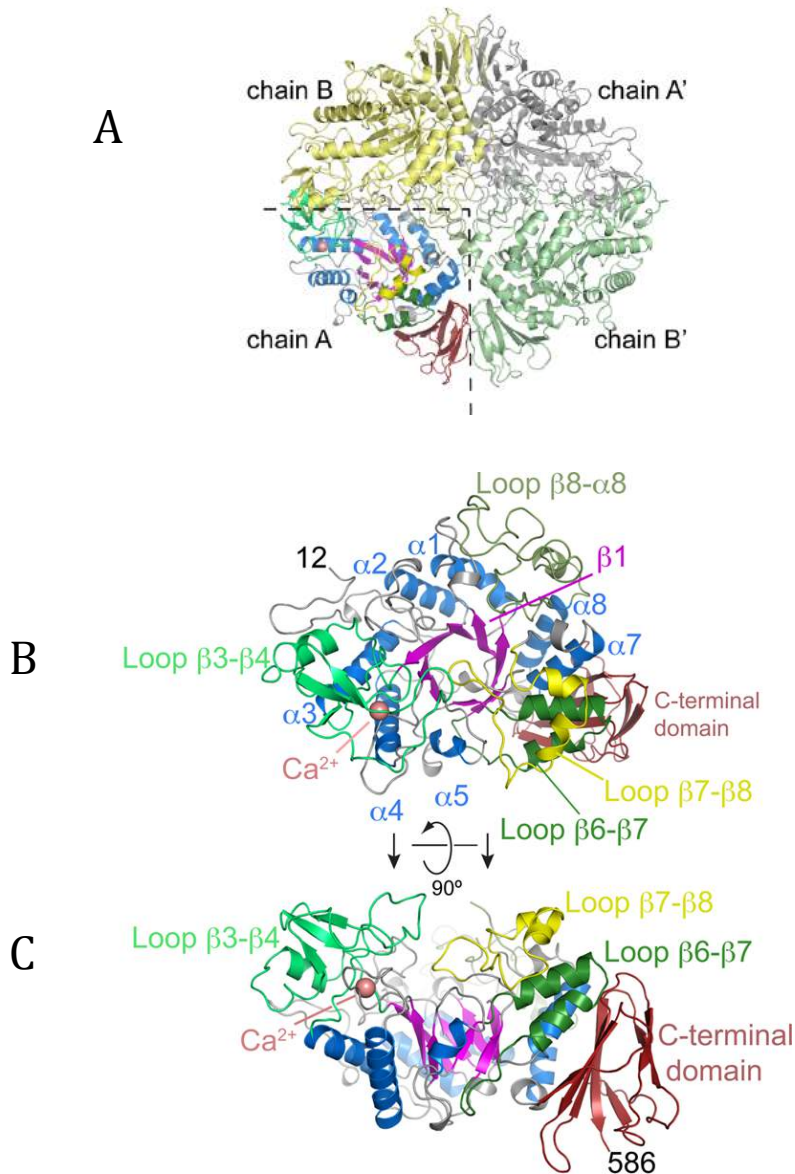


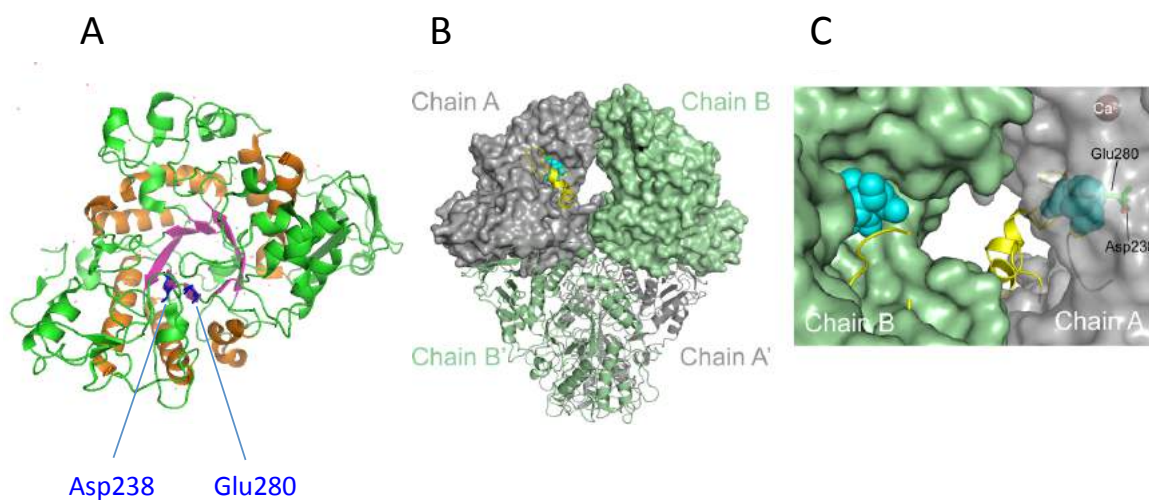
Figure 1-12. Schematic mechanisms of glycoside hydrolases. (A) Inverting and (B) retaining mechanisms. Adapted from (McCarter and Withers, 1994).



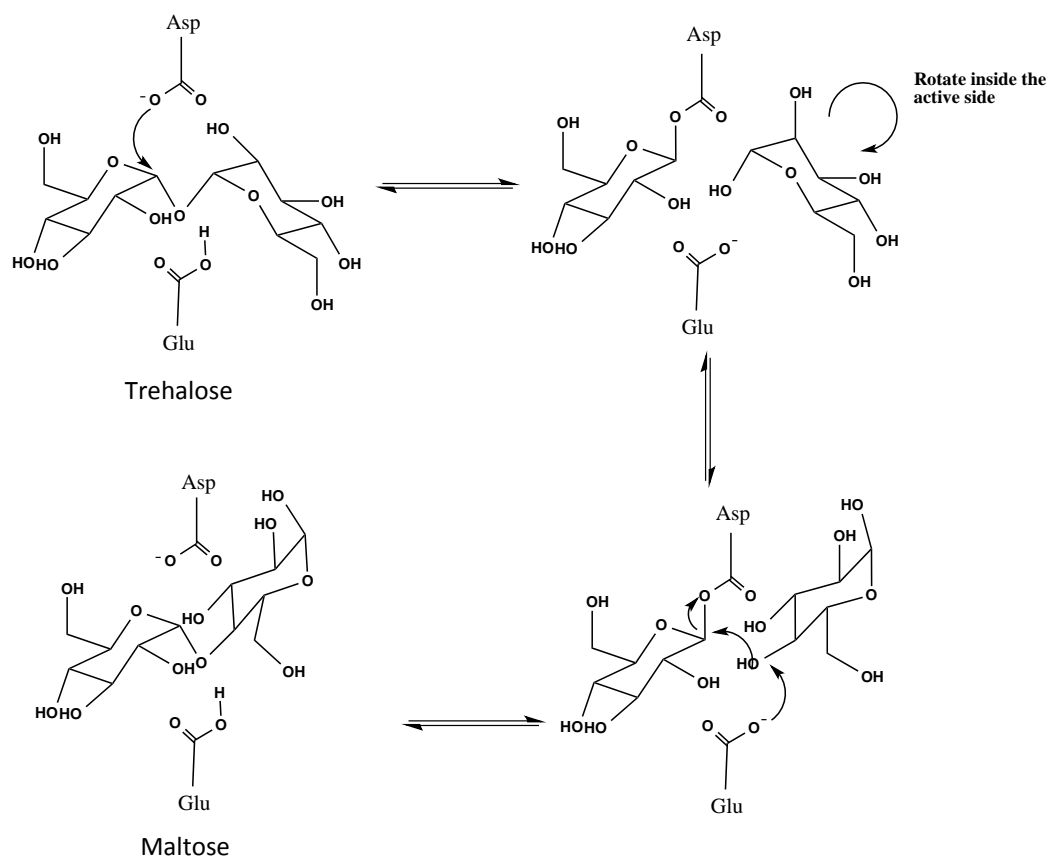
The active site of TreS is situated at the top (C-terminal) rim of the  $\beta$ -barrel, composed of Asp<sup>238</sup> (nucleophile), Glu<sup>280</sup> (general acid) and Asp<sup>350</sup> (unknown function) (Figure 1-13). The structure of *Mtb*TreS catalytic site is consistent with the catalytic mechanism of *Msm*TreS proposed by Zhang and colleagues (2011) (Figure 1-15). Asp<sup>238</sup> attacks the glycosidic bond and forms a covalent- $\beta$ -glucosyl-enzyme intermediate. The released glucose remains within the active site and reorients in such a way that 1-hydroxyl group or the 4-hydroxyl group is ready for re-attack in a productive complex. Although glucosyl-enzyme intermediate is protected in an enclosed site, it goes under some transfer to water and about 10% hydrolysis occurs. In *Pseudomonas mesoacidophila* two aromatic residues, Phe<sup>256</sup> and Phe<sup>280</sup>, form a clamp at the entrance of the enzymatic active site (Zhang et al., 2011). Mutation of any of these clamping phenylalanines increases the enzymatic hydrolysis and double mutant leads to exclusive hydrolysis. In *Mtb*TreS,  $\beta$ 7- $\beta$ 8 loop acts as a clamp to obstruct the diffusion of glucose moieties (Roy et al., 2013).



**Figure 1-13.** The crystal structure of trehalose synthase (TreS) from *M. tuberculosis*. (A) The crystal structure of TreS with two 2-fold rotational symmetry axes, chain A maps to chain A' and B into B' and subsequently chain A symmetrically superimposes to chain B. (B) Illustrating the TIM barrel fold as the conserved core of the TreS structure. (C) Representing an antiparallel  $\beta$ -sandwich domain at the C-terminus. This domain mediates intersubunit contacts. Adopted from Roy et al., 2013, with permission.



**Figure 1-14.** Position of active site and topology of solvent channel formed between chains A and B of *MtbTreS*. (A) Illustrates the position of active site of TreS at the top rim of  $\beta$ -barrel. Asp<sup>238</sup> and Glu<sup>280</sup> constitute the catalytic residues of active site. (B) Molecular surfaces of chains A and B in the TreS tetramer.  $\beta$ 7-  $\beta$ 8 loop is shown as a yellow ribbon. (C) Close up view of the solvent channel formed between active sites of chains A and B. The chain A is shown transparent so it can illustrate the catalytic residues of active site and putative substrate. Adapted from Roy et al., 2013. Panel A of this figure was generated with MacPyMOL software, with permission.



**Figure 1-15.** The catalytic mechanism of TreS proposed by Miah et al., 2013 (generated with ChemBioDraw).

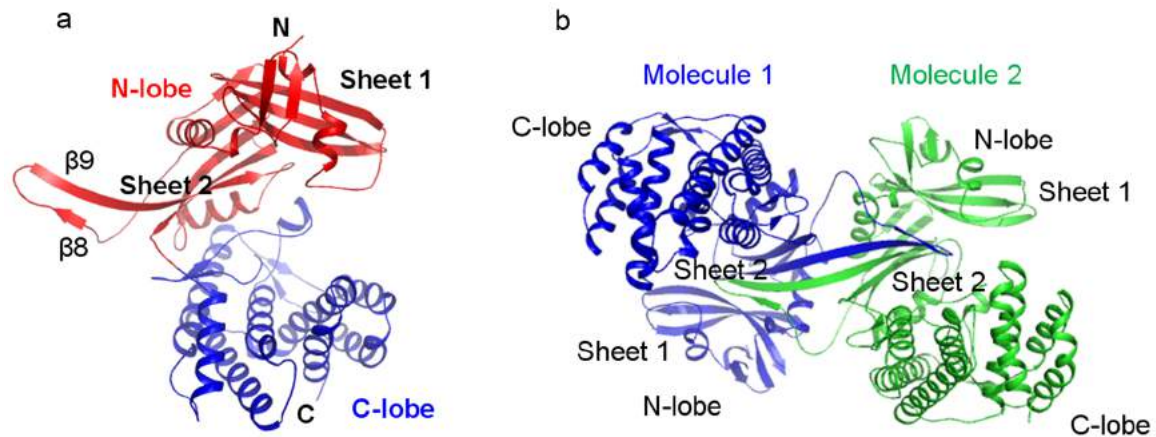
Pan and colleagues (2008) suggested that TreS has also  $\alpha$ -amylase activity, in addition to maltose to trehalose interconversion activity (referred to as MTase), and can convert glycogen to maltose and subsequently to trehalose, suggesting this enzyme is providing another pathway to convert glycogen to trehalose. However, solving the *Msm*TreS structure co-crystallized by acarbose, a potent  $\alpha$ -glucosidase inhibitor, has casted doubt on this model (Caner, et al. 2013). Although, a binding site was identified well remote from the active site, it didn't contain any catalytic residues required for amylase activity. Even in case that TreS has minor amylase activity, it is not physiologically significant, since it has been shown that wild type *M. smegmatis* accumulates large amounts of glycogen when grown in high concentrations of trehalose but *treS* mutants do not accumulate glycogen, supporting the dominant anabolic role of TreS (Kalscheuer et al., 2010).

According to PISA ([http://www.ebi.ac.uk/msd-srv/prot\\_int/pistart.html](http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html)) molecular assembly analysis data, it is suggested that TreS forms a tetramer in the solution. Roy and his colleagues using analytical ultracentrifugation in sedimentation equilibrium mode and solution X-ray scattering, confirmed the self-assembly of *Mtb*TreS into a tetrameric state, consistent with the crystal structure (Roy et al., 2013 & Caner et al., 2013).

### 1.7.2 Maltokinase Pep2

The crystal structure of Pep2 from *M. tuberculosis* (Li et al., 2014) and *M. vanbaalenii* (Fraga et al., 2015) was published recently. Li and colleagues described the structure of *MtbPep2* as two lobes architecture, with the N-terminal lobe mainly made up of  $\beta$ -strands and the C-terminal lobe mainly composed of  $\alpha$ -helices (Figure 1-16 A). The N-terminal lobe contains two sheets. The sheet close to the N-terminus of protein embraces five anti-parallel  $\beta$ -strands ( $\beta 1$ -  $\beta 5$ ) while the second sheet is formed of seven anti-parallel strands ( $\beta 6$ -  $\beta 7$ ,  $\beta 9$ -  $\beta 11$ ,  $\beta 8'$ -  $\beta 9'$ ). The strands  $\beta 8$  and  $\beta 9$  from second sheet protrude out of the protein, enter into the adjacent monomer and form an unusual homotypic dimer (Figure 1-15 B).

Structural alignment of *MtbPep2* with homologous kinases, using DALI (Holm and Rosenstrom, 2010), revealed that some parts of N-terminal lobe ( $\beta 1$ -  $\beta 3$  strands) are unique to *M. tuberculosis* and are missing from other homologue structures (Li et al., 2014). In contrast, C-lobe architecture is similar to the C-lobes of most other kinases (Li et al., 2014). However, they could not show the function of this unique  $\beta$  sheet in the structure of *MtbPep2*. Moreover, unlike other kinases where C-lobe is involved in dimerization, in *MtbPep2* only N-lobe forms a dimer with adjacent molecule (Figure 1-16 B).

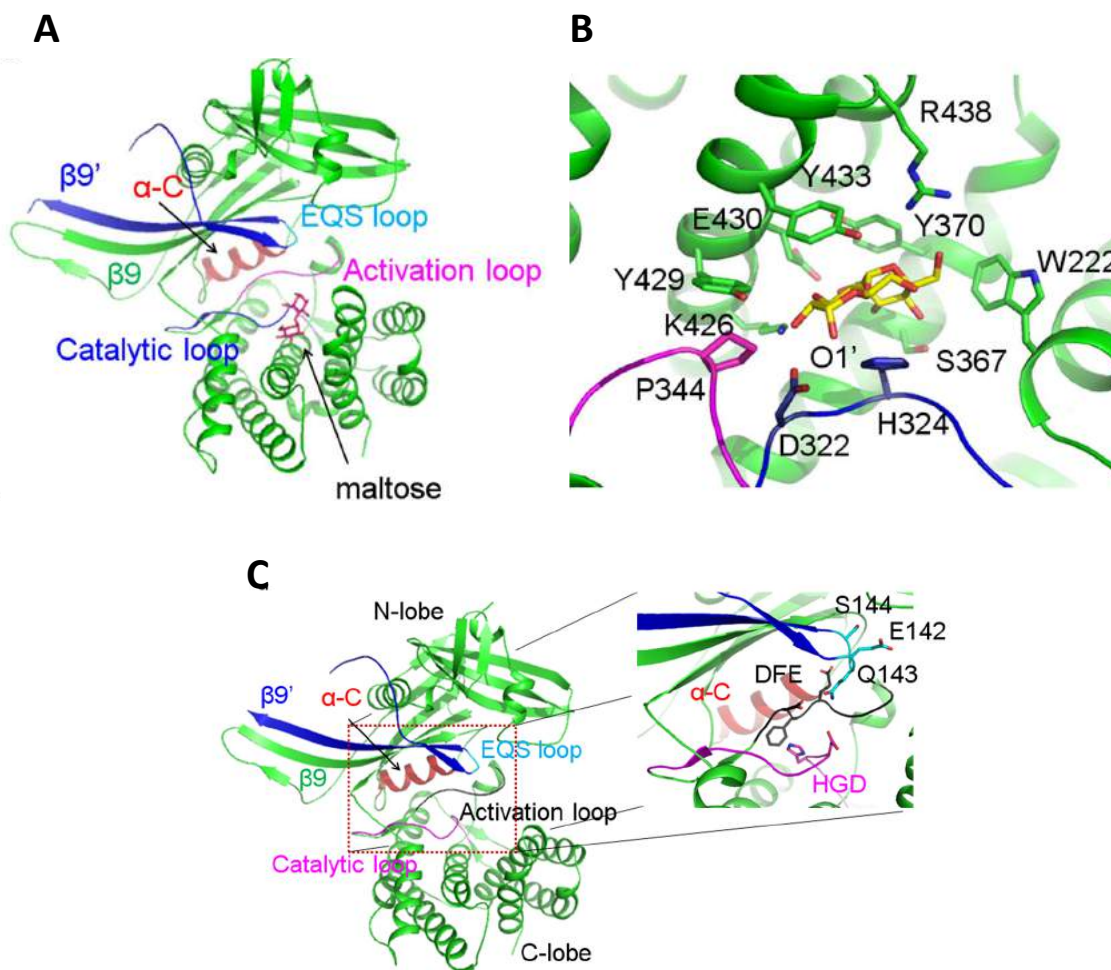


**Figure 1-16.** The crystal structure of *MtbPep2*. (a) Cartoon representation of Pep2 in monomeric state with two distinct lobes, N and C-lobes. The N-lobe, depicted in red, is mainly composed of  $\beta$ -sheets while C-lobe, depicted in blue, is primarily made up of  $\alpha$ -helices. (B) Cartoon representation of homotypic dimer of Pep2. The strands  $\beta 8$  and  $\beta 9$  from each monomer protrude into the adjacent monomer and form a dimer. Adopted from (Li et al., 2014), with permission.

Solving the crystal structure of *MtbPep2* in complex with maltose demonstrated that 11 residues, W222, D322, H324, P344, S367, Y370, K426, Y429, E430, Y433 and R438, interact with maltose (Li et al., 2014) (Figure 1-17 A-B). Alanine mutagenesis of these amino acids revealed that 9 residues are essential for phosphorylation of maltose (except P344 and K426). In addition, in order to identify the ATP-binding site of *MtbPep2*, the nucleotide-binding site of amino glycoside phosphotransferase (APH; PDB code 1J7U, RMSD of 2.6 Å for 90 aligned Ca positions) was compared with *MtbPep2*, revealing that three amino acids E142, Q143 and S144 from the EQS loop are interacting with ATP (Figure 1-17 C).

Before solving the structure of Pep2 by Li and colleagues, Mendes and colleagues characterized Pep2 from *M. bovis* BCG, an homologue species of *M. tuberculosis* (Mendes, et al. 2010). They showed that Pep2 is dependent on maltose and ATP, although GTP and UTP can also be used to produce MIP. They determined its  $K_M$  as  $2.52 \pm 0.40$  mM for maltose and  $0.74 \pm 0.12$  mM for ATP. Divalent cations are essential for enzyme activity and  $Mg^{2+}$  is the best activator. Pep2 has its maximal activity at 60 °C, between pH 7 and 9 (at 37 °C) and is unstable on ice and upon freeze/thawing (Mendes, et al. 2010).





**Figure 1-17.** Maltose and nucleotide binding sites of *MtbPep2*. (A) Maltose binding site in C-terminal lobe and in close contact with the conserved motifs essential for catalysis. (B) An illustration of maltose in contact with catalytic residues. Catalytic residues and P344 from the activation loop are shown in blue and magenta, respectively. Maltose is shown in sticks in both panels. (C) Cartoon representation of nucleotide binding pocket at the junction of N- and C-lobes. The activation loop (shown in black),  $\alpha$ -C helix (red color) and catalytic loop (magenta color) are highly conserved between kinases. The EQS loop (cyan sticks) contains the amino acids necessary for ATP binding and catalysis. Adopted from Li et al., 2014, with permission.

### 1.7.3 Maltosyltransferase GlgE

GlgE has become the focus of attention as a potential new class of drug targets as it is a part of pathway, which has never been targeted by antimicrobials and especially doesn't have any homologues in humans (Kalscheuer et al., 2010). Studying GlgE (encoded by Rv1327c) represents two novel bactericidal mechanisms, which can lead to death in *M. tuberculosis*. In first mechanism, inactivation of *glgE* leads to self-poisoning and death due to accumulation of M1P. Hyperaccumulation of M1P triggers the DNA damage responses, indicating DNA damage is likely the main factor causing death in Mtb. It is also possible that M1P is indirectly involved in toxicity through some complex mechanisms such as unidentified degradation or side products. The second mechanism is based on synthetic lethality with the glucosyltransferase Rv3032. Synthetic lethal interaction occurs when disrupting two non-essential genes can lead to death, while inhibiting neither of these genes on its own is lethal. This shows that these two pathways generate similar products and disturbing one pathway can be compensated by the activity of other pathway (Kalscheuer et al., 2010).

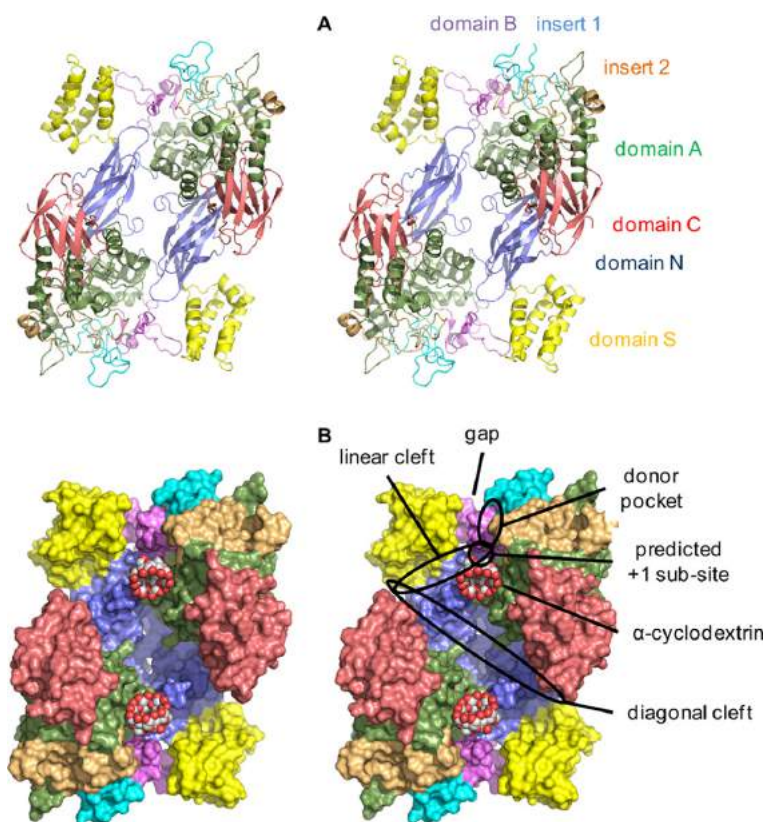
Kalscheuer and colleagues in a detailed study in 2010 showed there is a functional link between TreS, Pep2, GlgE and GlgB (Kalscheuer et al., 2010). To achieve this, they deleted *glgE* in *M. smegmatis* and subjected the mutant to the different concentrations of trehalose. They found *MsmΔglgE* mutant is sensitive to the trehalose. Although this sugar is abundant in mycobacteria, even small amount of trehalose, 0.5 mM, can completely inhibit the growth of the mutant. This growth inhibition in *MsmΔglgE* correlates to the M1P hyperaccumulation. Deletion of either trehalose synthase gene (*tres*) or *pep2* gene

confers trehalose resistance and suppresses M1P accumulation (Kalscheuer et al., 2010). The *glgE* gene clusters with *glgB* gene (encoded by Rv1326c), an essential gene in Mtb H37Rv which introduces  $\alpha$ -1,6-linked branches into linear  $\alpha$ -1,4-glucans. Deletion of *glgB* gene in *M. smegmatis* also, confers sensitivity to trehalose. This sensitivity in *Msm* $\Delta$ *glgB* can be suppressed by deletion of *treS* (Kalscheuer et al., 2010).

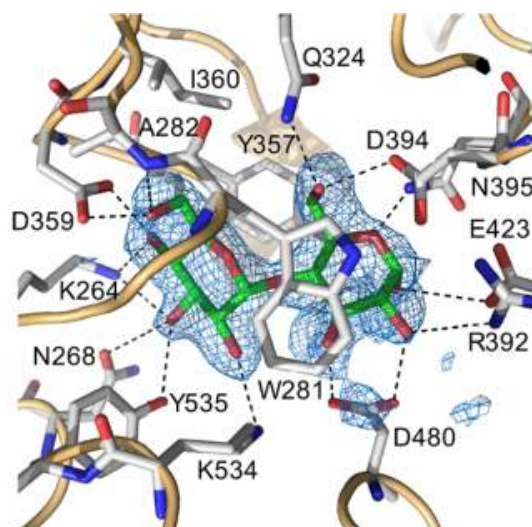
Trehalose synthase (TreS) and Pep2 catalyze the interconversion of trehalose and maltose and phosphorylation of maltose to M1P using ATP, respectively. Exploring the genomes of different organisms indicates that these two genes are clustered together in many cases, suggesting they catalyze two consecutive reactions, trehalose is first converted to maltose and latter is then turns over to M1P. In order to support this, Kalscheuer and his colleagues monitored the metabolism of  $^{14}$ C-labeled trehalose in *Msm* $\Delta$ *glgE* mutant. Radiolabeled trehalose was immediately consumed and converted to M1P (Kalscheuer et al., 2010). On the other hand, they showed that perturbing *treS* or *pep2* hinders M1P production in the  $\Delta$ *glgE* mutant. Moreover, deletion of *pep2* leads to accumulation of large amounts of radiolabeled maltose, implying the role of Pep2 in rapid conversion of maltose to M1P. These findings together represent the direct connection between TreS, Pep2, GlgE and GlgB in *M. smegmatis* (Kalscheuer et al., 2010).

## Structure of *Streptomyces* GlgE

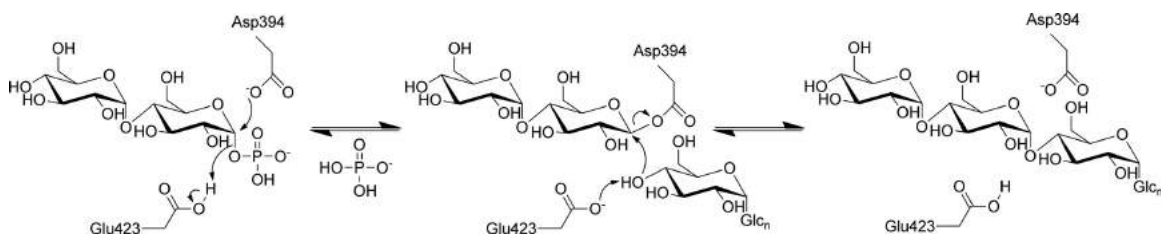
The crystal structure of *Streptomyces coelicolor* GlgE has shown that this protein is a homodimer composed of five domains, A, B, C, N and S (Figure 1-18) (Syson et al., 2011). Domain A is a  $(\beta/\alpha)_8$  barrel and acts as the catalytic domain. The role of domain B, involves a pair of anti-parallel strands and one short helix, is not clear but in other members of GH13 family is responsible for  $\text{Ca}^{2+}$  ion binding. The C-terminal domain C possesses a  $\beta$ -sandwich fold and contributes in stabilizing domain A and is involved in substrate binding. The N-terminal domain N, which also contains a  $\beta$ -sandwich, forms the core of the dimer interface. The final domain, referred to as domain S, is unique to GlgE and participates in dimer interface by interacting with domain B of the next subunit. M1P enters into the donor pocket, situated in the center of domain A, in such a way that its reducing end sits between Asp<sup>394</sup> and Glu<sup>423</sup> as nucleophile/base and proton donor, respectively (Figure 1-19). Asp<sup>394</sup> attacks the reducing end of M1P and forms a maltosyl-enzyme intermediate, and phosphate group is released as an inorganic phosphate (Figure 1-20). The non-reducing end of acceptor is deprotonated and attached to the C-1 position of maltose. A pair of anti-parallel  $\beta$ -strands of domain B and Pro<sup>353</sup> and Pro<sup>357</sup> form a lid at the entrance of donor pocket and control the access to this pocket (Syson et al., 2011).



**Figure 1-18.** The crystal structure of *Streptomyces coelicolor* GlgE. (A) Shows the protein as a homodimer with 5 domains. (B) A space-filling representation showing different features of the protein. Adopted from Syson et al., 2011, with permission.



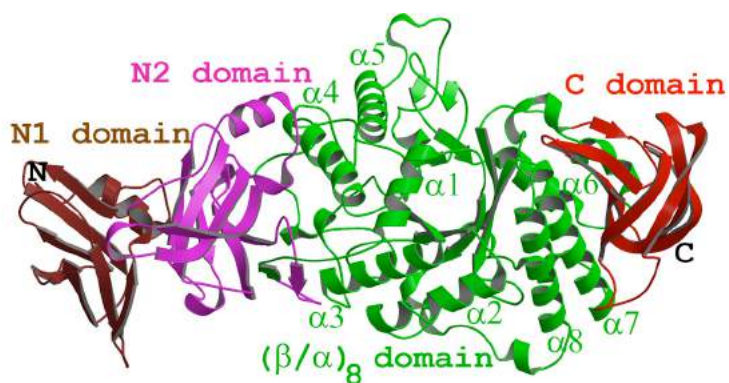
**Figure 1-19.** Co-crystallization of maltose and GlgE. Adopted from Syson et al., 2011, with permission.



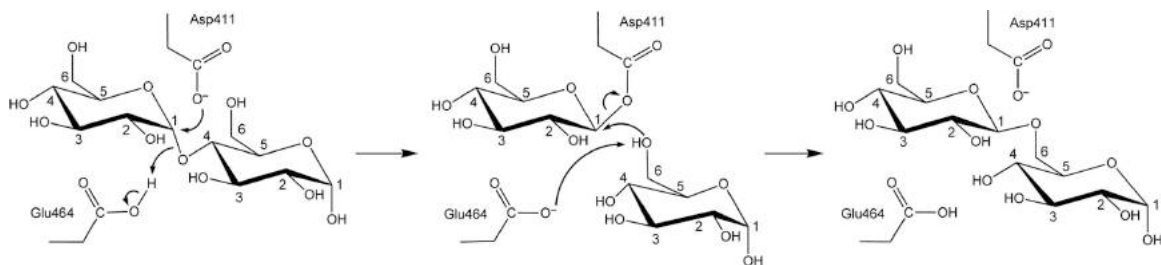
**Figure 1-20.** Double displacement mechanism of GlgE proposed by Syson et al. 2011, with permission.

#### 1.7.4 Branching enzyme GlgB

GlgB has also shown to be essential for *Mtb* viability through having a secondary effect on GlgE (Kalscheuer et al., 2010). In the absence of branching enzyme GlgB, the produced linear glucan has fewer non-reducing ends and become insoluble by reaching the degree of polymerization (DP) of about 20. In this case GlgE will not be able to extend the glucan, and reducing the GlgE activity similarly leads to toxic accumulation of M1P (Kalscheuer et al., 2010). *MtbGlgB* is a member of glycoside hydrolase family GH13 composed of four main domains: N1, N2, central  $(\beta/\alpha)_8$  domain and C domain. N1, N2 and C domains have a similar structure consist of a  $\beta$ -sandwich and central domain with a  $(\beta/\alpha)_8$  structure harbors the catalytic site (Pal et al., 2010) (Figure 1-21). Based on the structure of active site a substrate-binding model has been proposed (Figure 1-21). Asp<sup>411</sup> attacks the 1-4 glycosidic bond and forms a glucosyl-enzyme intermediate and Glu<sup>464</sup> protonates the 4<sup>th</sup> glycosidic oxygen. The substrate fragment is then transferred to the hydroxyl group at C6 of a glucose unit on the same or another chain to form 1-6 bond (Figure 1-22).



**Figure 1-21.** The crystal structure of *M. tuberculosis* GlgB. The N1  $\beta$ -sandwich domain (residues 1-105), the N2  $\beta$ -sandwich domain (residues 106-226), the central  $(\beta/\alpha)_8$  domain (residues 227-630) and the C-terminal  $\beta$ -sandwich (residues 632-731) are shown in brown, magenta, green and red, respectively. Adopted from Pal et al., 2010, with permission.

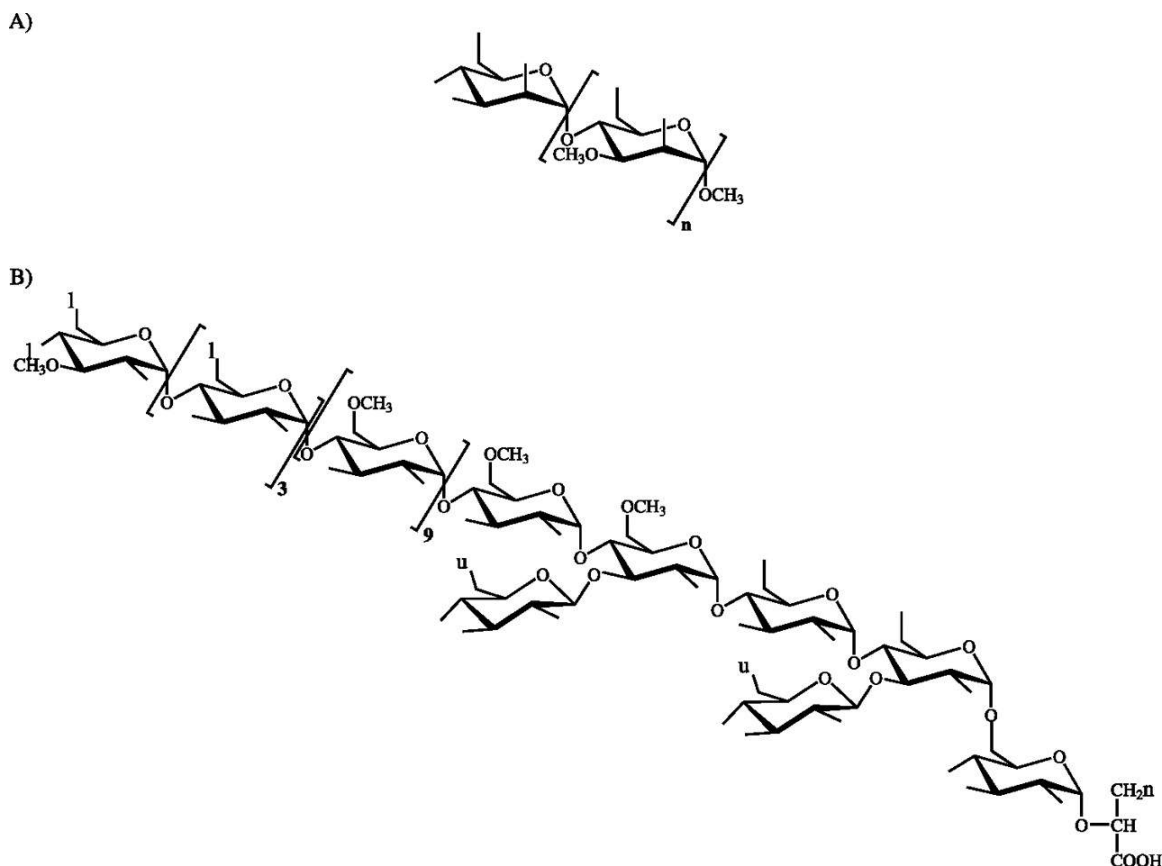


**Figure 1-22.** The substrate-binding model for *Mycobacterium tuberculosis* GlgB proposed by Pal et al., 2010, with permission.



## 1.8 Rv3032 pathway

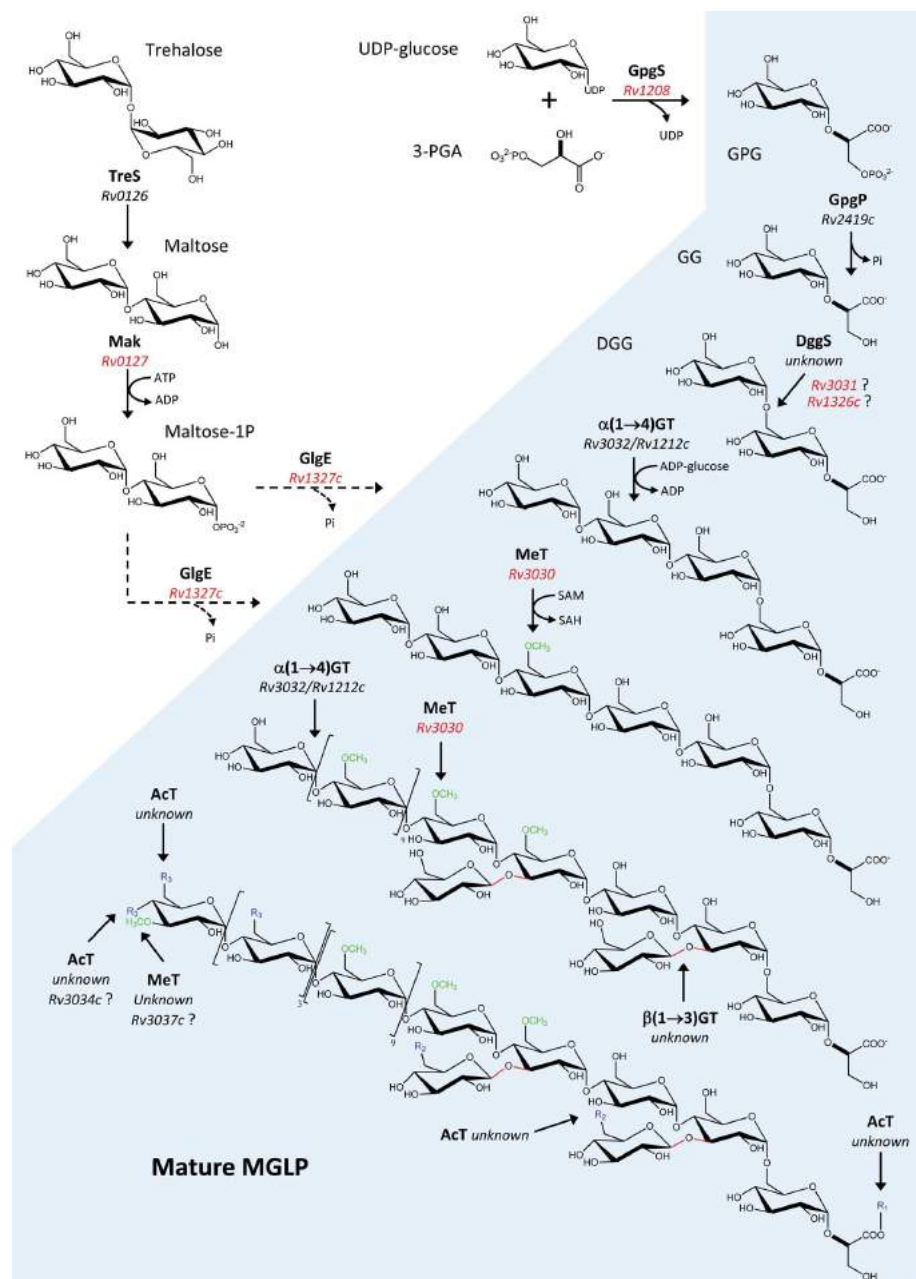
Mycobacteria produce two classes of cytoplasmic polymethylated polysaccharides (PMPS) including 6-O-methylglucosyl containing lipopolysaccharides (MGLP) and 3-O-methylmannose polysaccharides (MMP) (Stadthagen et al., 2007) (Figure 1-23). The role of these molecules is not very clear but they are proposed to be involved in regulating fatty acid biosynthesis and protecting latter molecules from degradation (Stadthagen et al., 2007). PMPS might also act as general lipid carriers to facilitate the synthesis of very large and insoluble mycolic acid esters (Stadthagen et al., 2007). In order to identify the genes encoding PMPS in mycobacteria, Stadthagen and colleagues searched for putative glucosyltransferases that use ADP-D-Glc or UDP-D-Glc as substrates in Carbohydrate-Active enZymes database (CAZy). Out of seven potential NDP-sugar-dependent glucosyltransferases found, they chose Rv3032 and Rv1212c for further experimentation due to greatest level of sequence similarity with *E. coli* GlgA. Scrutinizing the genes in close vicinity to Rv3032 on the *M. tuberculosis* genomic DNA revealed two putative S-adenosyl-methionine-dependent methyltransferase genes (Rv3030 and Rv3037c) and a putative branching enzyme gene (Rv3031).



**Figure 1-23.** Structures of mycobacterial polymethylated polysaccharides (PMPS). (A) Structure of methylmannose polysaccharides (MMP) composed of 10-13  $\alpha$ -(1-4) linked 3-O-methylmannoses terminated at the non-reducing end by a single  $\alpha$ -(1-4) linked unmethylated D-mannose and an  $\alpha$ -methyl group at reducing end. (B) Structure of methylglucosyl lipopolysaccharides (MGLP) composed of 10  $\alpha$ -(1-4)-linked 6-O-methylglucosyl residues with a non-reducing end made of tetrasaccharide glucosyl residues methylated at position 3. Reducing end consists of 4  $\alpha$ -(1-4) and one  $\alpha$ -(1-6) linked glucose residues linked to the position 2 of D-glyceric acid. Position 3 of the second and fourth  $\alpha$ -D-glucosyl residues are substituted by single  $\beta$ -D-glucosyl residues. Adopted from Stadthagen et al., 2007, with permission.

Inhibiting Rv3032 in *M. tuberculosis* reduced lipopolysaccharides and glycogen contents by ~50%, indicating the involvement of this gene in MGLP synthesis. It is noteworthy to mention that Stadthagen and colleagues were able to detect only MGLP in the extracts or in the purified PMPS fractions of *M. tuberculosis*, but not MMP. Rv3032 deletion, however, did not affect capsular glucan synthesis in *M. tuberculosis*. Likewise, knocking out the orthologous of Rv3030 in *M. smegmatis* (MSMEG2349) significantly reduced the production of PMPS in *M. smegmatis*, confirming the participation of this gene in the biosynthesis of MGLP. Furthermore, based on sequence similarity to GH-57 family branching enzyme from *Thermococcus kodakaraensis*, Stadthagen suggested that Rv3031 should be involved in generating  $\alpha$ -(1-6) glycosidic bonds in MGLP molecules. Rv3032 and Rv3030 are proposed to catalyze the elongation and 6-O-methylation of MGLP respectively in *M. tuberculosis* (Mendes et al., 2011) (Figure 1-24).

In another attempt to identify other genes involved in MGLP synthesis, Kaur and colleagues identified Rv1208 (GpgS) in the vicinity of Rv1212c as a glucosyl-3-phosphoglycerate catalyzing the transfer of the first glucosyl residue of MGLPs (Figure 1-24) (Kaur et al., 2009). The crystal structure of GpgS, which is essential for the survival of *M. tuberculosis*, was solved as apo enzyme and in complex with both UDP and 3-phosphoglycerate by Jose Barbosa Pereira et al. (2008). Mendes showed that Rv2419c (encoding a glucosyl-3-phosphoglycerate phosphatase) is responsible for converting glucosyl-3-phosphoglycerate to glucosylglycerate (Mendes et al., 2011) (Figure 1-24).



**Figure 1-24.** Proposed pathway for the biosynthesis of MGLP in *M. tuberculosis*. Genes highlighted in red are essential genes for *M. tuberculosis* growth and dashed arrows represent putative biosynthetic steps. GpgS: glucosyl-3-phosphoglycerate synthase, GpgP: glucosyl-3-phosphoglycerate phosphatase, DggS: diglucosylglycerate synthase, GT: glucosyltransferases, MeT: methyltransferase, AcT: acyltransferase, 3-PGA: 3-phosphoglyceric acid, GPG: glucosyl-3-phosphoglycerate, GG: glucosylglycerate, DGG: diglucosylglycerate. Adopted from Mendes et al., 2010, with permission.

A comparative genomic study indicated that unlike GlgC-GlgA and GlgE pathways, which are relatively abundant between bacterial genomes, Rv3032 pathway is unique with a few exceptions only to mycobacteria (Chandra, Chater and Bornemann, 2011).

## **1.9 Aims and objectives**

GlgE has been genetically validated as a new anti-TB drug target; especially lack of homologous of this gene from humans make it even more attractive (Kalscheuer et al., 2010). The primary aim of this project was to solve the crystal structure of GlgE from Mtb using X-ray crystallography and exploiting the structural information of this protein to find its inhibitors. Moreover, given that the accumulation of MIP is toxic to bacilli, we hypothesized that mycobacterium might facilitate the turn over of this substrate into  $\alpha$ -glucan chain through forming a complex between Pep2 and GlgE. Probing this complex formation could provide us with new information about the metabolic flux in GlgE pathway.

The second aim of this project was to describe the crystal structure of Pep2 on its own or in complex with TreS. Prior to this study there was not any information on the crystal structure of Pep2 and modeling predictions of tertiary structure of this protein would strongly suggest that this protein has a novel structure. In particular, solving the crystal structure of Pep2 in complex with TreS could explain the Pep2 rate-enhancement effects following complex formation with TreS.

Finally, we were interested to explain the synthetic lethal interactions between GlgA and Rv3032 in Mtb. Synthetic lethal interactions have also been reported for a dozen of other genes and explaining this effect could provide us with new insights into the biochemistry of mycobacteria and subsequently new therapeutic agents. We thought that solving the crystal structure of both proteins might be a good start to better understand this effect.

## **2 Materials and methods**

## 2.1 Cloning of *Mtb* and *MsmgI**E* and *glgB*

The *M. tuberculosis* Rv1327c gene was cloned from a pET28a vector containing Rv1327c gene (provided by Dr. Rana Roy) using the Phusion High-Fidelity DNA Polymerase kit (New England Biolabs, UK) and forward and reverse primers (Table 2-1), followed by PCR amplification (Table 2-2). The PCR products were analyzed on a 1% agarose gel stained with SYBR Safe DNA gel stain (Invitrogen, USA). The PCR products were purified from gel using QIAquick Gel Extraction protocol using a microcentrifuge. *Nde*I and *Hind*III restriction enzymes were used to digest purified PCR products to generate overhang ends. The reaction was set up in a final volume of 75 µL as 50 µL PCR products, 7.5 µL 10x FastDigest Buffer, 3 µL of each *Nde*I and *Hind*III restriction endonucleases, for 3 hour at 37 °C. The reaction sample was loaded on a 1% agarose gel and purified using QIAquick Gel Extraction protocol using a microcentrifuge.

Overhang PCR products were ligated into pET28a vectors digested with *Nde*I and *Hind*III restriction enzymes as explained earlier. Ligation reactions were set up with 2 µL T4 DNA Lygase buffer, 1 µL T4 DNA Lygase (New England Biolabs, UK) and sufficient amount of insert and plasmid in a final volume of 20 µL at room temperature for 30 minutes. The required amount of insert to plasmid was calculated using [www.insilico.uni-duesseldorf.de/Lig\\_Input.html](http://www.insilico.uni-duesseldorf.de/Lig_Input.html) program in a molar ratio of 3:1. 2 µL of ligation reaction was transformed into the Top10 cells using heat-shock protocol and cultured on LB-Kan plates (25 µg/mL kanamycin) overnight at 37 °C. The colonies were screened using restriction enzyme digestion and DNA sequencing.



**Cloning of *MsmglgE* and *glgB* and *MtbglgB* using In-Fusion HD cloning kit.** The *M. smegmatis* MSMEG\_4916 gene, MSMEG\_4918 and Rv1326c were amplified from *M. smegmatis* and *M. tuberculosis* genomic DNAs using the Phusion High-Fidelity DNA Polymerase kit (New England Biolabs, UK) and forward and reverse primers (Table 2-1), followed by PCR amplification (Table 2-2). The primers were designed in such a way that would contain 15 nucleotides extensions homologous to pET28a ends. The PCR products were analyzed and purified as explained earlier, followed by setting up the ligation reaction as 1 µL pET28a vector digested with NdeI and HindIII, 1 µL purified PCR product, 2 µL 5X In-fusion HD enzyme premix (Clontech) and 6 µL water in a final volume of 10 µL at 50 °C for 15 minutes. 2.5 µL of each cloning reaction were transformed into Top10 cells using heat-shock protocol and cultured on LB-Kan plates (25 µg/mL kanamycin) overnight at 37 °C. The colonies were screened using gel electrophoresis and DNA sequencing.

**Truncating *MsmPep2*.** *M. smegmatis* MSMEG\_6514 gene was cloned from a pET28a vector containing MSMEG\_6514 gene using the Phusion High-Fidelity DNA Polymerase kit (New England Biolabs, UK) and seven forward and one reverse primers, as explained earlier (Table 2-1 and 2-2).

**Table 2-1.** Sequence of primers for cloning. Underlined and bold sequences represent extensions homologous to pET28a ends and restriction sites, respectively.

Gene	Primer	Sequence
<i>M. tuberculosis</i> <i>glgE</i> cloned in pET28a	Forward	GATCCATATGGTGAGTGGCCGGGCAATCGGAAC
	Reverse	GATCGATCAAGCTTTCACCTCCTGCGCAGCAGCGTGTTC
<i>M. smegmatis</i> <i>glgE</i> cloned in pET28a	Forward	<u>CGCGCGGCAGCCATA</u> GTGGCCGGTCGTATCGGA
	Reverse	<u>GTGCGGCCGCAAGCTT</u> CATTCCCTGCGTAGCAAG
<i>M. tuberculosis</i> <i>glgB</i> cloned in pET28a	Forward	<u>CGCGCGGCAGCCATA</u> ATGAGTCGATCCGAGAAAC
	Reverse	<u>GTGCGGCCGCAAGCT</u> CTAGGCGGGCGTCAGCCA
<i>M. smegmatis</i> <i>glgB</i> cloned in pET28a	Forward	<u>CGCGCGGCAGCCATA</u> ATGACGAGAAGCAGCAATC
	Reverse	<u>GTGCGGCCGCAAGCT</u> CTAAGCCGGCTCGAACCA
<i>M. tuberculosis</i> <i>treS</i> cloned in BACTH plasmids	Forward	GATCTCTAGAAATGAACGAGGCAGAACACAGCGTC
	Reverse	GATC <b>GGT</b> TACCTCATAGGCGCCGCTCTCCCCGCA
<i>M. tuberculosis</i> <i>pep2</i> cloned in BACTH plasmids	Forward	GATCTCTAGAAATGACTCGGTCTGGACACGCTGGCA
	Reverse	GATC <b>GGT</b> TACCTCAGCTAGCGGTCTAGGCGGGCGA
<i>M. tuberculosis</i> <i>glgE</i> in cloned in BACTH plasmids	Forward	GATCTCTAGAGTGAGTGGCCGGGCAATCGGAAC
	Reverse	GATC <b>GGT</b> TACCTCACCTCCTGCGCAGCAGCGTGTTC
<i>M. tuberculosis</i> <i>glgB</i> in cloned in BACTH plasmids	Forward	GATC <b>GGATCC</b> ATGAGTCGATCCGAGAAAC
	Reverse	GATC <b>GGT</b> TACCCTAGGCGGGCGTCAGCCA

Table 2-1 continued.

Gene	Primer	Sequence
<i>MsmPep2</i> Δ10	Forward	GATCCATATGCAGCAGCGCTGGTACGCCGGACG
<i>MsmPep2</i> Δ20	Forward	GATCCATATGGAATTGGTCTCGGCCACAACCG
<i>MsmPep2</i> Δ30	Forward	GATCCATATGGTGCGGCTGCGCGACGGGCTGG
<i>MsmPep2</i> Δ40	Forward	GATCCATATGCTGCTGCAGGCGAATTACGCCGACG
<i>MsmPep2</i> Δ50	Forward	GATCCATATGGATGAGCGTTACCAGGTGATCGTCG
<i>MsmPep2</i> Δ60	Forward	GATCCATATGGGAAGCGGACCGATCGACGAGTAC
<i>MsmPep2</i> Δ70	Forward	GATCCATATGGTCGCCACGATCGGCATCGCCGAC
<i>MsmPep2</i> Δ10-Δ70	Reverse	GATCAAGCTTTCACTCCAGAAGACGTGCGATGGAGCC CAG
<i>Rv3031</i>	Forward	GATCCATATGTTGAACACGTCCGCAAGCCCGGTG
	Reverse	GATCAAGCTTTCACTTGGGCAGCCTCCGAGCGTC
<i>Rv3032</i>	Forward	GATCCATATGATGAGGATCCTCATGGTGTCGTGGGA
	Reverse	GATCAAGCTTCTACCGATCGGGAAGAGCGTGCTCGAC
<i>Rv1212c</i> (GlgA)	Forward	GATCCATATG ATGCGGGTGGCGATGTTGACTCG
	Reverse	GATCAAGCTT TTACGCGCACACCTTCCGGTAG
<i>Rv1213</i> (GlgC)	Forward	GATCCATATG ATGAGAGAAGTGCCGCACGTGCTG
	Reverse	GATCAAGCTT CTAGATCCAAACACCCTTGCCAC

**Table 2-2.** Thermocycler settings for PCR amplifications.

	<i>M. tuberculosis</i> <i>glgE</i>	<i>M. smegmatis</i> <i>glgE</i>	<i>M. tuberculosis</i> and <i>M. smegmatis</i> <i>glgB</i>	<i>M. smegmatis</i> Pep2 $\Delta 10- \Delta 70$
Initial denaturation	98 °C/2 min	98 °C/2 min	98 °C/2 min	98 °C/2 min
Denaturation	98 °C/30 s	98 °C/30 s	98 °C/30 s	98 °C/30 s
Annealing	-	64 °C/30 s	72 °C/30 s	-
Extension	72 °C/1 min	72 °C/1 min	72 °C/90 s	72 °C/90 s
Final extension	72 °C/20 min	72 °C/20 min	72 °C/10 min	72 °C/10 min
Number of cycles	34	34	35	35

## 2.2 Testing the expression and solubility of recombinant proteins

The recombinant plasmid of protein of interest was transformed into different competent cells, BL21 (DE3), BL21 (DE3) pLysS, C41 (DE3) or Rosetta using heat-shock protocol. Recombinant plasmid of Rv3031 was co-transformed with 60.1 GroES and 60.2 GroES plasmids. 1  $\mu$ L of miniprep DNA sample was added to the 50  $\mu$ L of competent cells and incubated on ice for 20-30 minutes. Cells were heat-shocked by placing them in a 42 °C water bath for 45 seconds, followed by incubation on ice for 3 minutes. Cells were recovered by adding 350  $\mu$ L of room temperature S.O.C Medium and placing them in a shaking incubator at 200 rpm for 1 hour at 37 °C. 50  $\mu$ L of transformed cells were plated on LB agar plates containing the appropriate antibiotics (Table 2-3) and incubated overnight at 37 °C.

**Table 2-3.** List of competent cells/plasmids used during transformation with appropriate antibiotics.

Competent cells/ Plasmid	Antibiotic	Final concentration
pET28a plasmid	Kanamycin	50 µg/mL
BL21 (DE3) pLysS	Chloramphenicol	25 µg/mL
Rosetta	Chloramphenicol	25 µg/mL
60.1 GroES plasmid	Ampicillin	100 µg/mL
60.2 GroES plasmid	Ampicillin	100 µg/mL

Next day 5 mL of overnight culture was prepared using a single colony from LB agar plate in LB broth with appropriate antibiotics (Table 2-3) and incubated at 37 °C. Glycerol stocks were prepared by adding 500 µL of overnight culture to 500 µL of 50% glycerol and stored at -80 °C. 2 mL of overnight culture was added to 100 mL LB broth supplemented with appropriate antibiotics and allowed to grow at 37 °C until the OD<sub>600</sub> of 0.4-0.6. In this stage the culture was induced with ~ 1 mM IPTG and allowed to grow for additional 3 hours at 37 °C; or incubated at 16 °C for additional 3 hours and then induced with IPTG, followed by overnight growth at 16 °C. Cells were harvested (7000 g, 10 min at 4 °C), re-suspended in lysis buffer, 25 mM Hepes-NaOH pH 7.6, 500 mM NaCl and 10% (v/v) glycerol, and lysed by a sonicator (10 cycles, 20 seconds on, 40 seconds off). Recombinant protein was purified as explained in section 2.3.

### 2.3 Protein expression and purification

All the proteins used during this thesis were purified as described below unless it is stated. 10 mL of overnight culture was prepared using glycerol stock of protein of interest in LB broth with appropriate antibiotics (Table 2-3) and 1% glucose (only for Pep2 and TreS) and incubated at 37 °C.

The overnight culture was centrifuged and the pellet was re-suspended in fresh Terrific Broth (TB) or LB with appropriate antibiotics (Table 2-3). The culture was allowed to grow at 37 °C until the OD<sub>600</sub> of 0.6-0.8, cooled down to 16 °C for 3 hours and protein expression was induced with 0.1 mM (*MsmPep2*), 0.6 mM (*MsmTreS*), 1 mM IPTG (*MsmGlgE* and *GlgB*) followed by further incubation at 16 °C overnight. Cells were harvested (7000 g, 10 min at 4 °C), washed with phosphate buffered saline (PBS), re-suspended in 30 mL lysis buffer, 25 mM Hepes-NaOH pH 7.6, 1 M NaCl and 10% (v/v) glycerol for Pep2 and 25 mM Hepes-NaOH pH 7.6, 300 mM NaCl and 10% (v/v) glycerol for TreS, and frozen at -80 °C until needed.

The cells were lysed using a French Press (Thermo Spectronic FA-078). Protease inhibitor cocktail (Roche), 1 mM PMSF, 10 mM MgCl<sub>2</sub> and 10 µg/mL DNase I were added to the re-suspended cell pellets prior to French Press. The lysed solution was centrifuged at 20,000 rpm for 30 min at 4 °C and clear lysate was diluted 4 times with lysis buffer, filtered with 0.45 µm Millipore filter and loaded into a pre-equilibrated Ni-NTA column (1 mL, GE Healthcare) in room temperature. The column was eluted with a gradient of imidazole from 0 mM to 500 mM. The fractions were tested with SDS-PAGE.

The Pep2 and TreS eluates were diluted with enough 50 mM MES pH 5.8, 10% (v/v) glycerol; and 20 mM Bis-Tris pH 6.0, 10% (v/v) glycerol, respectively, to reduce the NaCl concentration to 50 mM prior to load into the HiTrap Q-column (1 mL, GE Healthcare Life Sciences). The protein was applied into the HiTrap Q-column pre-equilibrated with 50 mM MES pH 5.8, 10% (v/v) glycerol, 50 mM NaCl for Pep2 and 20 mM Bis-Tris pH 6.0, 10% (v/v) glycerol, 50 mM NaCl for TreS. The protein was eluted with a gradient of NaCl from 0 mM to 500 mM NaCl. Fractions were analyzed with SDS-PAGE and concentrated using Amicon Ultra-4 centrifugal filter units prior to load onto the gel filtration column.

**Purification of *MsmPep2*( $\Delta$ 70).** This mutant protein was overexpressed as mentioned earlier with 0.2 mM IPTG and 25 mM Hepes-NaOH pH 7.6, 1 M NaCl and 10% (v/v) glycerol as lysis buffer. Fractions containing Pep2  $\Delta$ 70 from Ni-NTA chromatography column were mixed together and diluted with enough 50 mM MES pH 6.0 and 10% (v/v) glycerol to reduce NaCl concentration to 300 mM prior to load into the HiTrap Q-column pre-equilibrated with latter buffer. The protein was eluted with a gradient of NaCl from 150 mM to 1 M NaCl.

#### **Purification of Rv3032 and Rv3031**

Rv3032 and Rv3031 were grown in TB media and LB media, respectively, using 1 mM and 0.2 mM IPTG. Lysis buffer was 25 mM Hepes-NaOH pH 7.6, 500 mM NaCl, 10% (v/v) glycerol (for Rv3032) and 25 mM Hepes pH 7.6, 1 M NaCl, 10% (v/v) glycerol (for Rv3031) supplemented with 33  $\mu$ M ADP-Glucose (A0627- Sigma-Aldrich). Fractions

containing either Rv3032 or Rv3031 were pulled down from Ni-NTA column, mixed together and diluted with enough 25 mM Tris/HCl pH 8.2, 10% (v/v) glycerol to reduce the NaCl concentration to 50 mM prior to load into the HiTrap Q-column pre-equilibrated with latter buffer. Both proteins were eluted from the Hi-Trap Q-column using a gradient of NaCl from 100 mM to 1 M NaCl.

## 2.4 Size exclusion chromatography

Size exclusion chromatography column was calibrated using albumin (1 mg/mL) as a standard protein, prior to running other proteins, and the graph containing void volume and elution volumes of both dimer and monomer peaks was recorded. *Mtb* and *MsmPep2* and *Pep2* ( $\Delta 70$ ) were applied to a pre-equilibrated HiPrep Sephacryl 26/60 S300 HR (GE Healthcare) column. Equilibration was carried out with 50 mM MES pH 5.8, 200 mM NaCl and 10% (v/v) glycerol. In case of *Pep2*  $\Delta 70$  column was equilibrated with 50 mM MES pH 6.0, 500 mM NaCl and 10% (v/v) glycerol. Proteins were eluted at a flow rate of 0.5 mL/min and monitored by UV absorbance at 280 nm. Collected fractions were analyzed with SDS-PAGE. For the elution of *MsmegGlgE* and *GlgB* from the same size exclusion matrix, the proteins were in 50 mM Bis-Tris propane pH 7.0 (*GlgE*) or 20 mM Tris/HCl pH 8.0 (*GlgB*), 150 mM NaCl, 10% (v/v) glycerol.

In order to study the complex formation between *TreS* and *Pep2* under different pH conditions, both proteins dialyzed against 50 mM Na-phosphate pH 6.0-9.0 in 0.5 unit intervals, 300 mM NaCl and 10% (v/v) glycerol. *TreS* and *Pep2* were mixed together in 1:1 molar ratio (30  $\mu$ M each) in 2 mL final volume and loaded onto the pre-equilibrated



HiPrep Sephacryl 26/60 S300 HR (GE Healthcare) column. Proteins were eluted and analyzed as explained earlier.

The interaction between Pep2 and GlgE was explored using 20 mM Bis-Tris pH 6.0, 10% Glycerol and 300 mM NaCl buffer.

Rv3032 and Rv3031 were eluted from SEC using 25 mM Tris/HCl pH 8.2, 150 mM NaCl, 10% v/v glycerol and 25 mM Tris/HCl pH 8.2, 500 mM NaCl, 10% v/v glycerol, respectively.

## **2.5 Checking the nature of *Msm*TreS and *Msm*Pep2 degradation with western blot**

Three different fractions of *Msm*TreS and Pep2 eluted from HiTrap Q HP column were reduced in sample buffer and electrophoresed on a 10% SDS-PAGE gel. For western blotting, protein bands were transferred from gel to a hydrophobic polyvinylidene difluoride (PVDF) membrane (Amersham Hybond-P, GE Healthcare). The membrane was washed with PBS (two times, each 10 min) and PBST (0.1% Tween 20 in PBS) and blocked overnight at 4 °C in 3% albumin in PBST buffer on a shaker. The membrane was washed once with PBST, two times with PBS (each 10 min) and incubated with anti-His6 (1:1000 added into the 3% albumin in PBS) for 1.5 h at room temperature on a shaker. Membrane was washed once with PBST, and two times with PBS. Membrane was incubated with secondary antibody, Anti Mouse IgG alkaline phosphatase, (1:25000 added into the 3% albumin in PBS) for 1.5 h at room temperature on a shaker. Membrane was

washed three times with PBST and PBS as mentioned earlier and developed in 10 ml of SigmaFast BCIP/NBT in a dark room.

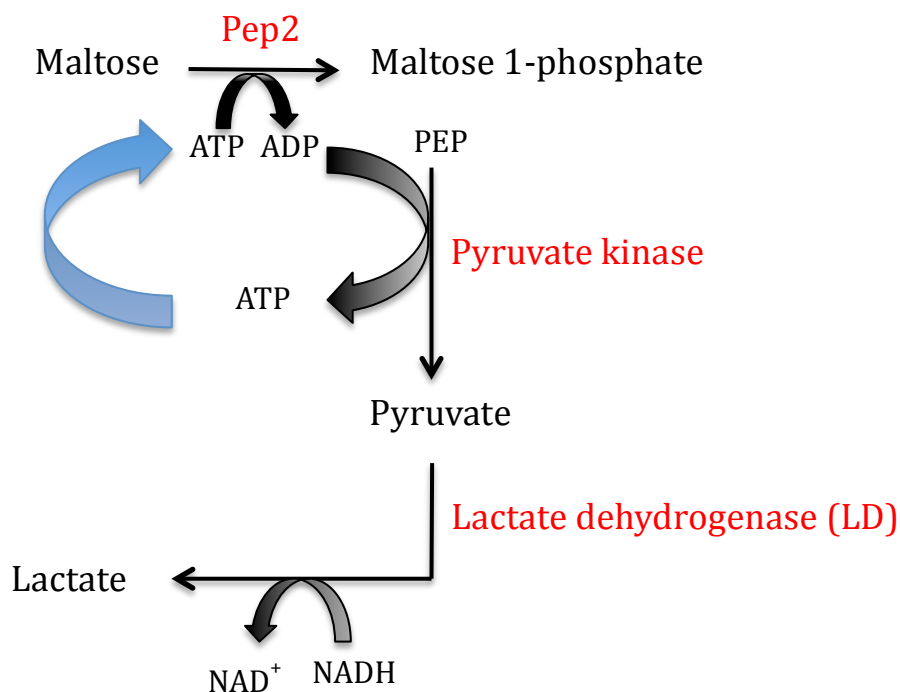
## **2.6 M1P purification from *MsmΔglgE***

4 mL of overnight culture was prepared using a loopfull of *MsmΔglgE* in 3.6 mL of Middlebrook 7H9 (MB7H9) supplemented with 0.5% glycerol, 0.25% Tween 80 and 10% OADC (Oleic Albumin Dextrose Catalase). Two shake flasks each containing 250 mL of MB7H9, were inoculated with 2 mL of overnight culture, followed by adding 0.25% Tween 80 into each. The cells were incubated at 37 °C with 120 rpm for 24 h (OD  $\approx$  0.8), harvested at 5000 rpm, 10 min, 4 °C, and kept at -80 °C until purification. Each pellet was re-suspended in 25 mL of PBS and 0.05% Tween 80, centrifuged for 10 min at 4000 rpm. The pellets were re-suspended in 4 mL of mili-Q water and boiled at 95 °C for 4 hrs. The samples were centrifuged at RT at 4000 rpm for 15 min, the supernatant was transferred to a new tube and dried overnight using a vacuum dryer (GeneVac, EZ-2, Suffolk, UK). Each sample was re-suspended in 400  $\mu$ l of mili-Q water and spotted on two Silica gels 60 F<sub>254</sub>, TLC plastic sheets (MERCK) and developed three times using solvent system butanol: ethanol: water (5:3:2). A thin layer (about 2 cm) of each TLC plate was cut and visualized using  $\alpha$ -naphthol sulfuric acid stain and used as a guideline to determine the M1P position on the TLC plates. M1P was scratched from TLC plates using a sharp blade and re-suspended in 3 mL of mili-Q water on a circular rotator for 15 min. Each sample was re-suspended and centrifuged three times at 4000 rpm at RT for 10 min and each time the supernatant was collected. The total volume (36 mL) was dried

overnight by vacuum dryer (GeneVac, EZ-2, Suffolk, UK) and re-suspended in 400  $\mu$ L of milli-Q water.

## **2.7 Activity assay**

Reaction mixtures probing TreS and Pep2 activity were prepared as 10  $\mu$ L of TreS (1 mg/mL), 100 mM trehalose or maltose, 50 mM Hepes buffer pH 7.5, 75  $\mu$ L H<sub>2</sub>O and 5.7  $\mu$ L of Pep2 (0.5 mg/mL), 100 mM maltose, 50 mM Hepes-KOH pH 7.5, 10 mM MgCl<sub>2</sub>, 5 mM ATP and H<sub>2</sub>O to reach the final volume of 100  $\mu$ L, respectively. The reaction mixture for each protein was incubated at 37 °C for 1 hour. The reactions were stopped by adding 100  $\mu$ L of ice cold acetone to each tube. Each tube was kept at -20 °C for at least 30 min, centrifuged at 14,000 rpm for 10 min at 4 °C, supernatant was collected and dried using a speed vacuum. The dried reaction product was re-suspended in 200  $\mu$ L of H<sub>2</sub>O, and 2  $\mu$ L of the solution applied on silica gel plate (Merck, Darmstadt, Germany) and developed using solvent system butanol: pyridine: water (7:3:1 v/v) for TreS and in butanol:ethanol:water (5:3:2 v/v) for Pep2. The plate containing Pep2 and M1P was developed three times in the same solvent system. Plates were visualized using  $\alpha$ -naphthol sulfuric acid stain and dried using a heat gun.



**Figure 2-1.** Measuring phosphorylation of maltose to maltose 1-phosphate using an enzymatically coupled reaction between ATP to ADP conversion and NADH to NAD<sup>+</sup> oxidation in the presence of pyruvate kinase and lactate dehydrogenase. PEP: phosphoenol pyruvate.

In order to measure the phosphorylation of maltose by Pep2, conversion of ATP to ADP was enzymatically coupled to the oxidation of NADH to NAD<sup>+</sup> in the presence of pyruvate kinase and lactate dehydrogenase (Figure 2-1). Pyruvate kinase catalyzes the transfer of one phosphate group from phosphoenol pyruvate (PEP) to ADP, generating one molecule of ATP and one molecule of pyruvate, which in turn is converted to lactate by lactate dehydrogenase (LD). NADH is fluorophore and its oxidation to NAD<sup>+</sup> can be monitored fluorimetrically by measuring its excitation at 340 nm and emission at 450 nm. The experiment was performed in duplicates in 96 well plates, using a BMG PHERAstar

FS microtiter plate reader. Data were recorded and analyzed using MARS and GraphPad Prism softwares, respectively.

The activity of Pep2 was monitored in a series of reactions where Pep2 concentration varied, from 0 to 2  $\mu\text{M}$ , at a fixed ATP concentration (0.2 mM). The reaction was set up as 50 mM Hepes-NaOH pH 7.5, 100 mM NaCl, 10% (v/v) glycerol, 10 mM  $\text{MgCl}_2$ , 0.1 mM NADH, 0.2 mM ATP, 4 mM phosphoenolpyruvate and 2 mM pyruvate kinase: lactate dehydrogenase couple (Sigma) at a final volume of 100  $\mu\text{L}$  and initiated with 20 mM maltose. In order to reduce the effect of dilution on pH, the buffer was prepared three times more concentrated. The effect of pH (6.0 to 7.5 in 0.5 unit increments) on the activity of Pep2 in the presence and absence of TreS (in 1:0.5, 1:1 1:2 and 1:4 molar ratios, respectively) was measured using the same set up and 0.02  $\mu\text{M}$  Pep2. When varying ATP (0 mM – 0.2 mM), maltose was present at 20 mM and when varying the maltose concentration (0 mM – 20 mM), the reaction was initiated by adding 0.3 mM ATP.

### **2.7.1 Enzymatic activity of Pep2 ( $\Delta 70$ )**

Reaction mixtures contained 50 mM Na-phosphate, pH 6.5, 300 mM NaCl, 10% (v/v) glycerol, 10 mM  $\text{MgCl}_2$  and Pep2 (0.45  $\mu\text{M}$  – 1.0  $\mu\text{M}$ ) or Pep2 ( $\Delta 70$ ) (0.45  $\mu\text{M}$  – 1.0  $\mu\text{M}$ ), 4 mM phosphoenolpyruvate, 2 units pyruvate kinase, 2 units lactate dehydrogenase and 0.1 mM NADH. When varying ATP (0 mM - 0.4 mM), maltose was present at 20

mM. Furthermore, the activity of both proteins were measured in the presence of 0.5 mM ATP and maltose varied from 0.1 mM – 3 mM.

## 2.8 Crystallization and data collection

Crystals of *MsmPep2* were grown by vapor diffusion in 96-well plates, using a Mosquito liquid handling system (TTP Labtech) to set up crystallization drops containing 100 nL of *MsmPep2* (40 mg/mL) and 100 nL of reservoir solution. JCSG+, Morpheus, MIDAS and Structure Screen (Molecular Dimensions) were used as reservoir solutions. In order to crystallize *MsmPep2* in complex with *MsmTreS*, both proteins were purified and concentrated separately and mixed in 1:1 (400  $\mu$ M each) and 1:2 (400  $\mu$ M: 200  $\mu$ M) molar ratios, respectively, just before crystallization. In addition to above screening solutions, Proplex was also included in the crystallization trials of *MsmTreS*-*Pep2* complex.

*MsmPep2* crystals were obtained from 2.1 M malic acid pH 7.0 and 2.4 M malonic acid pH 7.0. *MsmPep2* crystals were frozen in liquid nitrogen and transported in Unipuck cassettes to Diamond Light Source, beamline I03 (Oxford, UK) for data collection. *MsmTreS*:*Pep2* complex crystals were obtained from a reservoir of 9-10% polyethylene glycol 8000, 4% v/v glycerol, 200 mM  $\text{MgCl}_2$  and 0.1 M Tris-HCl pH 6.7-7.0. *MsmTreS*:*Pep2* crystals were briefly immersed in a cryoprotectant solution (the same mother liquor supplemented with 15% - 20 % glycerol in final concentration) and flash frozen in liquid nitrogen. Diffraction data were collected and processed using XDS/XSCALE. Phases were calculated using PHASER (McCoy et al., 2007) by

molecular replacement (MR) using PDB entry 3ZO9 (Caner et al., 2013) as the initial search model.

The first copy of TreS tetramer was placed using a tetrameric search model for PHASER (with a Z score > 30 after translation function). In order to place the second TreS tetramer, the TreS monomer was used as a search model and the three other monomers were found by performing NCS operations. The individual monomers were adjusted to the electron density map using rigid body fitting in COOT. Similarly, the first copy of *MsmPep2* was placed using molecular replacement (MR) and PDB entry 4O7O (*MtbPep2*) (Li et al., 2014). Additional copies of Pep2 were placed by performing NCS operations, real space rigid body fitting in COOT, refining the combined model in REFMAC5 and repeating the cycle.

## **2.9 Checking the expression and solubility of C-terminal his-tagged *MsmPep2***

The recombinant C-terminal his-tagged Pep2 cloned into the pET23b vector (provided by Dr. Rana Roy) was expressed as explained earlier with 100 µg/mL ampicillin in place of kanamycin. The N-terminal his-tagged *MsmPep2* was also expressed with 1 mM IPTG as a positive control.

## **2.10 Overexpression of Pep2 in *M. smegmatis***

pSD26 plasmid carrying *M. smegmatis pep2* gene was transformed into the *M. smegmatis* mc<sup>2</sup>155 cells by electroporation (Roy et al., 2013). 500 µL of LB media was added to the

transformation mixture and incubated at 37 °C for 4 hours. Transformation mixture was cultured on LB-hygB<sup>+</sup> plates (50 µg/mL hygromycin B) at 37 °C for four days. Several colonies (8-10 colonies) were used to inoculate a 4 mL culture (50 µg/mL hygromycin B and 0.05% Tween80) and grown for two days at 37 °C. The culture was propagated (1% v/v inoculation) into 50 mL LB (50 µg/mL hygromycin B and 0.05% Tween80) and grew for two days (OD<sub>600</sub> of 0.3-0.6). 5 mL of this culture was used to inoculate 1 L of LB (50 µg/mL hygromycin B and 0.05% Tween80) and further growth to the OD<sub>600</sub> of 0.3-0.6. The protein expression was induced with 0.2% w/v acetamide and further incubated at 16 °C overnight. Cells were harvested (7000 rpm, 10 min at 4 °C), washed with phosphate buffered saline (PBS), re-suspended in 30 mL lysis buffer, 25 mM Hepes-NaOH pH 7.5, 1 M NaCl and 10% (v/v) glycerol and frozen at -80 °C until needed. The cells were lysed by sonication (12 cycles of 60 s on, 90 s off at maximum amplitude). Protein purification was carried out as it was explained earlier.

### **2.11 Limited proteolysis of *MsmPep2***

Stocks of trypsin, α-chymotrypsin and proteinase K were prepared as 1 mg/mL in 1 mM HCl or 50 mM Tris/HCl pH 8.5 (for proteinase K). Limited proteolysis using trypsin (in molar ratios of trypsin: Pep2 as 1:20, 1:40 and 1:80), α-chymotrypsin (0.2-2\*10<sup>-6</sup> mg/mL) and proteinase K (0.2-2\*10<sup>-5</sup> mg/mL) was carried out for 1 hour at 37 °C. Pep2 concentration was about 3 mg/mL. Protease activity was inhibited by adding protease inhibitor cocktail and PMSF. Digestion of Pep2 was tested on SDS-PAGE.



## 2.12 Thermal shift assay

The thermal shift assay protein-dye solution was prepared using 35 µg/mL of protein of interest, 8x SYPRO Orange (Life Technologies) and appropriate amount of buffer 20 mM MES pH 5.8, 300 mM NaCl and 10% glycerol (for Pep2) and 20 mM Bis-Tris propane pH 7.0, 250 mM NaCl and 10% glycerol (for GlgE). 10 µL of thermal shift assay protein-dye solution was mixed with 20 µL of each screen (Table 2-4), to provide a final concentration of 100 mM Buffer and either 100 mM or 250 mM NaCl. Controls were included in the experiment and using water instead of reaction solution. The resulting solutions were loaded into a 96 well PCR plate. The plate was sealed with Optical-Quality Sealing Tape (Bio-Rad) and centrifuged at 2000 rpm for 4 min, followed by heating in a QPCR instrument (Mx3005P, Agilent Technologies) from 25 °C to 85 °C in 1 °C increments and staying at each temperature for 1 minute. The excitation and emission wavelengths for SYPRO Orange are 490 and 575 nm, respectively. Melting temperatures of each protein were calculated from cycle thresholds using Agilent software (Mx3005P Software).

**Table 2-4.** List of the buffers and their pH values used in the screen

Number	Buffer (150 mM)	pH
1	Sodium citrate/citric acid	4.0
2	Sodium acetate	4.2
3	Sodium acetate	4.8
4	Sodium citrate/citric acid	5.0
5	Sodium acetate	5.4
6	MES	5.7
7	Sodium citrate/citric acid	6.0
8	Bis-Tris	6.0
9	MES	6.1
10	Imidazole	6.4
11	MES	6.5
12	Bis-Tris	6.5
13	Bis-Tris	7.0
14	Hepes	7.0
15	Imidazole	7.1
16	Tris	7.2
17	Hepes	7.5
18	Imidazole	7.6
19	Hepes	8.0
20	Tris	8.0
21	Tris	8.6
22	CHES	8.6
23	CHES	9.1
24	CHES	9.6

### 2.13 Small Angle X-ray Scattering

Solution scattering data of *MsmTreS*, *MsmPep2* and the complex of these two proteins were recorded on beamline BM29 at ESRF, Grenoble. The protein samples were dialyzed against 50 mM Na-phosphate buffer (pH 6.5), 300 mM NaCl and 10% (v/v) glycerol. SAXS data were collected on five different concentrations for each protein sample (1.25, 2.5, 5, 7.5 and 10 mg/mL). The data for each protein concentration was collected in the following order, buffer measurement, protein solution and final buffer measurement. Each exposure contained 10 frames, each for 2 s in flow mode. The detector images were

averaged and reduced to 1-dimensional scattering curves, and buffer contributions to scattering subtracted using the beamline software BsxCuBE.

All SAXS data were analyzed using programs of ATSAS suite ([www.embl-hamburg.de/biosaxs/software.html](http://www.embl-hamburg.de/biosaxs/software.html)). Global parameters ( $R_g$  and  $D_{max}$ ) and pair distribution were calculated using PRIMUS (Konarev et al., 2003). Fitting curves were calculated using CRY SOL (Svergun, Barberato and Koch, 1995). Molecular envelopes were generated by both DAMMIF (Franke and Svergun, 2009) and DAMAVER (Volkov and Svergun, 2003).

#### **2.14 Analytical ultracentrifugation**

Sedimentation equilibrium was carried out in a Beckman Optima XL-A analytical ultracentrifuge equipped with absorbance optics. Protein samples were dialyzed overnight in 50 mM  $\text{NaPO}_4$  pH 6.5, 300 mM NaCl and 10% glycerol (Table 2-5). Equilibrium experiment was performed using 6-channel Epon centerpieces with quartz windows at a rotor temperature of 4 °C for a total duration of 116 hours. The data were recorded at the UV of 280 nm at 8000, 9000, 10000 and 11000 rpm. At each rotation speed, the sample was allowed to reach equilibrium during a 24 hours period. Finally, the sample was run at 25,000 rpm for 20 hours. The data were analyzed using SEDPHAT. Parameters for solvent density and viscosity and for the partial specific volume of the proteins were calculated using SEDNTERP.

**Table 2-5.** The concentrations of *Msm*TreS, TreS-Pep2 complex in 1:1 and 2:1 molar ratio before loading on analytical ultracentrifuge.

TreS	TreS: Pep2 (1:1 molar ratio)	TreS: pep2 (2:1 molar ratio)
2 $\mu$ M	1.25 $\mu$ M: 1.25 $\mu$ M	1.5 $\mu$ M: 0.75 $\mu$ M
4 $\mu$ M	2.5 $\mu$ M: 2.5 $\mu$ M	3 $\mu$ M: 1.5 $\mu$ M
6 $\mu$ M	3.75 $\mu$ M: 3.75 $\mu$ M	4.5 $\mu$ M: 2.25 $\mu$ M

## 2.15 Bacterial two-hybrid screen

The *Mtb* Rv0126, Rv0127, Rv1327c and Rv1326c genes encoding TreS, Pep2, GlgE and GlgB, respectively, were amplified by PCR (primers in Table) from *M. tuberculosis* H37Rv genomic DNA. The PCR amplification was carried out during 35 cycles: initial denaturation at 98 °C for 2 min, denaturation at 98 °C for 30 s, extension at 72 °C for 90 s (for *pep2*) and 59 °C for 1 min (for *glgB*) and final extension at 72 °C for 10 min. The PCR products were analyzed on a 1% agarose gel stained with 2% ethidium bromide. The PCR products were purified from gel using QIAquick Gel Extraction protocol using a microcentrifuge. *Xba*I and *Kpn*I restriction enzymes (for *treS*, *pep2* and *glgE*) and *Bam*HI and *Kpn*I (for *glgB*) were used to digest purified PCR products to generate overhang ends. The reaction was set up in a final volume of 50  $\mu$ L as 40  $\mu$ L PCR products, 5  $\mu$ L 10x FastDigest Buffer, 1  $\mu$ L of each restriction endonucleases, for 1 h at 37 °C. The reaction sample was loaded on a 1% agarose gel and purified using QIAquick Gel Extraction protocol using a microcentrifuge.

Overhang PCR products were ligated into pKT25, pKNT25, pUT18 and pUT18C vectors digested with appropriate restriction enzymes. Ligation reaction was set up different amounts of digested plasmids and PCR products (the ratio between vector and insert was calculated using [Insilico.uni-dusseldorf.de/lig\\_input.html](http://Insilico.uni-dusseldorf.de/lig_input.html) online program), 2  $\mu$ L T4 DNA Lygase buffer and 1  $\mu$ L T4 DNA Lygase (New England Biolabs, UK) in a final volume of 20  $\mu$ L incubated at room temperature for 10 minutes. One  $\mu$ L of each ligation reaction was used to transform 50  $\mu$ L of XL1-Blue cells using heat-shock protocol and cultured on LB plates with appropriate antibiotics (50  $\mu$ g/mL kanamycin for pKT25 and pKNT25; 100  $\mu$ g/mL ampicillin for pUT18 and pUT18C) overnight at 37 °C. The colonies were screened using gel electrophoresis and DNA sequencing.

The test of interaction between two pairs of proteins was performed as described by Battesti and Bouveret (2012). BTH101 and DHM1 competent cells were co-transformed with 0.5  $\mu$ L of two recombinant plasmids. One plasmid carrying the first gene fused to pKT25 or pKNT25 and second plasmid carrying the other gene of interest fused to pUT18 or pUT18C (8 plasmid combinations). 60  $\mu$ L of transformation reaction was cultured on LB plates supplemented with 100  $\mu$ g/mL ampicillin and 50  $\mu$ g/mL kanamycin and incubated at 30 °C for 48 hours. 5 mL overnight culture was prepared from transformation plates (several colonies were picked in order to reduce heterogeneity) with appropriate antibiotics and 0.5 mM IPTG, incubated at 30 °C with shaking. Next day, the interaction between two proteins of interest was measured using  $\beta$ -galactosidase assay on the same overnight liquid using Miller protocol. A loopfull from overnight culture was streaked on the LB plate containing appropriate antibiotics, 40  $\mu$ L X-gal and 0.5 mM

IPTG. Plates initially incubated overnight at 30 °C but next day transferred to room temperature for another 72 hours. Positive and negative controls included throughout the assay using two empty plasmids as negative control and plasmids pKT25-zip and pUT18C-zip as positive control.

## **2.16 $\beta$ -galactosidase activity assay**

The overnight culture was diluted 1 to 5 with M63 medium and the optical density at 600 nm of each culture was recorded. Bacterial cells were permeabilized with 30  $\mu$ L of toluen and 0.1% SDS solution per 2.5 mL of the diluted cells in a Falcon tube. Tubes were vortexed for 10 seconds and left open at 37 °C for one hour to allow the toluen to evaporate. In a glass tube, 0.5 mL of the permeabilized cells were mixed with 0.5 mL of PM2 buffer, followed by incubation in water bath at 28 °C for 5 minutes. A tube containing 1 mL of PM2 buffer was also prepared and used as blank. The enzymatic reaction was started by adding ONPG substrate, prepared in PM2 buffer without  $\beta$ -mercaptoethanol, in 0.4% final concentration. After 10 minutes the reaction was stoped by adding 500  $\mu$ L of 1 M Na<sub>2</sub>CO<sub>3</sub> stop solution. The optical density of each reaction was measured at 420 nm. The enzymatic activity of each sample was calculated using this equation:  $A = (200 \times OD_{420} \text{ of each sample} / \text{min of incubation}) \times \text{dilution factor}$ . The results were recorded as units/ mL that one unit represent hydrolysis of 1 nmol of ONPG in 1 min at 28 °C (BACTH System Kit, Euromedex).

## 2.17 Isothermal titration calorimetry assay

Isothermal titration calorimetry (ITC) experiments were performed using a VP-ITC Microcalorimeter from MicroCal, LLC (Northampton, MA). Both proteins, *MsmTreS* and *MsmPep2*, were dialysed against 50 mM NaPO<sub>4</sub> pH 6.5, 7.5 and 8.5, 300 mM NaCl and 10% v/v glycerol and degassed at 2 °C below the experimental temperature prior to use. Around 290 µL of *MsmPep2* (500 µM) was titrated into 1.4 mL of *MsmTreS* (75 µM as monomer) during 29 injections (10 µL each time) and 300 seconds time intervals between each injection. The experiment was conducted at 25 °C. The data were analyzed using Origin 7.0 software (Microcal, Northampton, MA) provided by the company. The experimental data were fitted to a theoretical titration curve using single and two sets of sites. The results were shown as  $n$  (Stoichiometry),  $K_a$  (association constant, in M<sup>-1</sup>),  $\Delta H$  (enthalpy change, in kcal mol<sup>-1</sup>) and  $\Delta S$  (entropy change, kcal K<sup>-1</sup> mol<sup>-1</sup>).

# Results



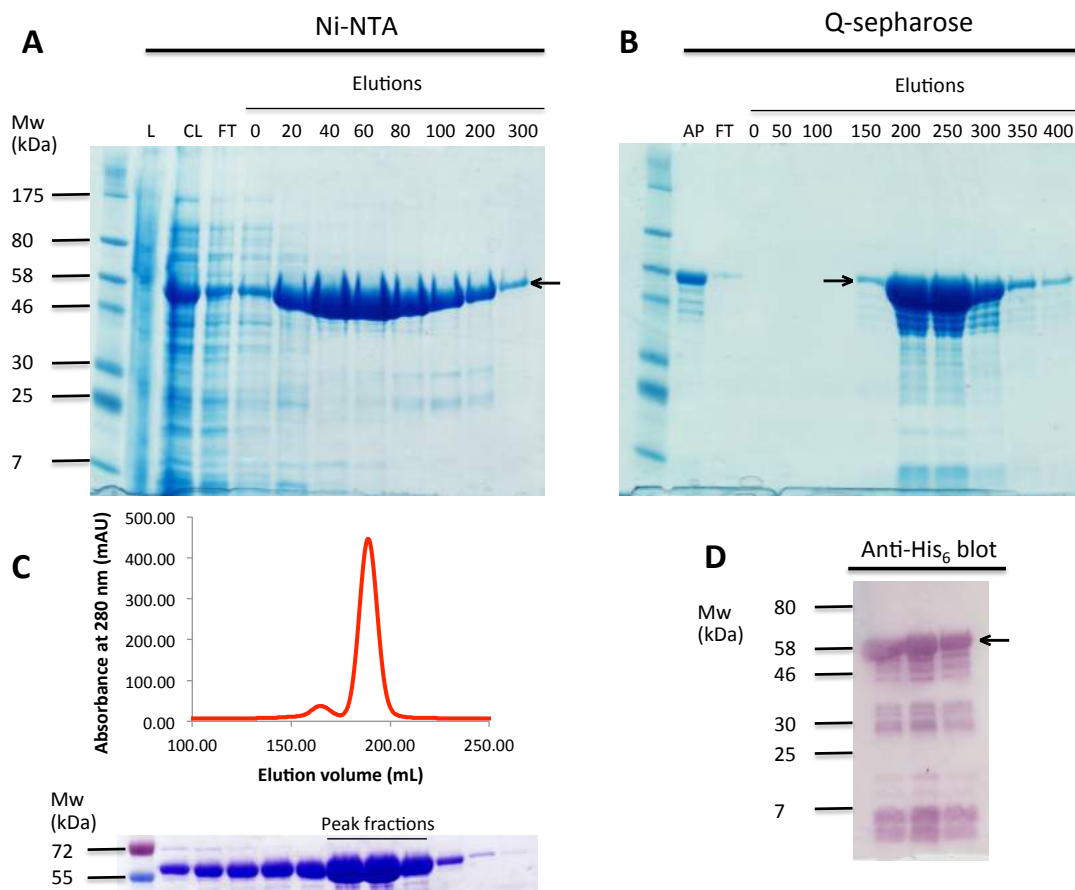
### **3 Crystallization of Pep2**

### 3.1 Crystallization of maltokinase

Maltokinase (Pep2) is a member of GlgE pathway that catalyzes the phosphorylation of maltose to maltose 1-phosphate in the presence of ATP. Although Pep2 from *M. tuberculosis* H37Rv is considered to be important for optimum growth (Sasseti, Boyd and Rubin, 2003), the crystal structure of this protein wasn't characterized prior to the start of this project. Roy and colleagues showed that *Mtb*Pep2 forms a complex with *Mtb*TreS and, surprisingly, this complex formation increases the activity of Pep2 in a dose-dependent manner (Roy et al., 2013). Due to lack of a crystal structure of Pep2 with TreS, we were not able to explain this rate enhancement. In order to answer these questions, I decided to attempt the crystallization of Pep2, initially working with the orthologue from *M. tuberculosis*, then that of *M. smegmatis*.

### 3.2 Expression and purification of *Mtb*Pep2 and *Msm*Pep2

Both *Mtb*Pep2 and *Msm*Pep2 were purified using immobilized metal ion affinity chromatography (IMAC) and anion-exchange chromatography (Figure 3-1 A-B). SDS-PAGE analysis of fractions collected from anion-exchange Q-Sepharose column showed that *Msm*Pep2 is amenable to degradation (Figure 3-1B). Degradation appears as lower bands with smaller molecular weights on the SDS-PAGE gel. In order to determine the nature of degradation, *Msm*Pep2 was analyzed with western blot using an antibody directed against the His-tag, demonstrating that degradation occurs from the C-terminal side (Figure 3-1D).

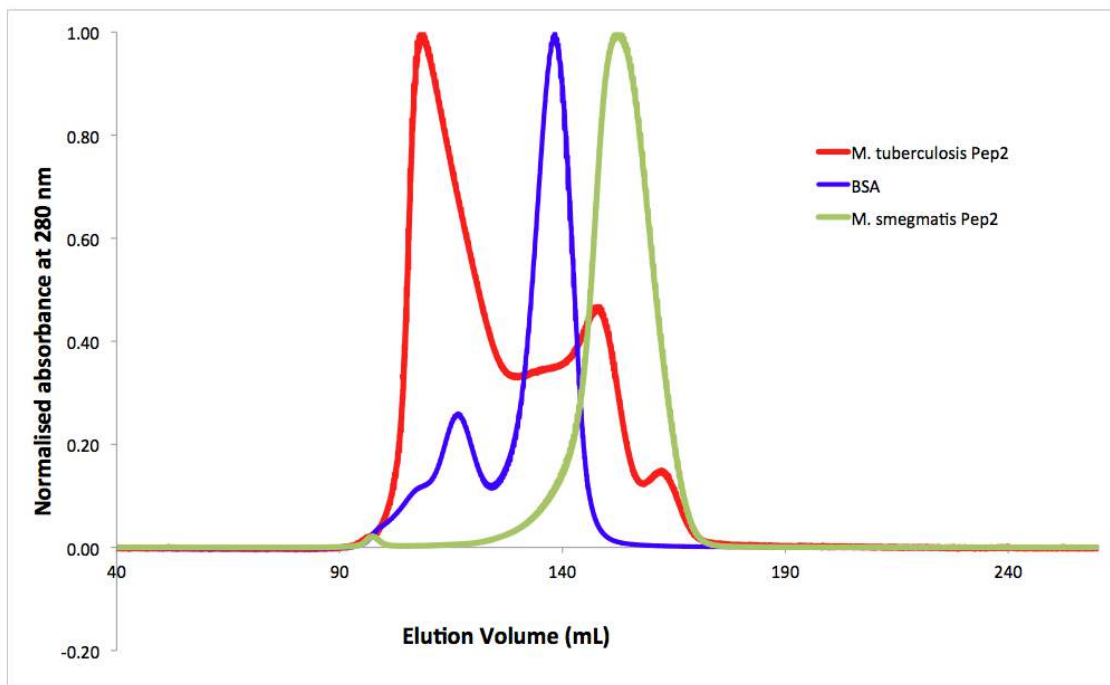


**Figure 3-1.** Purification of recombinant *MsmPep2*. (A-B) SDS-PAGE analysis of the fractions from affinity (Ni-NTA) and anion-exchange (Q-Sepharose) chromatography, respectively. L: crude lysate, CL: clear lysate, FT: flow through, AP: applied. Ni-NTA elution fractions (25 mM Hepes pH 7.5, 1 M NaCl, 10% (v/v) glycerol) supplemented with 0 to 300 mM imidazole and anion-exchange elution fractions (50 mM MES pH 5.8, 10% (v/v) glycerol) supplemented with 0 to 400 mM NaCl as indicated above the gels. (C) The size exclusion chromatogram of *MsmPep2* and SDS-PAGE analysis of the peak fractions from size exclusion chromatography. (D) Western blot analysis of *MsmPep2* (three protein-rich fractions from anion-exchange Q-Sepharose column) using an anti-His tag antibody. Arrows indicate His<sub>6</sub>-tagged Pep2 bands on the gel (~ 51 kDa with the His<sub>6</sub> tag).

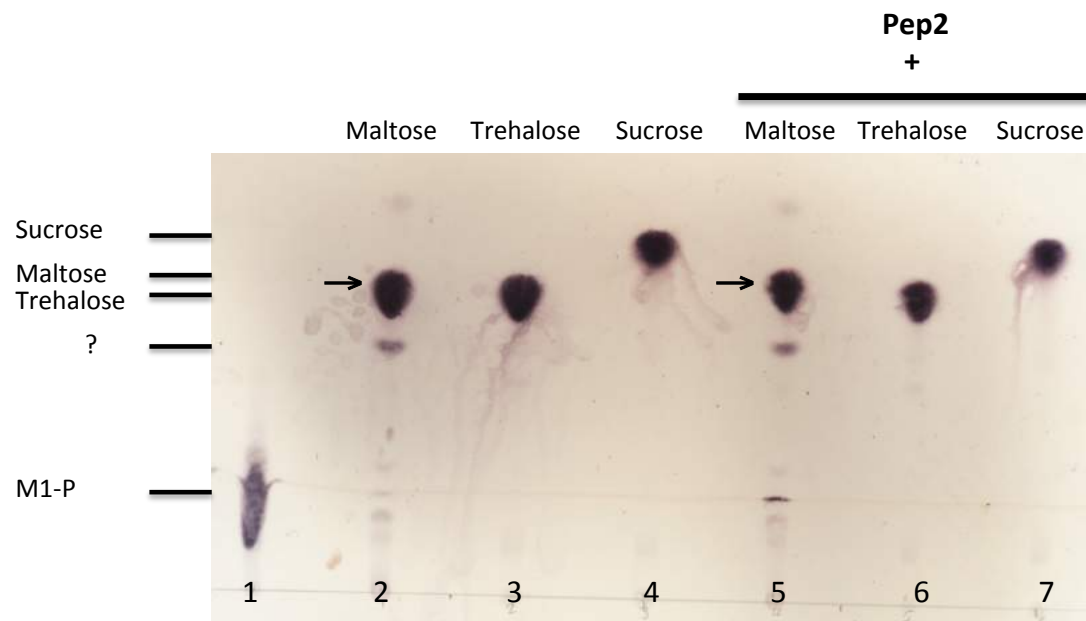
Size exclusion chromatography (SEC) was used to enhance the purity of *MtbPep2* and *MsmPep2* and also to determine the oligomeric state of each protein. *MsmPep2* was eluted as a dominant monomer peak at 151 mL, while *MtbPep2* showed a continuous size distribution with three major peaks at 108, 147 and 162 mL (using a Sephacryl S200 HR column) (Figure 3-2). Compare to the albumin standard, three peaks could represent the tetrameric, dimeric and monomeric state of *MtbPep2*. On a Sephacryl S300 column, *MsmPep2* elutes at 188 mL.

### **3.3 Verifying activity of *MsmPep2***

In order to probe the activity of *MsmPep2*, the reaction products were monitored in an end point assay using thin layer chromatography (TLC) (Figure 3-3). Incubating *Pep2* with maltose led to the production of maltose 1-phosphate, while neither trehalose nor sucrose were turned over.



**Figure 3-2.** The size exclusion chromatogram of *MtbPep2* and *MsmPep2* from a Sephacryl S200 column. Comparing the elution pattern of each protein with albumin demonstrates *MtbPep2* is in a dimer-tetramer configuration, while *MsmPep2* is monomer. The albumin double peak corresponds to masses of 133 and 66.5 kDa, respectively.



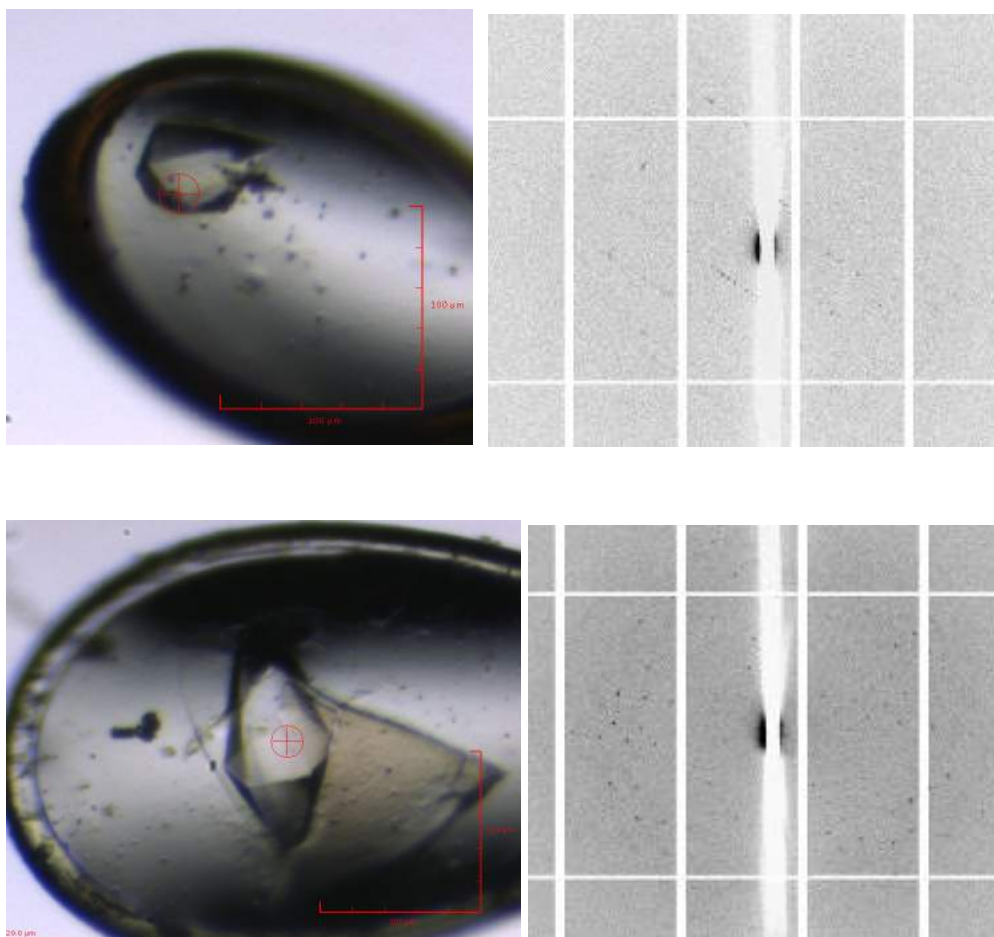
**Figure 3-3.** Thin layer chromatography of reaction mixtures probing the activity of *MsmPep2*. A volume of 2  $\mu$ L of the reaction mixture containing maltose, trehalose and sucrose, spotted on the lanes 5-7, respectively. Lane 1 represents maltose 1-phosphate and lanes 2-4 are the same activity reactions but without *Pep2* enzyme and used as a control to determine the place of saccharides on the plate. Arrows indicate the position of maltose on the plate.

### 3.4 X-ray diffraction of *MsmPep2*

*MsmPep2* crystals formed over a reservoir containing 2.1 M malic acid pH 7.0 and 2.4 M malonic acid pH 7.0 and diffracted to 7.5 Å. The unit cells parameters were determined as 123.2 Å, 123.2 Å, 265.9 Å (90.0°, 90.0°, 120.0°) (Figure 3-4). In order to improve the diffraction quality of the crystals, Pep2 protein was screened in a 2D screen (concentration of malic acid or sodium malonate vs. pH). In another attempt, available crystals were used as seeds against this optimization screen (streak seeding). In addition, Pep2 was also screened against the screening solutions supplemented with 20% (v/v) of 2.1 M malic acid pH 7.0 to promote crystallization, though, neither had any effects on the diffraction quality of *MsmPep2* crystals. In order to co-crystallize *MsmPep2* with maltose, the concentrated protein (33 mg/mL) was mixed with maltose (30 mM) and AMPPNP (5 mM) and crystal trays were set up. The resultant crystals diffracted very poor or did not diffract at all.

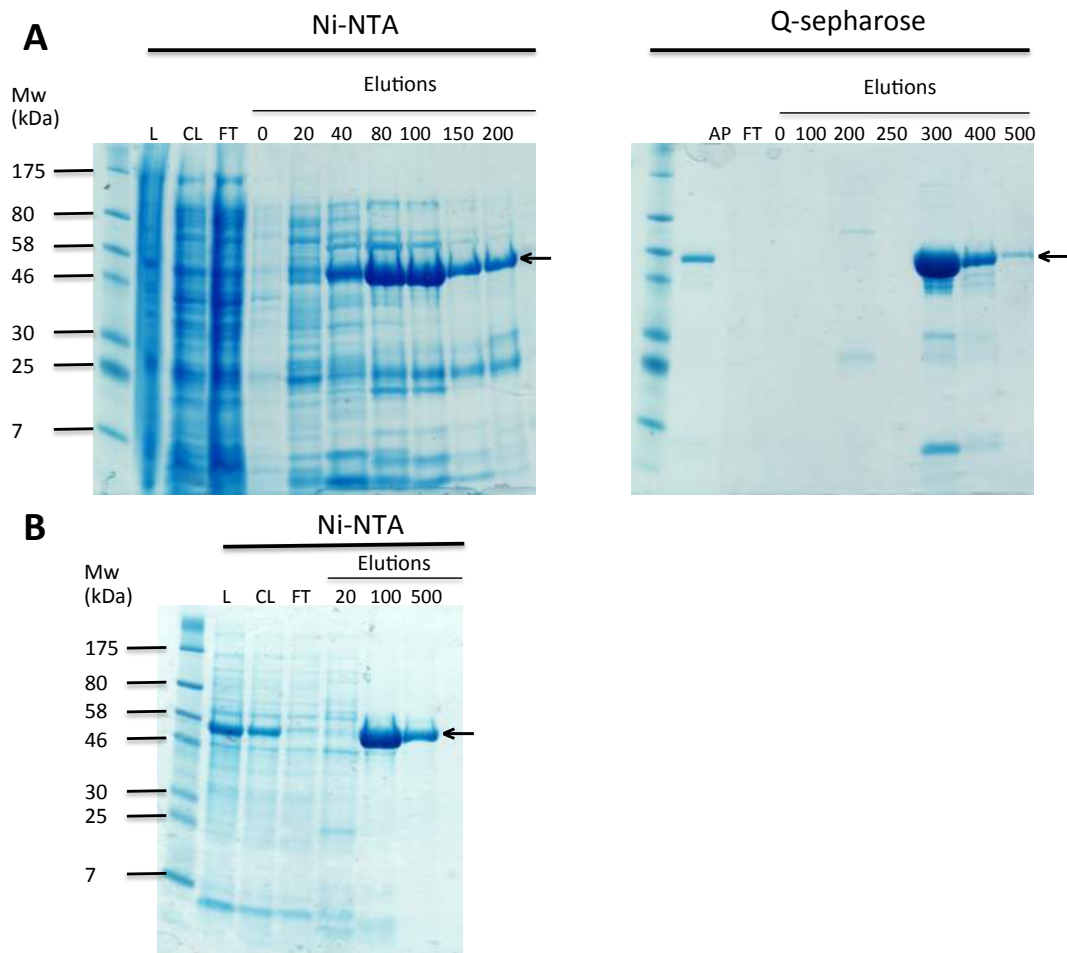
### 3.5 Expression and purification of C-terminally tagged *MsmPep2*

In an attempt to generate higher diffraction-quality crystals of *MsmPep2*, C-terminally tagged and *M. smegmatis* overexpressed Pep2 were purified using IMAC and anion-exchange chromatography (Figure 3-5A-B). Both proteins were concentrated to 27 mg/mL and 22 mg/mL, respectively, and used to set up crystallization screens. However, neither of these proteins was able to generate any crystals.



**Figure 3-4.** The diffraction pattern of *MsmPep2* crystals grown in 2.1 M malic acid pH 7.0 and frozen in liquid nitrogen. Crystals diffracted to 7.5 Å.



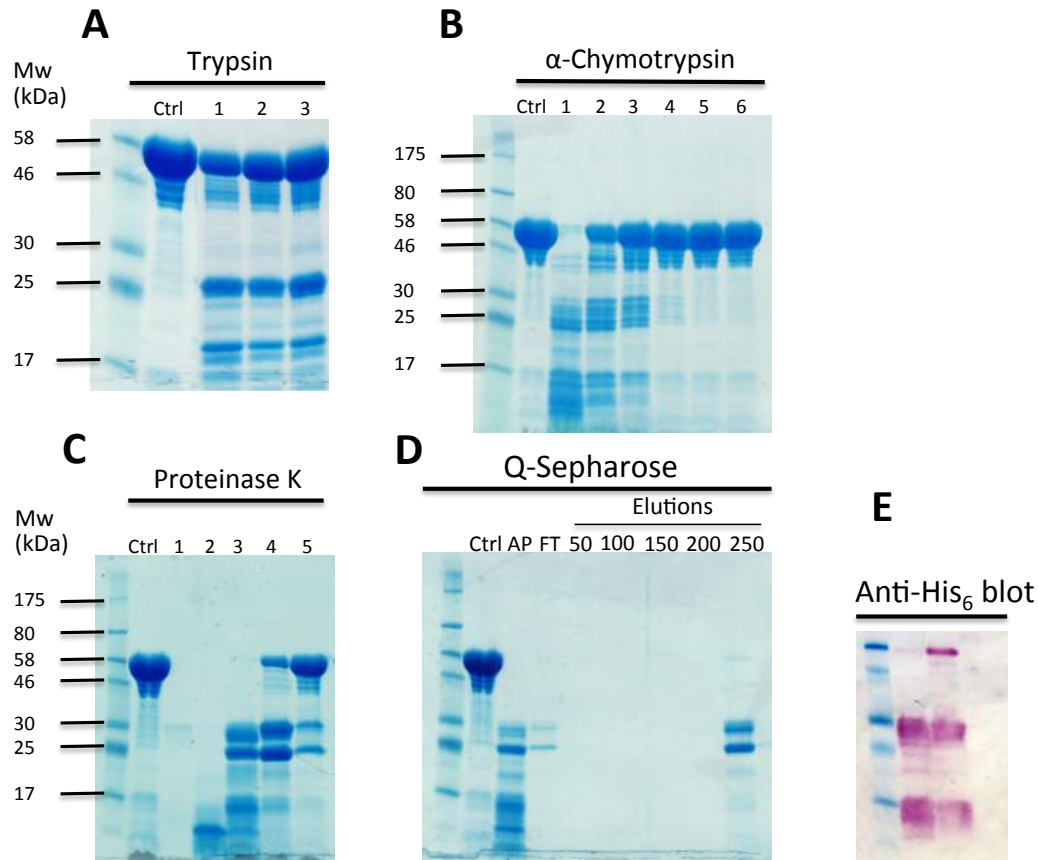


**Figure 3-5.** Purification of recombinant C-terminally tagged *MsmPep2* (A) and *Pep2* overexpressed in *M. smegmatis* (B). (A-B) SDS-PAGE analysis of the fractions from affinity (Ni-NTA) and anion-exchange (Q-Sepharose) chromatography. L: crude lysate, CL: clear lysate, FT: flow through, AP: applied. Ni-NTA elution fractions (25 mM Hepes pH 7.5, 1 M NaCl, 10% v/v glycerol) supplemented with 0 to 500 mM imidazole and anion-exchange elution fractions (50 mM MES pH 5.8, 10% v/v glycerol) supplemented with 0 to 500 mM NaCl as indicated above the gels. Arrows indicate His<sub>6</sub> tagged *Pep2* bands on the gel (~ 51 kDa with the His<sub>6</sub> tag).

### 3.6 Limited proteolysis of *MsmPep2*

In order to improve the crystallization of *MsmPep2* by removing potential unstructured fragments from either side of protein, limited proteolysis was carried out using three different proteases, trypsin,  $\alpha$ -chymotrypsin and proteinase K (Figure 3-6). Limited proteolysis of *MsmPep2* using trypsin led to the formation of two major smaller fragments with molecular weights of about 25 and 19 kDa, while digesting this protein using  $\alpha$ -chymotrypsin (final concentration of 0.2 mg/mL) led to formation of several smaller fragments with molecular weights ranging from 7 to 30 kDa. Similar to trypsin, proteinase K (final concentration of 0.002 mg/mL) produced two major smaller fragments with molecular weights of 25 and 30 kDa. In principle we were interested to generate only one or two fragments with a MW very similar to the undigested one but probably better folded or more stable. We stopped continuing with  $\alpha$ -chymotrypsin since it generated several fragments with small MWs. Although trypsin produced only two fragments during digestion (similar to proteinase K), it also failed to meet our criteria, as those two fragments were smaller than half of the size of undigested protein. Thus, we chose the proteinase K for further experimentation.

During anion-exchange chromatography of digested products with proteinase K, the two fragments 25 and 30 kDa pulled down together in one fraction but they appeared as two bands on the SDS-PAGE gel, suggesting these two fragments are linked together by non-covalent bonds. Furthermore, western blot using an anti-his6 tag antibody demonstrated that the cleavage site should be somewhere in the middle of protein as only the 30-kDa fragment contained his tag but not 25-kDa fragment.



**Figure 3-6.** SDS-PAGE analysis of *MsmPep2* digested with trypsin (A) α-chymotrypsin (B) and proteinase K (C); anion exchange chromatography (D) and western blot (E) of proteinase K digested products. Ctrl in all figures represent undigested *MsmPep2* as control. Lanes 1-3 in (A) represent reaction mixtures of *MsmPep2* with trypsin in 20:1, 40:1 and 80:1 ratios, respectively. Lanes 1-6 in (B) represent reaction mixtures of *MsmPep2* (2.6 mg/mL) with different concentrations of α-chymotrypsin from  $0.2 \cdot 10^{-6}$  mg/mL. Lane 1-5 in (C) represent reaction mixtures of *MsmPep2* (2.6 mg/mL) with different concentrations of proteinase K from  $0.2 \cdot 10^{-5}$  mg/mL. (D) Fractions from anion-exchange (Q-Sepharose) chromatography eluted with 50-250 mM NaCl as indicated above the gel, AP: applied, FT: flow through. (E) Western blot analysis of proteinase K digested *MsmPep2* using an anti-His<sub>6</sub> antibody.

proteinase K digested *MsmPep2* (both fragments 25 and 30 kDa) was concentrated to 27 and 54 mg/mL and exposed to crystallization screens but failed to produce any crystals.

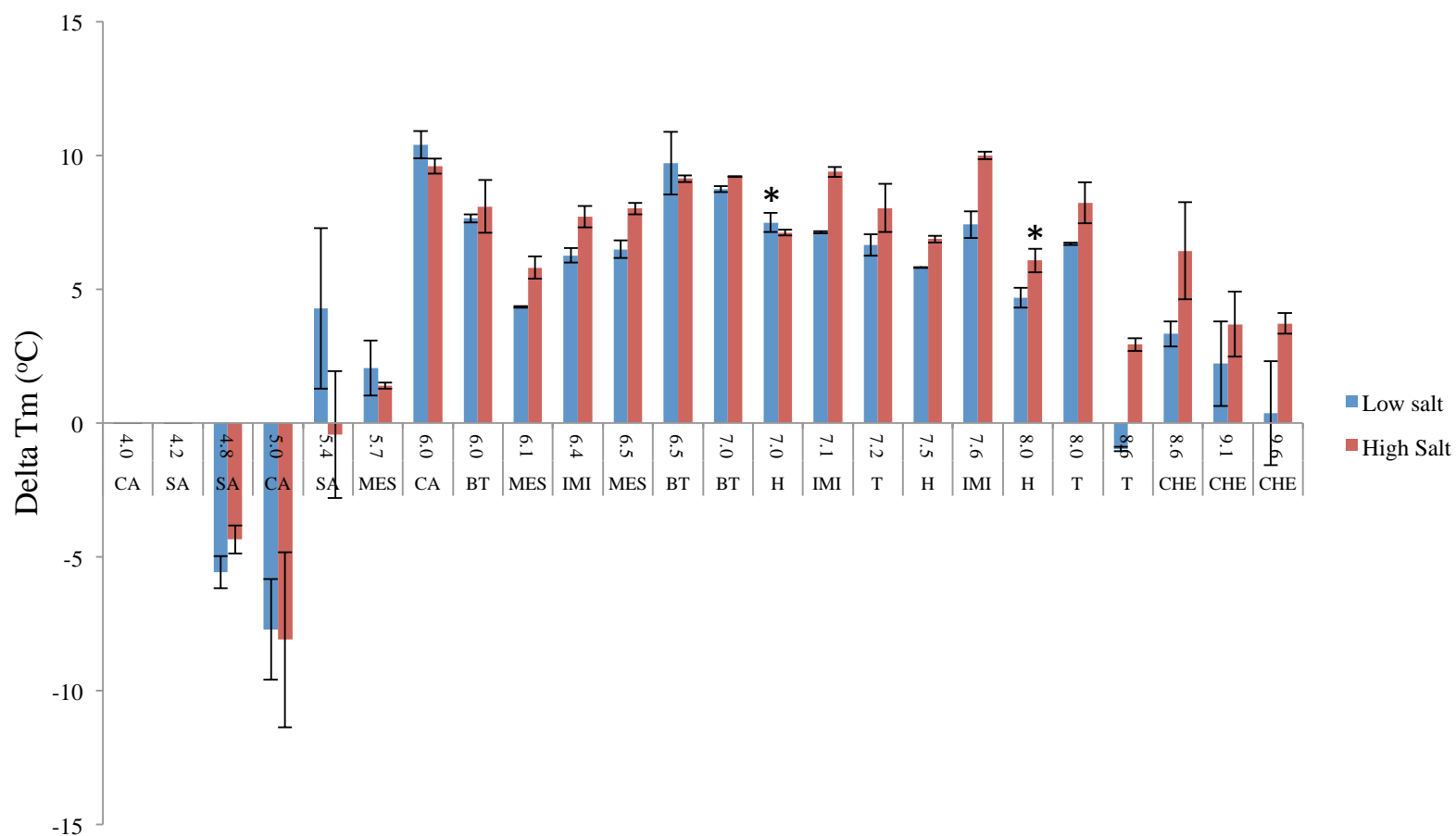
### **3.7 Increasing pH promotes *MsmPep2* stabilization**

In order to probe thermal stability of *MsmPep2* as a function of buffer composition, a thermal shift assay was performed using a qPCR thermal cycler and changes in melting temperature ( $T_m$ ) of this protein were monitored (Niesen, Berglund and Vedadi, 2007). Improving the buffer composition can increase homogeneity, stability and solubility of a protein and therefore, promoting the crystallization rate of a protein (Ericsson et al., 2006). This method is based on fluorescent properties of the dye SYPRO Orange in non-polar environments. SYPRO Orange is intrinsically a fluorescent dye, but its fluorescence emission significantly increases when it is exposed to non-polar environments.

Gradual rise in temperature causes the protein to start unfolding, which exposes hydrophobic amino acids, valine, leucine, isoleucine, phenylalanine and tryptophan, to the dye and increases the intensity of fluorescence. Fluorescence intensity starts decreasing after reaching a maximum point. This could be due to precipitation of the complex of SYPRO Orange and the denatured protein (Ericsson et al., 2006). The melting temperature correlates with stability of the protein, as higher melting temperature represents that the protein is better folded and therefore is in a more stable state.

The  $T_m$  value of *MsmPep2* was measured as 20 °C and its changes ( $\Delta T_m$ ) were monitored throughout the experiment. An increase in  $T_m$  could suggest an elevation in structural

order of *MsmPep2*, whereas a decrease in  $T_m$  might represent a disordered conformation or a sign of misfolding. Large variation in stability was observed in different screens. Some buffers had a higher impact on *MsmPep2* stability, mainly sodium citrate/citric acid pH 6.0, Bis-Tris pH 6.5 and 7.0, imidazole pH 7.1 and 7.6 with a  $\Delta T_m$  of about 10. The most destabilizing buffers are sodium acetate pH 4.8 and sodium citrate/citric acid pH 5.0 with a  $\Delta T_m$  of about -4 and -8, respectively. However, it is apparent that increasing the pH from 4.8 towards more basic pH values, up to 8.0, increases the melting temperature and promotes the stability of *MsmPep2*. In most screens there is not a significant difference between 100 mM and 250 mM NaCl on the stability of *MsmPep2*. *MsmPep2* was dialyzed against two buffers, Hepes pH 7.0 and 8.0; supplemented with 250 mM NaCl and 10% (v/v) glycerol and exposed to crystallization screens. Unfortunately, neither of these buffers improved the crystallization of *MsmPep2* protein.



**Figure 3-7.** *MsmegmPep2* stability versus buffer type determined using thermal shift assay. Stability was measured at 100 mM NaCl (■) and 250 mM NaCl (■). Stars above the bars indicate the buffers that were tested in the crystallization screenings. Error bars represent standard deviation. Abbreviations on X axis are explained in Table 2-4 of materials and methods.

### 3.8 Summary

Pep2 was successfully purified to more than 95% purity (single band on the SDS-PAGE gel) from *Mtb* and *M. smegmatis* (Figure 3-1). Oligomeric state of this protein in the solution was determined using size exclusion chromatography (Figure 3-2). *Mtb*Pep2 was eluted from SEC column as a continuous size distribution (multimer) at 108, 147 and 162 mL, while *Msm*Pep2 appeared as a dominant monomer peak at 151 mL. Given that *Msm*Pep2 behaves as a monomer in the solution, this protein was chosen for crystallization trials. *Msm*Pep2 crystallized over a reservoir containing 2.1 M malic acid pH 7.0 and 2.4 M malonic acid pH 7.0 and diffracted to 7.5 Å. In order to improve the diffraction of *Msm*Pep2 and to find the best condition for crystallizing this protein, folding and stability of Pep2 were further characterized by limited proteolysis and thermal shift assay. Limited proteolysis of *Msm*Pep2 by proteinase K produced two major smaller fragments with MWs of 25 and 30 kDa. Anti-His<sub>6</sub> blotting of these two fragments indicated that cleavage occurs in the middle of Pep2 and two fragments are linked together by non-covalent bonds (two fragments eluted together from anion-exchange chromatography column). However, attempts to crystallize these two fragments together were not successful. The thermal stability of *Msm*Pep2 as a function of buffer composition was explored in a thermal shift assay (Figure 3.7). The assay suggested that elevating the pH increases the melting temperature and therefore promotes the stability of *Msm*Pep2, with the most pronounced impact at sodium citrate/citric acid pH 6.0, Bis-Tris pH 6.5 and 7.0, imidazole pH 7.1 and 7.6 (Figure 3.7). Nevertheless, neither of these buffers were able to improve the crystallization of *Msm*Pep2.

## **4 Characterization of the *Mycobacterium* *smegmatis* TreS:Pep2 complex**

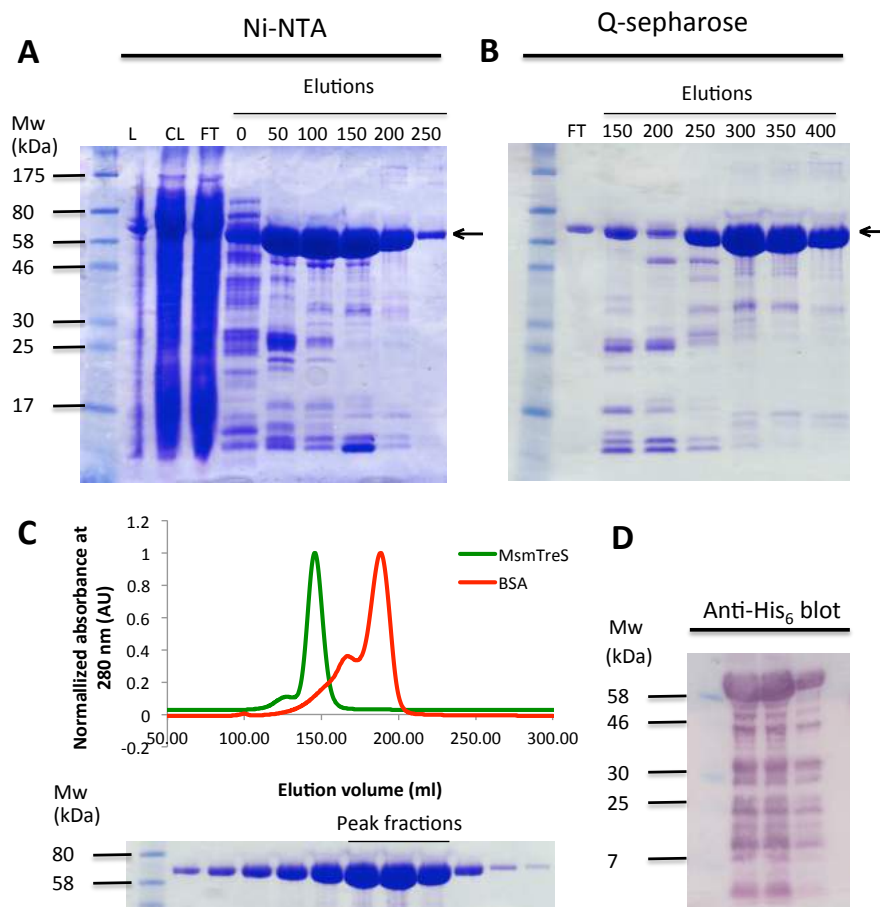


#### **4.1 Crystallization of *Msm*TreS:Pep2 complex**

The crystal structures of *Mtb*TreS (Roy et al., 2013), *Mtb*Pep2 (Li et al., 2014) and *M. vanbaalenii* Pep2 (Fraga et al., 2015) were determined by others prior and during the course of this project. Roy and colleagues demonstrated complex formation between *Mtb*TreS and *Mtb*Pep2, using size exclusion chromatography (SEC) and analytical ultracentrifugation (AUC). Using AUC in sedimentation equilibrium mode they determined the stoichiometry, but not the affinity of complex formation and were not able to solve the structure of neither Pep2 nor complex. In addition, they were not able to explain the *Mtb*Pep2 rate enhancement, since having structure of the complex was necessary. During this study I aimed to solve the structure of *Msm*TreS:Pep2 complex, determine the affinity of binding and how complex formation affects the activity of *Msm*Pep2.

#### **4.2 Protein expression and purification of *Msm*TreS and *Msm*Pep2**

Both *Msm*Pep2 and *Msm*TreS were purified using immobilized metal ion affinity chromatography (IMAC) and anion-exchange chromatography (Figures 3-1 and 4-1). SDS-PAGE analysis of fractions collected from anion-exchange Q-Sepharose column showed that, similar to *MsmPep2*, *Msm*TreS is susceptible to degradation (Figure 4-1B). Degradation appears as lower bands with smaller molecular weights. In order to determine the nature of degradation, both proteins were analyzed with western blot using an antibody directed against the His<sub>6</sub> tag, demonstrating that degradation occurs from the C-terminal side (Figure 4-1D).



**Figure 4-1.** Purification of recombinant *MsmTreS*. (A-B) SDS-PAGE analysis of the fractions from affinity (Ni-NTA) and anion-exchange (Q-Sepharose) chromatography, respectively. L: crude lysate, CL: clear lysate, FT: flow through. Ni-NTA elution fractions (25 mM Hepes pH 7.6, 300 mM NaCl, 10 % (v/v) glycerol) supplemented with 0 to 250 mM imidazole and anion-exchange elution fractions (20 mM Bis-Tris pH 6.0, 10% (v/v) glycerol) supplemented with 150 to 400 mM NaCl as indicated above the gels. (C) The size exclusion chromatogram of *MsmTreS* vs. albumin and SDS-PAGE analysis of the peak fractions from size exclusion chromatography. The albumin double peak corresponds to masses of 133 and 66.5 kDa, respectively. (D) Western blot analysis of *MsmTreS* (three protein-rich fractions from anion-exchange Q-Sepharose column) using an anti-His<sub>6</sub> tag antibody. Arrows indicate His<sub>6</sub>-tagged *MsmTreS* bands on the gel (~ 70 kDa with the His<sub>6</sub> tag).

*MsmTreS* eluted as a single peak from size exclusion chromatography at 144 mL, representing tetrameric state of this protein in the solution (Figure 4-1C). Likewise, *MsmPep2* was eluted as a dominant monomer peak at 188 mL from size exclusion column (Figure 3-2).

### **4.3 Verifying activity of *MsmTreS***

In order to probe the activity of recombinant *MsmTreS*, the reaction products were analyzed in an end point assay using thin layer chromatography (TLC) (Figure 4-2). Incubation of *MsmTreS* with trehalose or maltose interconverted the substrates. Also, a third molecule was observed on the TLC plate, which is likely glucose (glucose and fructose have similar positions on the TLC plate) suggesting hydrolysis also happens in addition to the catalysis reaction. Incubating sucrose,  $\alpha$ -1,4 glucosyl fructose, with *MsmTreS* produced trehalose as well as fructose consistent with the data for TreS from *M. smegmatis* (Zhang et al., 2011).

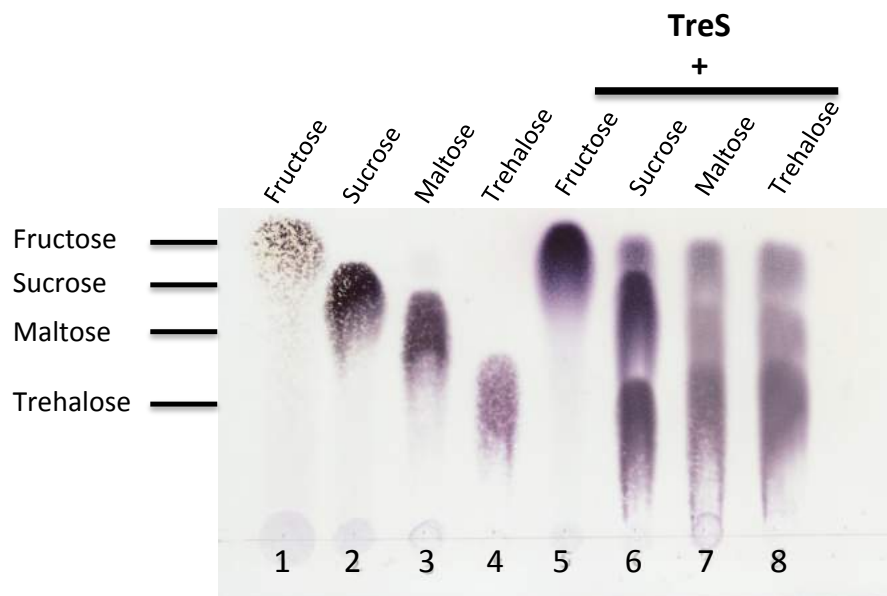
### **4.4 *MsmTreS* forms a complex with *MsmPep2***

*MsmTreS* and *MsmPep2* eluted individually as a single peak from size exclusion chromatography at 144 and 188 mL, respectively, representing tetrameric and monomeric state of these proteins in solution (Figure 4-3), consistent with published data (Roy et al., 2013). However, combining *MsmTreS* and *MsmPep2* in 4:1, 2:1, 1:1, 1:4 and 1:6 molar ratios, respectively, changed the elution pattern of both TreS and Pep2 proteins (Figure 4-3). They eluted as a single peak at lower elution volume indicating complex formation

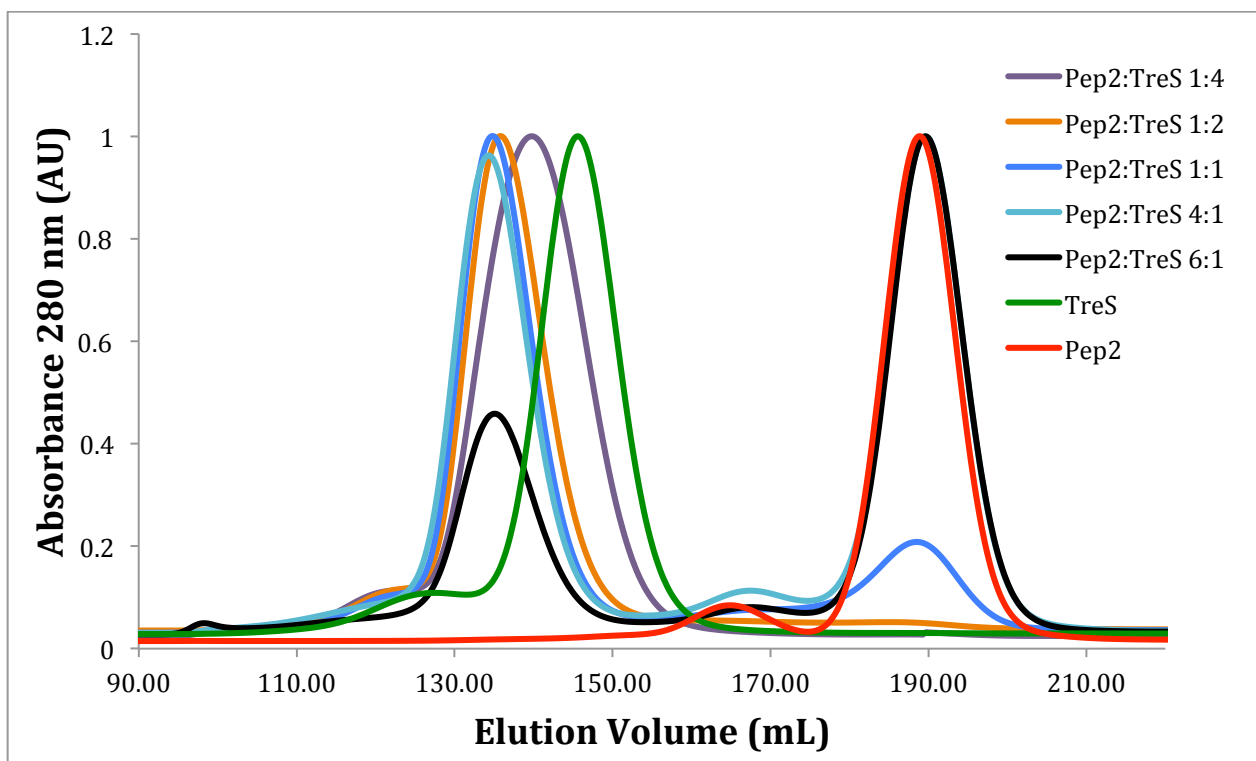
between these two proteins. Nevertheless, the excess amounts of Pep2 as free molecules in the solution eluted at the same elution volume as Pep2 (Figure 4-3). Adding *MsmPep2* to *MsmTreS* from 0.25 up to 0.5:1 molar ratio led to a shift of the *MsmTreS* tetramer peak towards shorter retention times (Table 4-1). Adding additional amounts of *MsmPep2* (up to 1:1 molar ratio) slightly shifted the peak (from 136 mL to 134.6 mL) but the free *MsmPep2* (not bound to *MsmTreS*) appeared as a second peak at 188 mL.

#### **4.5 *MsmTreS*:Pep2 complex formation is pH dependent**

*MsmTreS*:Pep2 complex formation was examined under different pH conditions, from 6.0-9.0, using size-exclusion chromatography in sodium phosphate buffer (Figure 4-5). At pH 9.0, *MsmTreS*:Pep2 complex eluted at 142 mL. At pHs below 9.0, the complex peak shifted towards lower elution volumes (134 mL at pH 6.0). Moreover, the peak height of the complex in acidic pH is higher than basic pH, demonstrating a distinct preference of complex formation at acidic pH. Plotting elution volume of complex (mL) vs. different pH values would indicate a sigmoidal relationship between these two parameters.



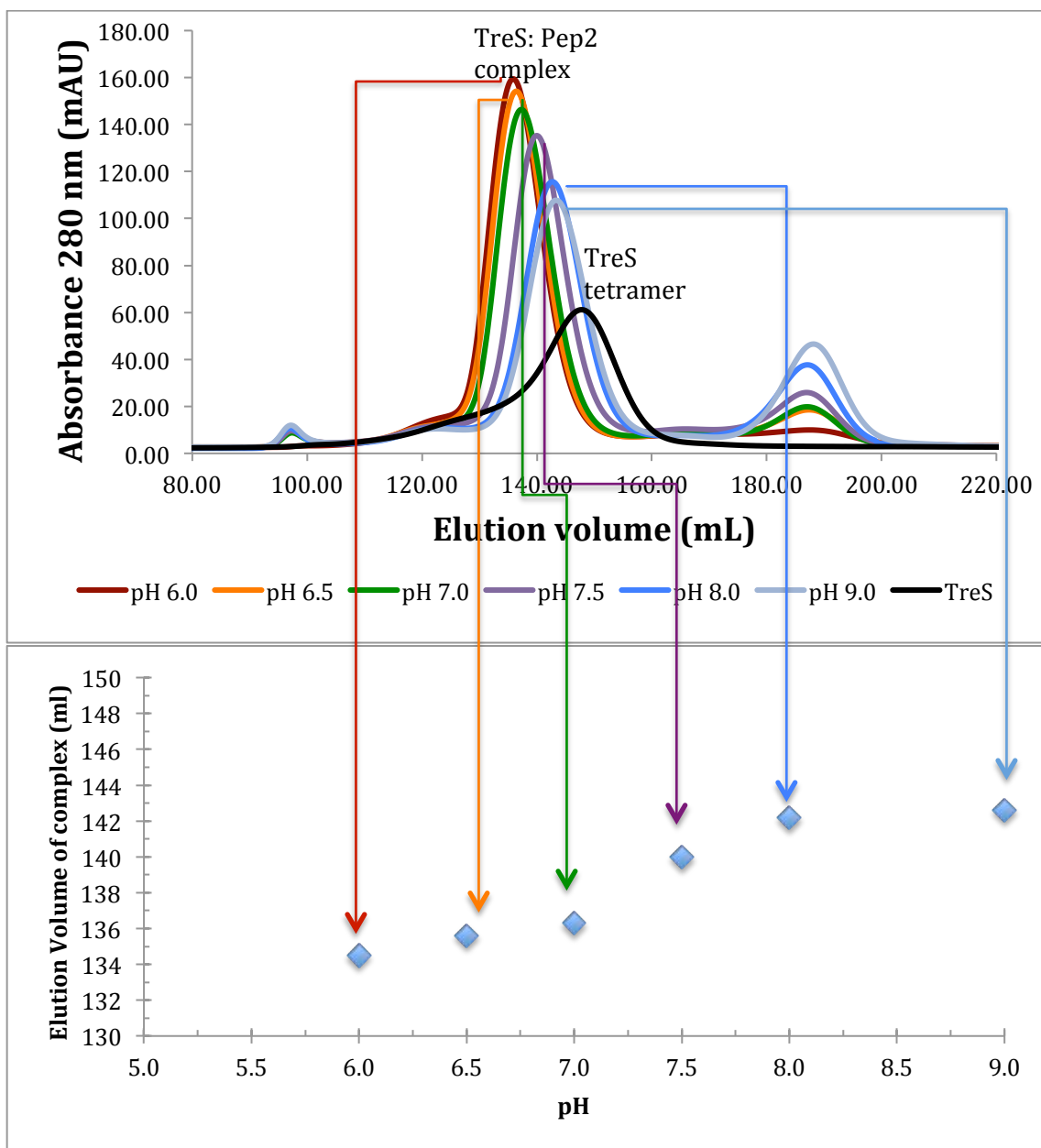
**Figure 4-2.** Thin layer chromatography analysis of reaction products probing the activity of *MsmTreS*. 2  $\mu$ L of the reaction mixture containing fructose, sucrose, maltose and trehalose, respectively, spotted on the lanes 5-8. Lanes 1-4 are the same activity reactions but without *MsmTreS* enzyme and used as a control to determine the place of saccharides on the plate.



**Figure 4-3.** Exploring complex formation between *Msm*TreS and *Msm*Pep2 in a range of molar ratios using size-exclusion chromatography.

**Table 4-1.** Size exclusion chromatography data of complex formation between *Msm*TreS and *Msm*Pep2.

Molar ratio <i>Msm</i> Pep2: <i>Msm</i> TreS	0.25:1	0.5:1	1:1	4:1	6:1
Elution volume (mL)	139.5	136	134.6	134	134.6



**Figure 4-4.** Exploring complex formation between *MsmTreS* and *MsmPep2* under different pH values using sephacryl S300-HR size exclusion resin. (Top section) Traces of absorbance (at 280 nm) vs elution volume. (Bottom section) The elution volume of each complex is plotted against corresponding pH value, demonstrating the shift of complex peak elution volume. Proteins were dialyzed against Na-phosphate buffer at specified PH values.

#### 4.6 Assessing binding affinity of the complex by isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a useful technique capable of determining the thermodynamic parameters (stoichiometry, affinity and enthalpy) of an interaction between macromolecule and ligand just in one single experiment (Pierce, Raman and Nall, 1999). The calorimeter is composed of two cells; a reference cell containing water and a sample cell containing the sample protein. Ligand is added in small aliquots to the sample cell with an automated syringe and due to its interaction with macromolecule heat will be released, exothermic reaction, or heat will be absorbed, endothermic reaction. ITC can measure this heat changes in sample cell by determining the electrical power required to keep both cells at the same temperature. As the ligand concentration increases the macromolecule becomes saturated and less heat is released or absorbed. Then the area of each peak is integrated and plotted against molar ratio of ligand to protein. A binding model (e.g., one set of binding site or two sets of binding sites) (equations 4-1 and 4-2) will be fitted into the data using a nonlinear regression procedure. The model is a mathematical description that helps to understand the actual physical, chemical or biological process taking place in the calorimeter (Freyer and Lewis, 2008). The molar ratio at the center of binding gives stoichiometry of the protein and ligand and  $K_d$  can be calculated from the model.

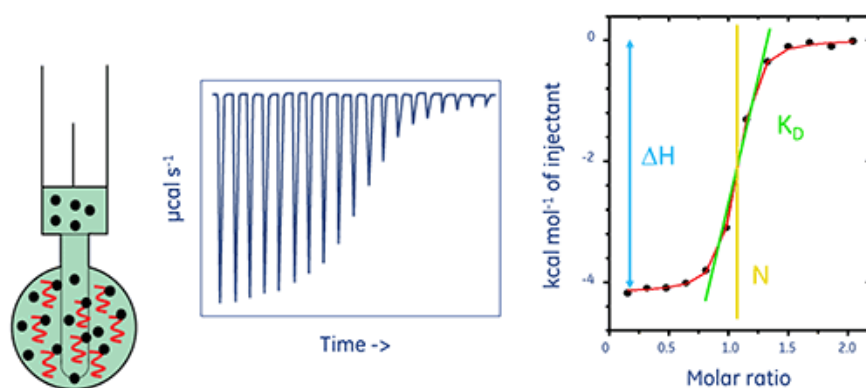
$$\theta_j = \frac{[L]K_j}{1 + [L]K_j}$$
$$Lt = [L] + Pt \sum_{j=1}^k (n_j \theta_j)$$



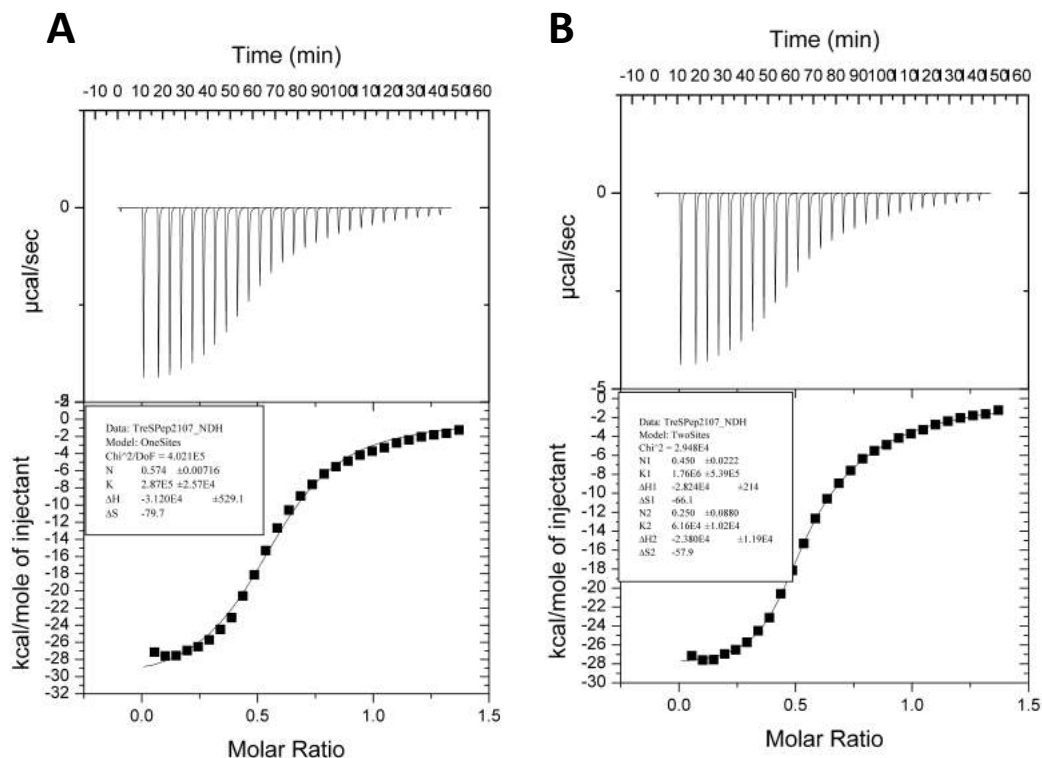
$\Theta_j$  is the fraction of site  $j$  occupied by ligand,  $[L]$  is the free ligand concentration,  $K_j$  is the binding constant of process  $j$ ,  $L_t$  is the total ligand concentration,  $P_t$  is the total macromolecule concentration and  $n_j$  is the total stoichiometric ratio for process  $j$ .

In order to determine the affinity between *MsmTreS* and *Pep2*, I measured the binding between these two proteins using ITC. *MsmPep2* (initial concentration of 500  $\mu$ M) was loaded into the syringe and titrated into the *MsmTreS* solution in the sample cell (initial concentration of 75  $\mu$ M) (Figure 4-6). The experiment was performed in Na-phosphate buffer and three different pH values, 6.5, 7.5 and 8.5 at 25 °C. Following each titration heat was released and as *MsmTreS* became saturated the amount of heat released decreased, suggesting that *MsmTreS* was interacting with *MsmPep2*. If the heat change was simply due to reagent mixing, the amount of heat change should have remained constant throughout the experiment. The area of each peak is integrated and plotted against molar ratio of *MsmPep2* to *MsmTreS*. One set of binding site or two sets of binding sites models were fitted very well into the data points (using equations 4-1 and 4-2) and binding parameters, the stoichiometry of *MsmTreS* and *Pep2* ( $n$ ), association constant ( $K_a$ ), the enthalpy change ( $\Delta H$ ) and the entropy change ( $\Delta S$ ) were determined from theoretical fit (Figure 4-5 and Table 4-2). The strongest affinity was measured at pH 6.5 with a  $K_d$  of 3.5  $\mu$ M. By increasing the pH the affinity of binding decreased. Therefore, the lowest affinity was measured at pH 8.5 with a  $K_d$  of 34.8  $\mu$ M. The stoichiometry ( $n$ ) at pH 6.5 was measured as 0.5, which indicates the *MsmTreS* tetramer binds 2 copies of *MsmPep2*. Likewise to affinity, by increasing pH the  $n$  dropped to 0.22 at pH 8.5, suggesting only one copy of *MsmPep2* binds to the *MsmTreS* tetramer.

The crystal structure of *Msm*TreS:Pep2 suggested that the previously observed rate enhancement of Pep2 in *Mtb*TreS:Pep2 complex might be due to conformational changes in Pep2 upon binding to TreS. In order to explore this we decided to measure the affinity of complex formation in the presence of trehalose and maltose using ITC. The binding measurement between two proteins was performed as explained earlier but 10 mM of each maltose and trehalose was included in the dialysis buffer. Similarly, the  $n$  was measured as 0.5, suggesting the binding of one TreS tetramer to 2 Pep2 monomers. But the affinity was increased to the  $K_d$  of 1.8  $\mu$ M, half of the affinity measured earlier (Table 4-2).



**Figure 4-5.** Schematic representation of the essential components of an ITC system. The stoichiometry of ligand and macromolecule ( $n$ ), association constant ( $K_a$ ), the enthalpy change ( $\Delta H$ ) and the entropy change ( $\Delta S$ ) can be calculated from the theoretical curve. Adopted from Malvern website.



**Figure 4-6.** Isothermal titration calorimetry (ITC) graphs of titration of *MsmTreS* with *MsmPep2*. Prior to experiment both proteins were dialyzed against Na-phosphate buffer at pH values of 6.5, 7.5 and 8.5, supplemented with 300 mM NaCl and 10% (v/v) glycerol. The initial concentration of *MsmTreS* in the reaction chamber was 75  $\mu\text{M}$  that was titrated with 500  $\mu\text{M}$  *MsmPep2*. The experimental data were fitted into the theoretical titration curves using one set of binding sites (A) and two sets of binding sites (B).

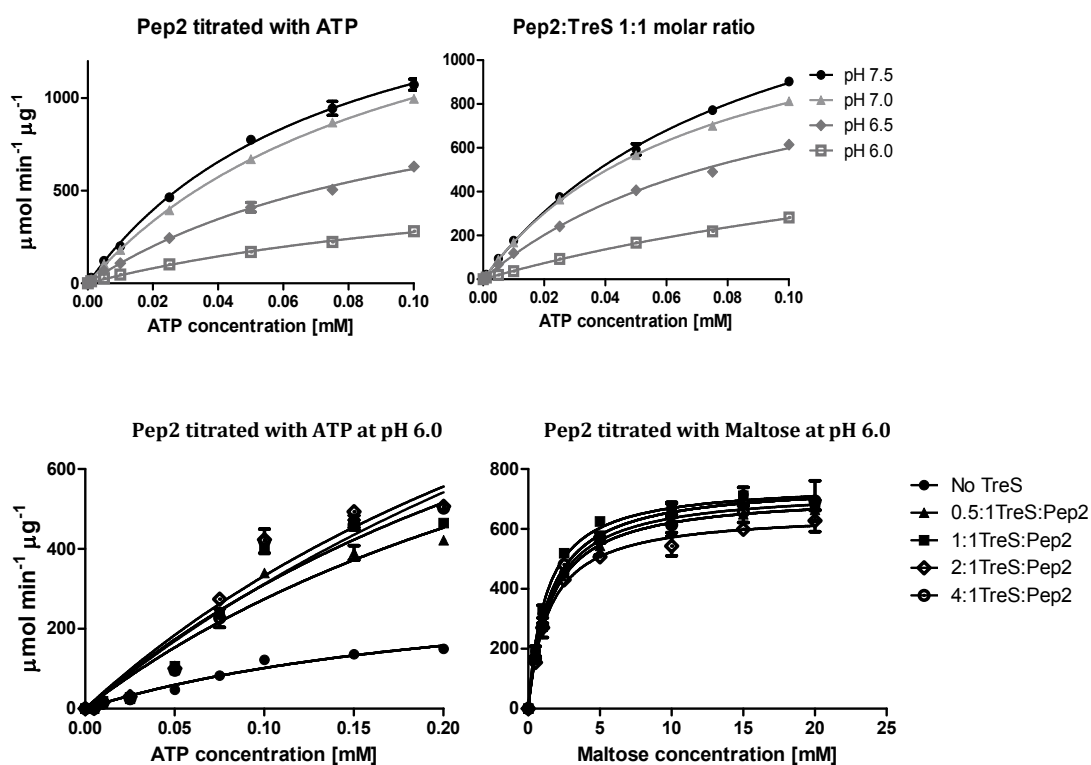
**Table 4-2.** Thermodynamic parameters of *Msm*TreS:Pep2 interaction determined using isothermal titration calorimetry. The data are represented as an average of triplicate experiments.

	Pep2 (wild type)			Pep2 (wild type) in the presence of substrates	Pep2 ( $\Delta 70$ )
<b>pH</b>	6.5	7.5	8.5	6.5	6.5
<b>T (K)</b>	298	298	298	298	298
Single-site binding model					
<b>n</b>	0.574 $\pm$ 0.007	0.459 $\pm$ 0.008	0.22 $\pm$ 0.04	0.536 $\pm$ 0.003	0.681 $\pm$ 0.004
<b>K<sub>a</sub> (10<sup>5</sup> M<sup>-1</sup>)</b>	2.87 $\pm$ 0.26	1.6 $\pm$ 0.12	0.287 $\pm$ 0.03	5.31 $\pm$ 2.82	0.35 $\pm$ 0.08
<b>K<sub>d</sub> (<math>\mu</math>M)</b>	3.48	6.25	34.8	1.88	26.8
<b><math>\Delta H</math> (10<sup>4</sup> kcal mol<sup>-1</sup>)</b>	-3.12	-1.81	-1.96	-3.39	
<b><math>\Delta S</math> (kcal K<sup>-1</sup> mol<sup>-1</sup>)</b>	-79.7	-36.7	-45.9	-87.7	
Two-site binding model					
<b>K<sub>a1</sub> (10<sup>5</sup> M<sup>-1</sup>)</b>	19.7 $\pm$ 3.2				
<b>K<sub>a2</sub> (10<sup>5</sup> M<sup>-1</sup>)</b>	0.55 $\pm$ 0.08				
<b>K<sub>d1</sub>, K<sub>d2</sub> (<math>\mu</math>M)</b>	0.5, 18				
<b>n<sub>1</sub>, n<sub>2</sub></b>	0.445 $\pm$ 0.004 0.457 $\pm$ 0.055				

#### 4.7 Complex formation increases the activity of *MsmPep2* in acidic pH

In order to test the effects of complex formation on catalytic activity of *MsmPep2* activity, a continuous enzyme assay was developed based on this fact that phosphorylation of maltose to maltose 1-phosphate is proportional to the conversion of ATP to ADP (Figure 2-1). The conversion of ATP to ADP was enzymatically coupled to the oxidation of NADH to NAD<sup>+</sup>. Varying ATP at fixed *MsmPep2* concentration (0.02  $\mu\text{M}$ ) at different pH values, in the absence of *MsmTreS*, resulted in Michaelis-Menton type kinetics (Figure 4-7 and Table 4-3). The assay indicated that the activity of *MsmPep2* is pH dependent, as lowering the pH from 7.5 to 6.0 dramatically reduced the activity of *MsmPep2* ( $V_{\text{max}}$  was measured at pH 7.5 and 6.0 as 1901 and 687  $\mu\text{mol min}^{-1} \mu\text{g}^{-1}$ , respectively) (Table 4-3). Then, I probed the activity of *MsmPep2* in the presence of *MsmTreS* in 1:1 molar ratios (Pep2: TreS). Adding *MsmTreS* to *MsmPep2* increased the  $V_{\text{max}}$  of latter at pH 6.0 from 687 to 912  $\mu\text{mol min}^{-1} \mu\text{g}^{-1}$ . Consistent with ITC results, which measured the highest and lowest affinity at pH 6.5 and 8.5 with a  $K_d$  of 3.5 and 35  $\mu\text{M}$ , respectively; we expected to see this rate enhancement only at acidic pH. As the affinity drops at basic pH, adding *MsmTreS* may not have a distinct effect on the activity of *MsmPep2*. To confirm this rate enhancement at pH 6.0, I explored the activity of Pep2 in the presence of four different concentrations of TreS at either fixed maltose or ATP concentrations (Figure 4-7 and Table 4-3). At fixed maltose concentration, adding TreS to Pep2 in 0.5:1, 1:1, 2:1 and 4:1, respectively, increased the activity of Pep2 ( $V_{\text{max}}$ ) from 352  $\mu\text{mol min}^{-1} \mu\text{g}^{-1}$  (in the absence of TreS) to 2040  $\mu\text{mol min}^{-1} \mu\text{g}^{-1}$  (in the presence of four fold TreS concentration). However, at fixed ATP concentration, adding TreS didn't

affect the activity of Pep2,  $754 \mu\text{mol min}^{-1} \mu\text{g}^{-1}$  (in the absence of TreS) to  $733 \mu\text{mol min}^{-1} \mu\text{g}^{-1}$  (in the presence of four fold TreS concentration). It is evident from Michaelis-Menton plots that in several cases the concentration of substrate is not enough to saturate the enzyme and the as a result the calculated  $K_M$  value is more than the highest concentration of substrate titrated to the assay (Figure 4-7). Therefore, it is necessary to repeat those activity assays with higher concentration of substrates to reach closer to equilibrium point and to be able to calculate  $K_M$  values with higher confidence of accuracy.



**Figure 4-7.** Michaelis-Menton kinetics of *MsmPep2* activity under different pH conditions, in absence and presence of *MsmTreS*, and measured by an enzyme-coupled assay (for details see section 2.7). Error bars represent standard deviation.

**Table 4-3.** Michaelis-Menton parameters of *MsmPep2*-catalyzed phosphorylation of maltose in absence and presence of *MsmTreS*.

pH	<i>MsmPep2</i>			<i>MsmTreS</i> : <i>Pep2</i> (1:1)		
	$V_{\max}$ ( $\mu\text{mol min}^{-1} \mu\text{g}^{-1}$ )	$K_M$ (mM)	$k_{\text{cat}}$ ( $\text{sec}^{-1}$ )	$V_{\max}$ ( $\mu\text{mol min}^{-1} \cdot \mu\text{g}^{-1}$ )	$k_{\text{cat}}$ ( $\text{sec}^{-1}$ )	$K_M$ (mM)
<b>6.0</b>	687.2	0.15	224.5	912	298	0.2274
<b>6.5</b>	1291	0.11	421.9	1173	383.3	0.09601
<b>7.0</b>	1993	0.01	651.3	1389	453.9	0.07218
<b>7.5</b>	1901	0.08	621.2	1716	560.8	0.09139

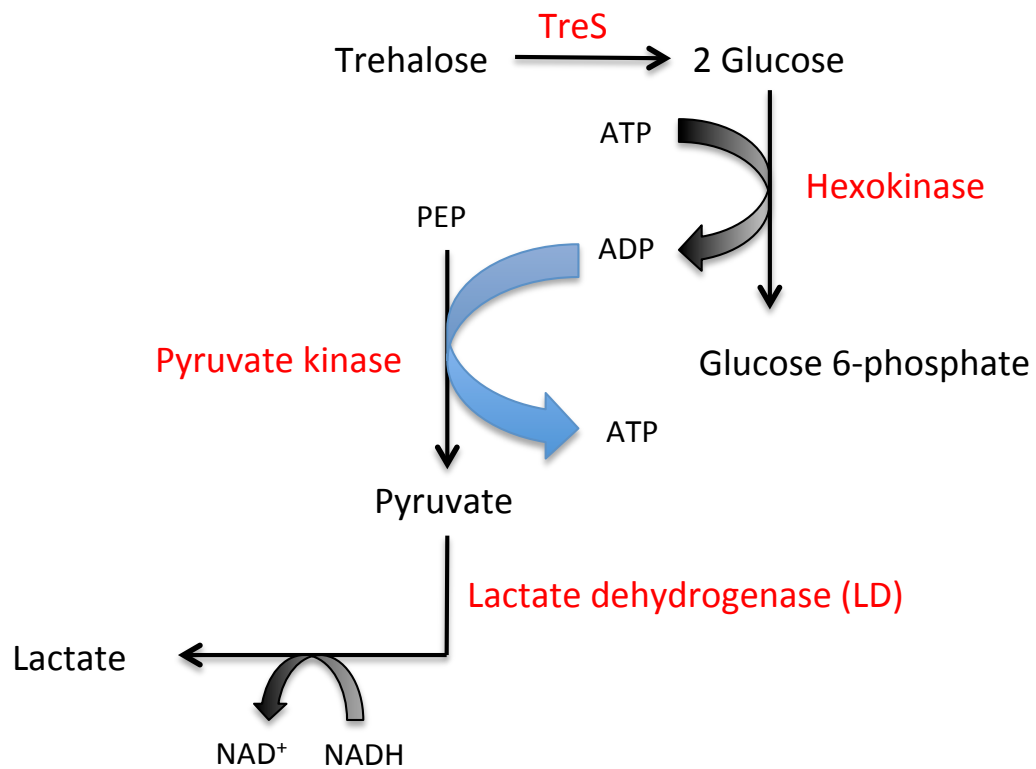
<i>MsmTreS</i> :	Titration with ATP at pH 6.0			Titration with maltose at pH 6.0		
<i>Pep2</i>						
(Molar ratio)						
	$V_{\max}$ ( $\mu\text{mol min}^{-1} \mu\text{g}^{-1}$ )	$K_M$ (mM)	$k_{\text{cat}}$ ( $\text{sec}^{-1}$ )	$V_{\max}$ ( $\mu\text{mol min}^{-1} \cdot \mu\text{g}^{-1}$ )	$K_M$ (mM)	$k_{\text{cat}}$ ( $\text{sec}^{-1}$ )
<b>No TreS</b>	352.1	0.2456	115	754.7	1.511	246.6
<b>0.5 : 1</b>	1306	0.3759	426.8	717.6	1.562	234.3
<b>1 : 1</b>	1515	0.3874	495	752.9	1.240	246
<b>2 : 1</b>	1697	0.4100	554.6	655.3	1.457	214.2
<b>4 : 1</b>	2040	0.5528	666.6	733.4	1.534	239.6

#### **4.8 *MsmTreS* forms a complex with inactive mutant *MsmPep2*-K145A**

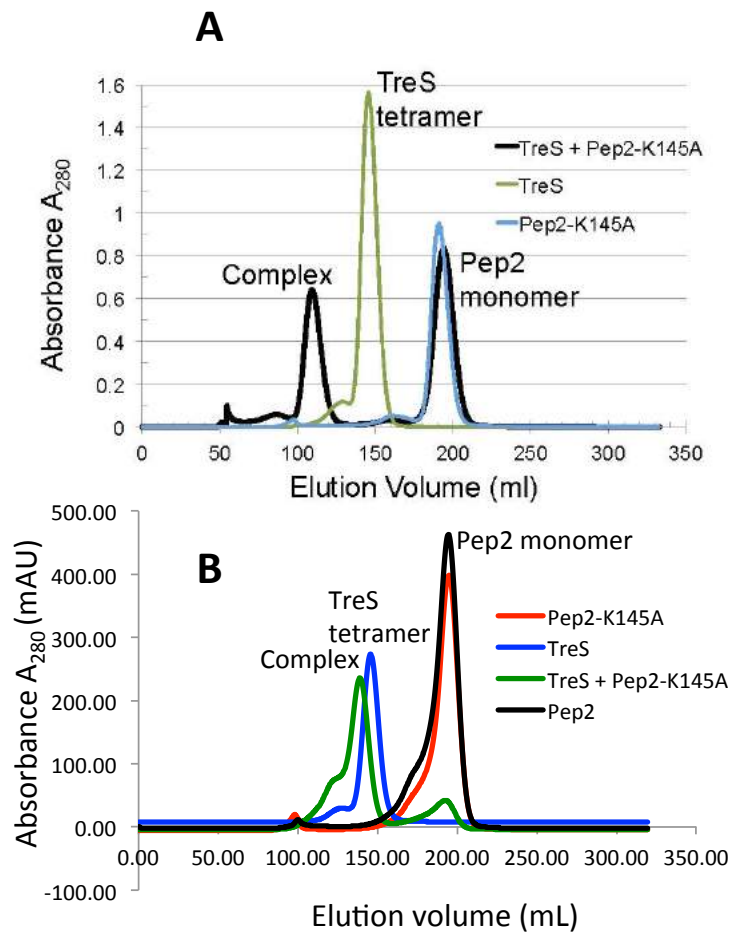
In order to test the effects of complex formation on *MsmTreS* activity, Roy and colleagues measured the enzymatic activity of *MsmTreS* by monitoring the conversion of trehalose to glucose (the hydrolysis side reaction) coupled to the oxidation of NADH to NAD<sup>+</sup> in the presence of hexokinase, pyruvate kinase and lactate dehydrogenase (Figure 4-8) (Roy et al., 2013). Since the coupling reactions involved conversion of ATP to ADP, it was necessary to prevent the consumption of ATP by Pep2 and subsequently phosphorylation of maltose to maltose 1-phosphate. Therefore, they generated an inactive mutant of *MsmPep2*, K145A, designed to abrogate ATP binding to Pep2. Using size exclusion chromatography Roy et al. showed that inactive mutant was still able to form a complex with *MsmTreS*, though, surprisingly; the complex eluted at very low elution volumes very close to the void volume (Figure 4-9 A).

In order to confirm this complex formation and determine its elution volume, both proteins were mixed together in 1:1 molar ratio (75  $\mu$ M each) and run at a pre-equilibrated Sephacryl S300 column (GE Healthcare) as explained in section 2.4. Equilibration was carried out using 50 mM Na-phosphate pH 6.5, 300 mM NaCl and 10% (v/v) glycerol. Two peaks appeared on the chromatogram at 137 and 193 mL, corresponding to the elution volumes of TreS:Pep2 complex and Pep2, respectively, indicating that Pep2-K145A mutant is still able to form a complex with *MsmTreS* and elutes at the same elution volume, contrary to Roy et al. findings (Figure 4-9 B).





**Figure 4-8.** Measuring the enzymatic activity of TreS using an enzymatically coupled reaction between trehalose to glucose conversion and NADH to NAD<sup>+</sup> oxidation in the presence of hexokinase, pyruvate kinase and lactate dehydrogenase. PEP: phosphoenolpyruvate.



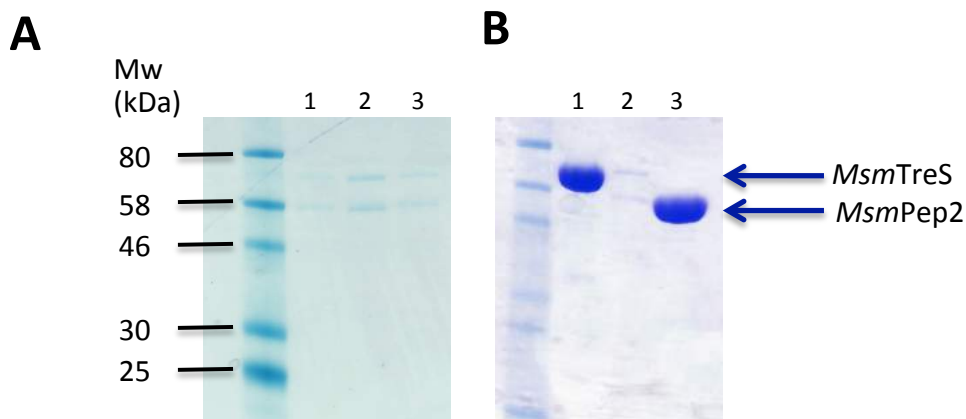
**Figure 4-9.** Exploring complex formation between *Msm*TreS and mutant Pep2-K145A. The size exclusion profile of complex between TreS:Pep2-K145A from (A) Roy et al, 2013 paper and (B) in a second study in order to validate the Roy et al. data. This confirmed that *Msm*TreS and mutant Pep2-K145A form a complex together, though, complex elutes at higher elution volume (138 mL).

#### 4.9 Crystallization of the *Msm*TreS:Pep2 complex

In order to co-crystallize *Msm*TreS and Pep2 together, both proteins were purified separately and mixed together in three different molar ratios of *Msm*Pep2 to TreS (1:1, 2:1 and 1:2) respectively, just before setting up crystal trays. Initial crystals appeared in 6 different conditions (Table 4-4). SDS-PAGE analysis of the four crystals selected from different conditions demonstrated that these crystals are formed from both TreS and Pep2 proteins (Figure 4-10). However, most of these crystals did not diffract using the in-house X-ray source. Only crystals grew in 0.2 M MgCl<sub>2</sub>, 0.1 M Tris pH 7.0 and 10% w/v PEG 8K (JCSG-plus B8 condition) diffracted (~8.5 Å). Using extensive screening and a very narrow pH range (pH 6.7 to 7.0) I was able to generate higher diffraction quality crystals. The structure was phased by molecular replacement and refined to a resolution of 3.6 Å (Table S9-4).

**Table 4-4.** List of conditions generating crystals of *Msm*TreS:Pep2 complex. <sup>1</sup> NPS: NaNO<sub>3</sub>, Na<sub>2</sub>HPO<sub>4</sub>, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; <sup>2</sup> Ethylen glycol, PEG 8K; <sup>3</sup> NaF, NaBr, NaI; <sup>4</sup> D-glucose, D-mannose, D-galactose, L-fucose, D-xylose, *N*-acetyl-D-glucosamine.

Salt	Buffer	pH	Precipitant
0.2 M MgCl <sub>2</sub>	0.1 M Tris	7.0	10% w/v PEG 8K
0.3 M C <sub>2</sub> H <sub>2</sub> MgO <sub>4</sub>	0.1 M Bis-Tris	5.5	
0.09 M NPS <sup>1</sup>	0.1 M imidazole-MES (acid)	6.5	30.0% EDO_P8K <sup>2</sup>
0.09 M Halogens <sup>3</sup>	0.1 M Tris (base)-Bicine	8.5	30.0% EDO_P8K <sup>2</sup>
0.12 M Ethylene glycols	0.1 M Tris (base)-Bicine	8.5	30.0% EDO_P8K <sup>2</sup>
0.12 M Monosaccharides <sup>4</sup>	0.1 M Tris (base)-Bicine	8.5	30.0% EDO_P8K <sup>2</sup>



**Figure 4-10.** SDS-PAGE analysis of crystals harvested from *MsmTreS*:Pep2 crystal trays. Crystals initially were washed in cryoprotectant solution, boiled in sample buffer for 5 minutes and ran on the SDS-PAGE. Crystals in (A) harvested from Lanes (1) 0.1 M Morpheus buffer system 3 pH 8.5, 0.09 M Halogens and 30.0% EDO\_P8K, (2) 0.1 M Morpheus buffer system 3 pH 8.5, 0.12 M monosaccharides and 30.0% EDO\_P8K, (3) 0.1 M Bis-Tris (pH 5.5) and 0.3 M  $C_2H_2MgO_4$  conditions. Lanes in B represent (1) TreS (2) crystals obtained from 0.2 M  $MgCl_2$ , 0.1 M Tris pH 7.0 and 10% w/v PEG 8K (JCSG-plus B8 condition) and (3) Pep2. These gels confirm that crystals grew in these conditions contain both *MsmTreS* and Pep2 proteins. Morpheus buffer system 3 is made of a mixture of imidazole and MES in a way to give the appropriate pH. Halogens: NaF, NaBr, NaI. Monosaccharides: D-glucose, D-mannose, D-galactose, L-fucose, D-xylose, *N*-acetyl-D-glucosamine. EDO\_P8K: Ethylen glycol, PEG 8K.

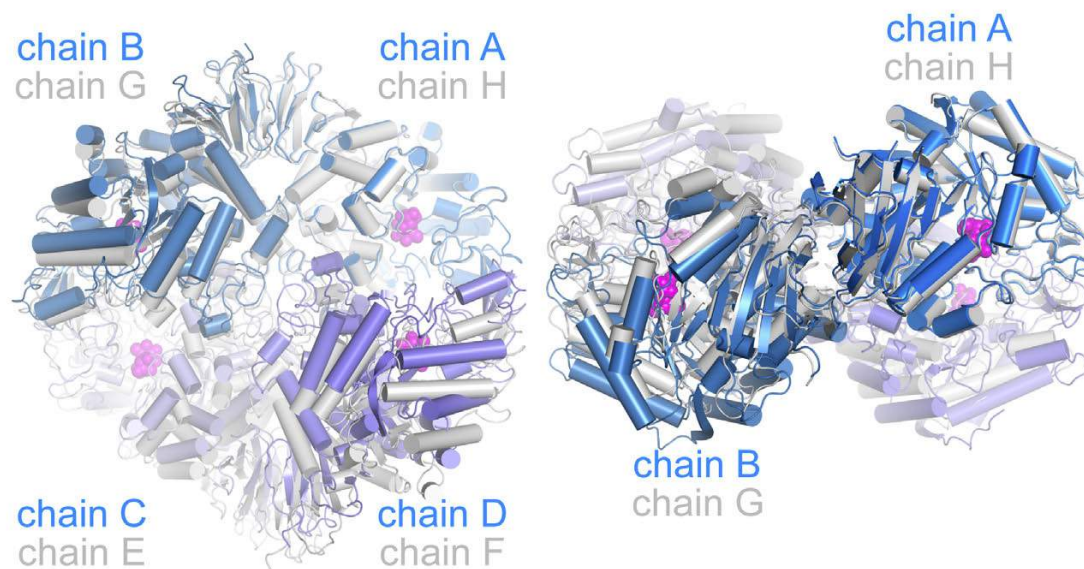
#### 4.10 X-ray crystal structure of the *Msm*TreS:Pep2 complex

The structure of *Msm*TreS:Pep2 complex was phased by molecular replacement, and refined to a resolution of 3.6 Å (Table 9.3). There are two copies of the *Msm*TreS:Pep2 complex in the asymmetric unit of the crystal lattice, which display differences in terms of subunit orientation and level of disorder. Each copy of the complex is composed of 4 TreS subunits bound to 4 Pep2 monomers (4 + 4 assembly). Except Ca<sup>2+</sup> sites in TreS, no other ligand was observed in the crystal structure of TreS:Pep2 complex. The backbone of the *Msm*TreS tetramer of the first complex of the asymmetric unit (chains A – D) superimposes well with the tetramer of the stand-alone TreS (PDB entry 3ZO9) with minimal root mean square deviation (RMSD) of 0.45 Å for 2243 aligned Cα positions (Figure 4-11). In contrast, the TreS tetramer from the second complex (chains E-H) do not superimpose well with either *Msm*TreS from the first complex or the stand-alone *Msm*TreS structure. However, if we superimpose chain H of complex 2 to chain A of complex 1, the three other subunits can be mapped accordingly (i.e. G → B, F → D, E → C), leading to an RMSD of 3.4 Å for 2175 aligned Cα positions (Figure 4-11). This mismatch can be explained as a result of rotation of the *Msm*TreS protomers against each other.

There are eight Pep2 protomers in the asymmetric unit, which show different levels of disorder. In first complex, three Pep2 chains I – L (Figure 4-12B) are completely folded but in fourth Pep2 subunit (chain K) residues 1 to 86 are largely disordered. In second

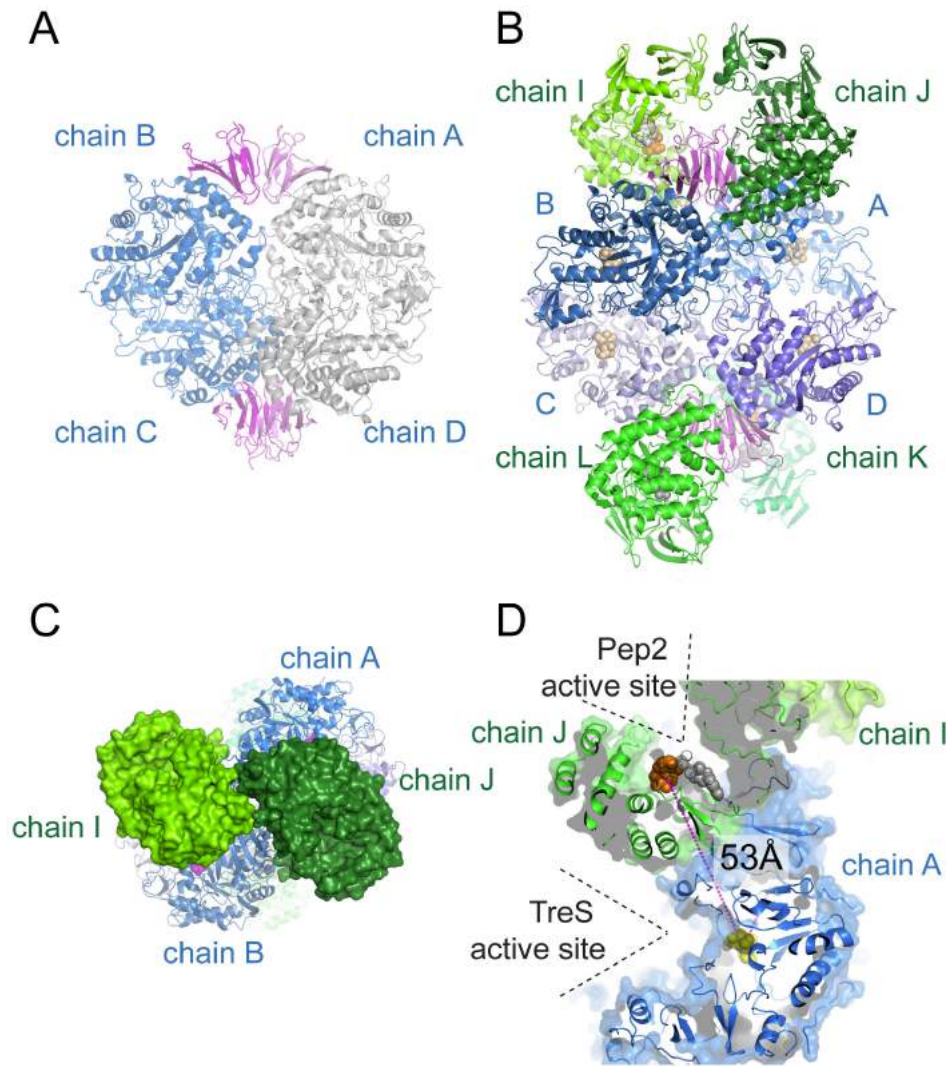
complex (Pep2 chains M – P) only one copy of Pep2 (chain M) could be traced fully, while in other three chains the disorder extends to about residue 195 in N-terminus.

The core of *Msm*TreS:Pep2 complex is previously described diamond shaped TreS tetramer (Figure 4-12A), which can be described as a dimer of dimers (Caner et al., 2013 and Roy et al., 2013). The  $\beta$ -sandwich domains, composed of seven-stranded anti-parallel  $\beta$ -sheets, (see section 1-7-1 for more information) of each dimer join together and form the top and bottom apex of the diamond (Figure 4-12A). Pairs of Pep2 monomers bind to each apex of the diamond in a clamp-like assembly (Figure 4-12 B), demonstrating an assembly of 4 + 4 complex. It might appear that Pep2 form dimers on TreS core but considering this fact that the contact surface between Pep2 monomers is small compared to than the contact surface between TreS and Pep2 implies that Pep2 binds as a monomer (Figure 4-12 C). More importantly, in those copies of Pep2 that contain disordered N-termini, the Pep2:Pep2 interface is absent.



**Figure 4-11.** The superimposition of the TreS tetramer of complex 1 (chains A – D, in hues of blue) to the TreS tetramer of the complex 2 (chains E- H, in gray). The superposition was performed such that chain A matches chain H, therefore, there is only minimal displacement between chains A and H. Spheres in the magenta indicate the location of the active sites.



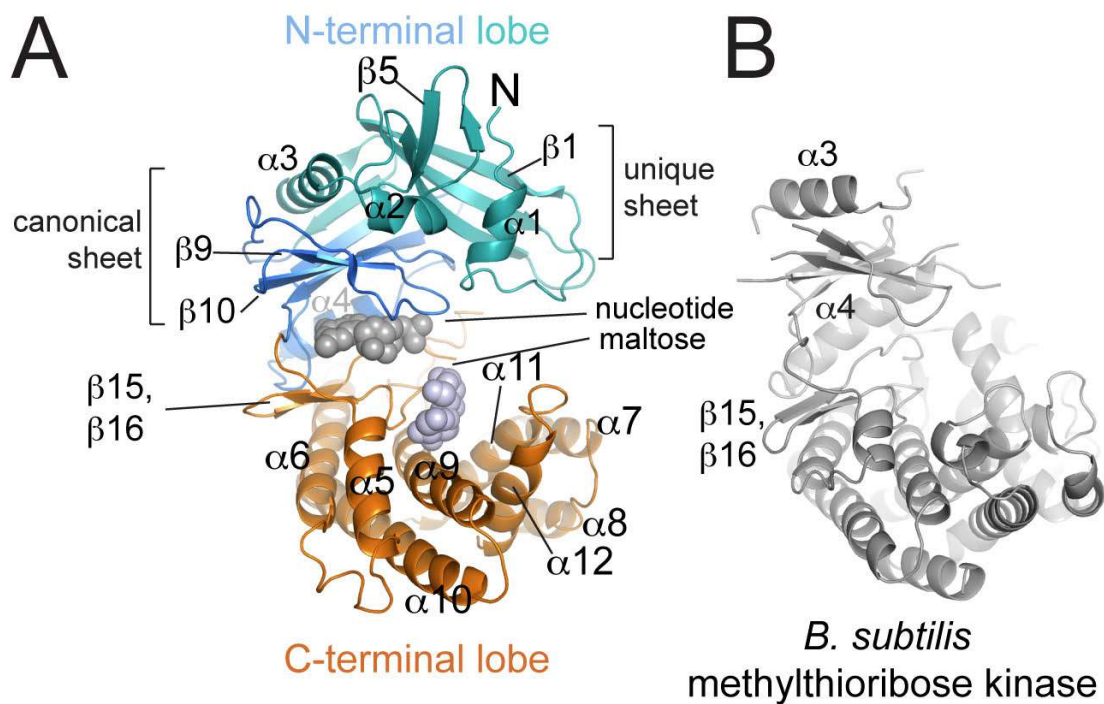


**Figure 4-12.** Crystal structure of the *MsmTreS*:Pep2 complex. (A) Crystal structure of *MsmTreS* tetramer that can be described as a dimer of dimer. Magenta ribbons show the interaction between C-terminal  $\beta$ -sandwich domains of the TreS subunits. (B) Assembly of the *MsmTreS*: Pep2 complex with the TreS tetramer at the core and pairs of Pep2 on each end of the diamond shaped TreS. Spheres indicate the location of active sites derived by superposition with homologous structures. (C) Contact surface between TreS and Pep2 in top view, illustrating that the contact surface between Pep2 monomers is smaller than the contact surface with TreS. (D) Illustration of location of TreS and Pep2 active sites in proximity of each other. The dashed line that connects TreS active site (chain A) and Pep2 active site (chain J) traverses through the hydrophobic core of Pep2, indicating there is no substrate channel between two protein's active sites.

#### 4.11 The structure of *MsmPep2*

During the course of this project the crystal structure of Pep2 from *M. tuberculosis* (Li et al., 2014) and *M. vanbaalenii* (Fraga et al., 2015) were published. The structural model of *MsmPep2* in the current complex is completely defined by electron density except for two residues at the N-terminus and 4-5 residues at the C-terminus. We were able to dock the amino acid sequence of Pep2 onto the backbone structure by using the coordinates from Pep2 structures of *M. tuberculosis* and *M. vanbaalenii*. However, in some positions where the corresponding density was missing, we were not able to dock the side chains.

Pep2 is a member of Pfam family PF01636, a database that classifies proteins according to their sequence homology (Finn et al., 2014). The main characteristic of the members of this family is having a two-lobe architecture, N-terminal and C-terminal lobe (Figure 4-13A). Comparing the structure of *MsmPep2* with closest structural neighbor (apart from the mycobacterial orthologues) *Bacillus subtilis* methylribose kinase (PDB code 2PUL, RMSD 3.9 Å, 262 aligned Cα) demonstrates a good match between the C-terminal lobe and parts of the N-terminal lobe of Pep2 (Figure 4-13). However, some parts of N-terminus, β1- β3, appear to be unique to mycobacterial Pep2 and are missing from homologue structures identified by DALI. In contrast, the consecutive 7-stranded β-sheet (β6-β12) is conserved. Here, we refer to it as the canonical β-sheet (Figure 4-13A).



**Figure 4-13.** Structure of *MsmPep2*. (A) Cartoon representation of *MsmPep2* in monomeric state with conserved and unique regions in the N-terminal lobe shown in blue and teal, respectively. Spheres depict the active sites of nucleotide and maltose obtained by superposition with GDP-bound aminoglycoside 2-phosphotransferase IIIa (PDB code 3TDW) and maltose-bound *MtbPep2* (PDB code 4O7P), respectively (B) Cartoon representation of methylribose kinase from *Bacillus subtilis* (PDB code 2PUL), the closest structural neighbor of *Pep2*.

#### 4.12 Truncating *MsmPep2*

Prior to experimental structures of Pep2 being available, predictions of the tertiary structure of this enzyme using the HHpred threading server ([toolkit.tuebingen.mpg.de/hhpred](http://toolkit.tuebingen.mpg.de/hhpred)) failed to predict the folding of the first 100 amino acids of the Pep2 sequence, while high-confidence models for the remainder of the sequence were obtained. This observation suggested that either the N-terminal sequence encodes a unique structure or this sequence is intrinsically disordered. In order to test the hypothesis of intrinsic disorder, we generated seven mutants and truncated the first 70 amino acids from the N-terminus of *MsmPep2*. The first mutant was generated by removing the first 10 amino acids and the proceeding mutants lost additional 10 amino acids in each time. Probing the oligomeric state of these truncations using size exclusion chromatography showed that they are highly prone to aggregation but the tendency to aggregate do not systematically increase with the extent of truncation (Figure 4-14).

Following solving the crystal structure of *MtbPep2* (Li et al., 2014) and then *MsmPep2* it became evident that the N-terminus of Pep2 is well folded but contains a unique 3-stranded  $\beta$ -sheet that is exclusive only to mycobacterial species. In order to explore the functional role of this unique  $\beta$ -sheet I chose Pep2( $\Delta$ 70) for further studies as it lost this region. In order to improve the solubility of this mutant the concentration of NaCl was increased from 200 mM to 500 mM. This improved the solubility of  $\Delta$ 70 mutant to some extent but it did not prevent the aggregation of this protein completely. However, size exclusion chromatography is quite useful in separating the aggregated from the non-aggregated protein. Pep2( $\Delta$ 70) was eluted from the SEC column as two major peaks at

100 mL and 188 mL representing aggregation and monomeric state of this protein, respectively. Although Pep2( $\Delta$ 70) is 70 amino acids smaller than wild type Pep2, it elutes almost at the same elution volume as *Msm*Pep2, suggesting that the shape of this mutant has been changed.

In order to test whether Pep2( $\Delta$ 70) is still able to bind *Msm*TreS, both proteins were mixed together in 1:1 molar ratio (50  $\mu$ M each) and loaded onto the pre-equilibrated HiPrep Sephacryl S300 HR (GE Healthcare) column (Figure 4-14). The complex peak was eluted at higher elution volumes (144 mL) indicating that the mutant Pep2 is still able to bind to its partner *Msm*TreS but the interaction is weaker. In addition, the catalytic activity of Pep2( $\Delta$ 70) was measured against the wild type *Msm*Pep2 using the continuous enzyme assay explained earlier in section 4-7. Varying ATP at a fixed initial maltose concentration or varying maltose at a fixed ATP concentration resulted in Michaelis-Menton type kinetics in both cases (Figure 4-15). Fixing maltose concentration at 20 mM and increasing ATP concentration up to 0.4 mM showed a reduction in  $V_{\max}$  from 2513 ( $\pm$  205) (for wild type Pep2) to 242 ( $\pm$  41)  $\mu$ mol min<sup>-1</sup> mg<sup>-1</sup> (for Pep2 $\Delta$ 70). Similarly, fixing ATP concentration at 0.5 mM and increasing the concentration of maltose up to 25 mM, indicated  $\sim$  36-folds reduction in the activity of Pep2( $\Delta$ 70) (Table 4-5). This data suggests that the unique  $\beta$ -sheet region of Pep2 is essential for structural stability of the N-terminal lobe and perturbing this region can affect nucleotide binding and subsequently activity.

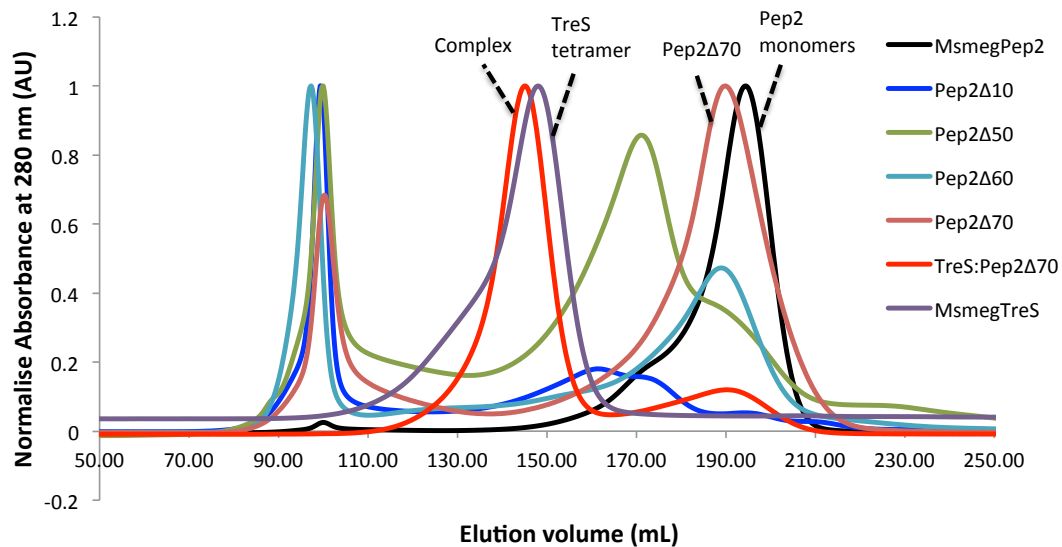
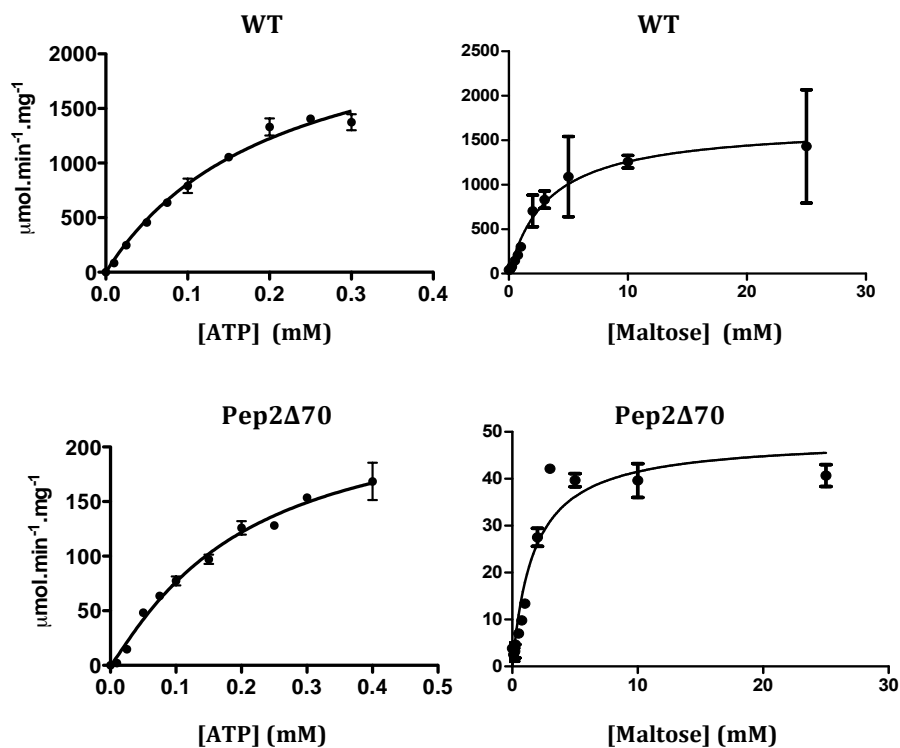


Figure 4-14. The size exclusion profile of *MsmPep2* ( $\Delta$ 10-  $\Delta$ 70) and evidence of complex formation between mutant *Pep2* ( $\Delta$ 70) and *MsmTreS*. Individual *TreS*, *Pep2* ( $\Delta$ 70) and the complex containing both proteins were eluted from column at 147, 193 and 144 mL, respectively.



**Figure 4-15.** Michaelis-Menton kinetics of *MsmPep2* wild type and *MsmPep2*( $\Delta$ 70) activity, measured by an enzyme-coupled assay (for details see section 2.7). Error bars are representing standard deviation.

**Table 4-5.** Michaelis-Menton kinetics of *MsmPep2* and Pep2( $\Delta$ 70) measured by an enzyme-coupled assay.

Parameters	Pep2 titrated with ATP	Pep2 $\Delta$ 70 titrated with ATP	Pep2 titrated with maltose	Pep2 $\Delta$ 70 titrated with maltose
$V_{\max}$ ( $\mu\text{mol}\cdot\text{min}^{-1}\cdot\mu\text{g}^{-1}$ )	$2513 \pm 205$	$242 \pm 41$	$1679 \pm 162$	$49 \pm 3.2$
$K_M$ (mM)	$0.2 \pm 0.03$	$0.16 \pm 0.1$	$3.3 \pm 0.87$	$1.7 \pm 0.36$
$k_{\text{cat}}$ ( $\text{sec}^{-1}$ )	821.2	79	548.7	16

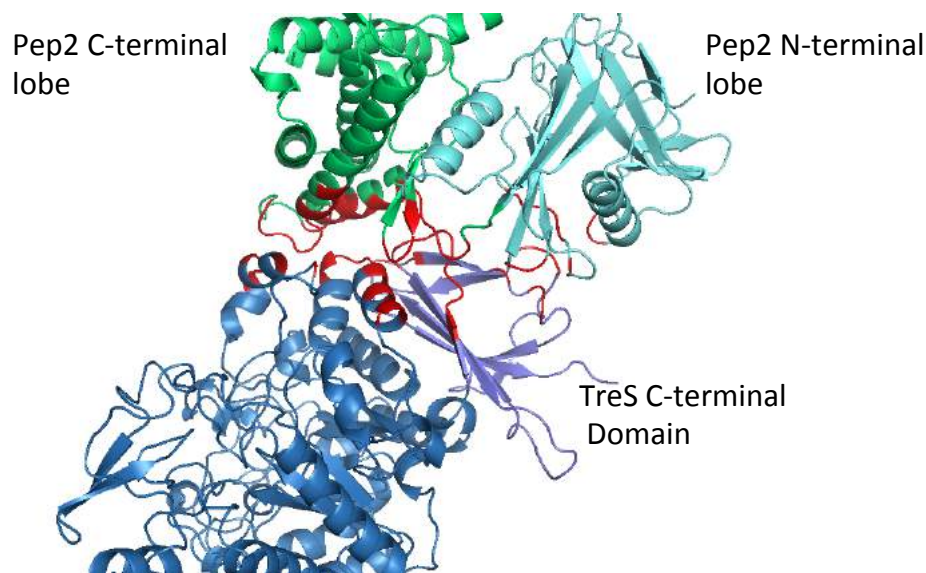
#### 4.13 TreS and Pep2 active sites

Roy and colleagues showed that complex formation between TreS and Pep2 in *M. tuberculosis* increases the activity of Pep2 in a dose-dependent manner (Roy et al., 2013). Therefore, we thought there might be a tunnel between two enzyme active sites, which could potentially facilitate substrate channeling. However, if we connect the active site of TreS to the active site of nearest Pep2 using a direct line we observe that the line passes through the hydrophobic core of complex and there is no substrate channel between the active sites of two proteins (Figure 4-12D). The linear distance between *Msm*TreS active site and the nearest *MsmPep2* active site (chain J) is about 53 Å and the next nearest *MsmPep2* active site (chain I) is 63 Å away. This fact that *Msm*TreS and *MsmPep2* active sites do not face each other and also both active sites are exposed to solvents suggest that maltose, following release from TreS active site, dissolves in solvent and enters the Pep2 active site.

#### 4.14 The TreS-Pep2 contact interface

According to PISA, the contact surface between chain B of TreS and chain I of Pep2 is composed of 37 and 33 amino acids of each subunit, respectively (Figure 4-16 and Table 4-6). However, only a few residues from each subunits, ASN199, SER200, ASP213, ALA216, GLY227, ASP228 and ARG235 (from Pep2) and ARG311, ARG312, ARG315, GLU360 and ARG534 (from TreS) are involved in interaction between two proteins, mainly by forming hydrogen bonds and salt bridges (Table 4-6).





**Figure 4-16.** Contact surface between TreS and Pep2. The interaction between chain I of *Msm*Pep2 and chain B of *Msm*TreS is stabilized by hydrogen bonds and salt bridges. Residues involved in surface interaction between two proteins are colored in red.

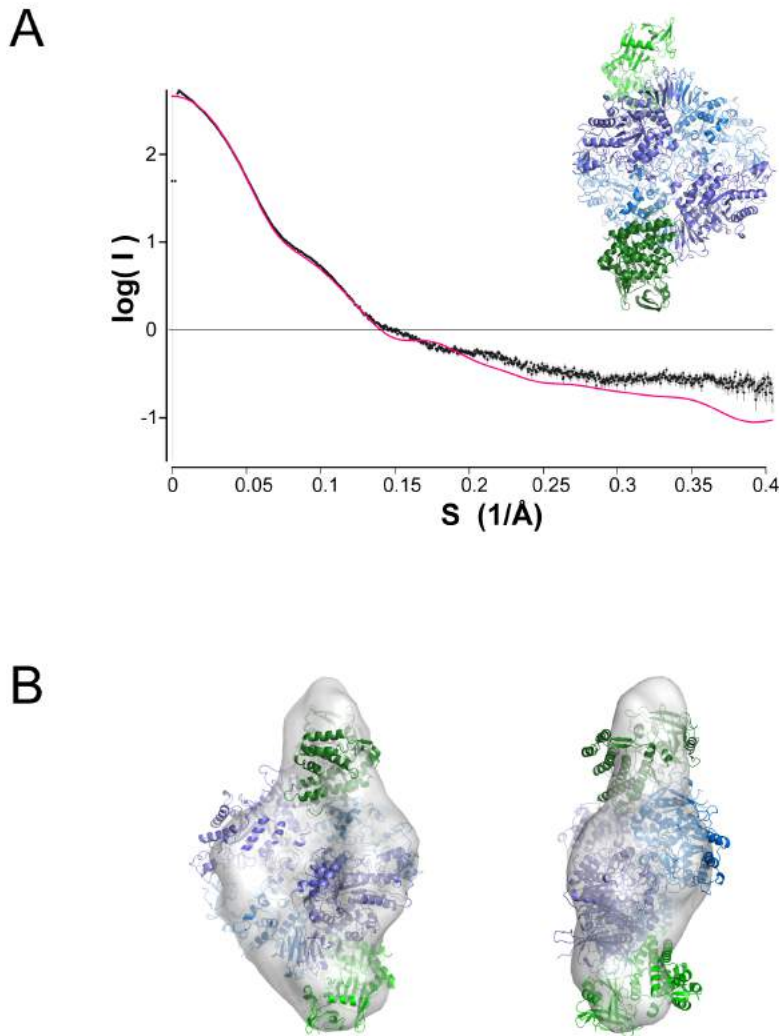
According to Figure 4-4, the complex formation is more likely affected by pH changes at pH 6.5 – 8.0. It has been well established that the charge of amino acids can change in the range of -1 to +1 by pH changes and therefore, it would affect the charge-charge interactions established between charged residues. It was speculated that Histidine by having a side chain with a  $pK_a$  value of 6.1 is more likely responsible for charge-charge interactions between TreS and Pep2. However, analyzing the interface between these two proteins using PISA server represented no Histidine in the binding interface, referring to the role of other forces in governing the interactions between TreS and Pep2 at the interaction interface.

**Table 4-6.** List of TreS and Pep2 residues in binding interface.

Pep2 residues at surface interface	Formed bond	TreS residues at surface interface	Formed bond
ILE65		ARG311	Hydrogen bond
GLU67		ARG312	Hydrogen bond/salt bridge
TYR68		GLU313	
ALA122		SER314	
PRO124		ARG315	Hydrogen bond
ARG125		PHE316	
VAL126		SER319	
PHE127		GLU320	
GLY139		ALA323	
GLU140		GLN324	
ALA198		TYR356	
ASN199	Hydrogen bond	GLU360	Hydrogen bond/salt bridge
SER200	Hydrogen bond	LYS363	
ALA201		ASP364	
GLU202		PRO365	
ASP205		ARG496	
MET206		GLU497	
THR209		LEU498	
SER210		GLY499	
ARG212		GLY500	
ASP213	Hydrogen bond/salt bridge	SER501	
ALA216	Hydrogen bond	ASN502	
GLU217		PRO503	
GLU223		SER504	
GLU224		ARG534	Hydrogen bond/salt bridge
VAL225		PHE535	
GLY226		GLN537	
GLY227	Hydrogen bond	ASN542	
ASP228	Hydrogen bond/salt bridge	GLN544	
GLY231		GLN545	
GLU232		GLN565	
ARG235	Hydrogen bond/salt bridge	LEU566	
ARG312		PRO567	
THR313			
PRO314			
LYS315			
ARG382			

#### 4.15 Small Angle X-ray Scattering

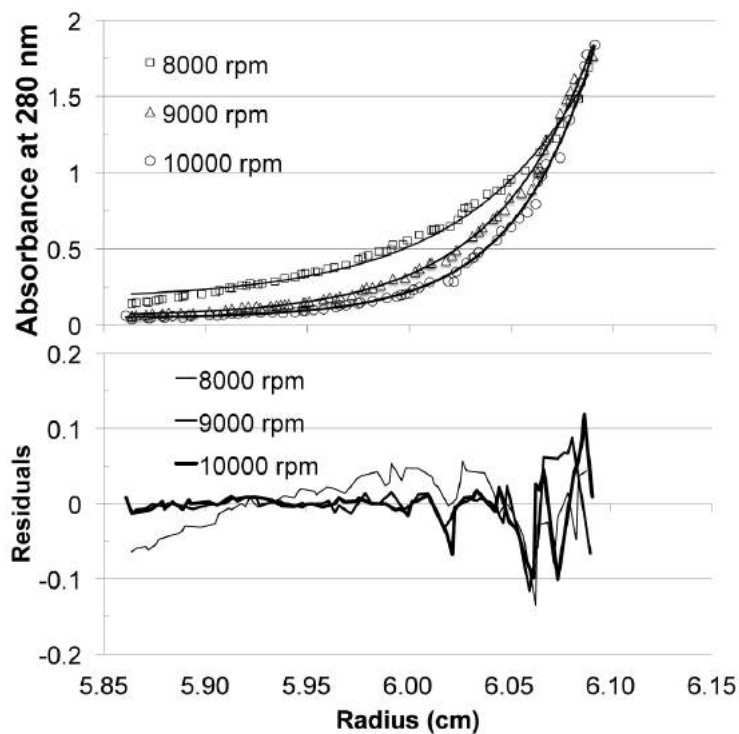
In order to obtain structural information and determine the stoichiometry of TreS:Pep2 complex in the solution state, a small angle X-ray scattering (SAXS) was performed (Jacques and Trewhella, 2010). We first examined the scattering profiles of individual *Msm*TreS (1.25 to 10 mg/mL) and *Msm*Pep2 (1.25 to 10 mg/mL). Guinier plots of the scattering curves indicated that both proteins are heavily aggregating in the solution with radii of gyration ( $R_g$ ) of 4.51 ( $\pm$  0.03) nm and 2.78 ( $\pm$  0.32) nm, respectively. Then, we examined the solution profiles of *Msm*TreS:Pep2 complex in 1:1 and 4:2 molar ratios of TreS to Pep2. It resulted in a bigger  $R_g$  ( $5.60 \pm 0.01$  nm). Fitting various structural models derived from the crystal structure of the TreS:Pep2 complex to the solution scattering data using CRY SOL (Svergun, Barberato and Koch, 1995), indicated a clear preference for a model consisting of the TreS tetramer bound to two copies of Pep2, attached to apposite apices of the TreS tetramer, respectively (Figure 4-17). Molecular envelope of the complex generated using DAMMIF (Franke and Svergun, 2009) and DAMAVER (Volkov and Svergun, 2003) suggested that the TreS:Pep2 complex in the solution is formed of 4 TreS bound to 2 Pep2 monomers (4+2 configuration) where chains I and L of Pep2 participate.



**Figure 4-17.** Small-angle X-ray scattering (SAXS) analysis of the *MsmTreS:Pep2* complex. (A) Fitting the atomic structure of *MsmTreS:Pep2* complex (4+2) into the experimental scattering curves from SAXS (using CRY SOL program). The magenta color depicts the fitting curve. (B) The ab initio model generated using DAMMIF (Franke and Svergun, 2009) and DAMAVER (Volkov and Svergun, 2003) and converted into a pseudo-density map using CCP4 software suite. See materials and methods for more information. It indicates that the model fits one *TreS* tetramer and 2 *Pep2* monomers.

#### 4.16 Stoichiometry of the *Msm*TreS:Pep2 complex

Molecular mass and stoichiometry of the complex between *Msm*TreS and *Msm*Pep2 were determined using analytical ultracentrifugation (AUC) in sedimentation equilibrium mode (Cole et al., 2008) (Figure 4-18). Mixtures of *Msm*TreS: Pep2 at molar ratios of 1:1 and 2:1 were analyzed at 3 different concentrations and 4 different rotation speeds. It had been previously established that *Mtb*TreS forms a tetramer in the solution with a molecular weight of ~280,000 Da (Roy et al., 2013) and was further confirmed during this project using size exclusion chromatography (SEC) and small angle X-ray scattering (SAXS). Knowing the molecular weight of TreS and considering that Pep2 behaves as a monomer in the solution, molecular weights obtained from AUC were analyzed. In *Msm*TreS:Pep2 in molar ratio of 2:1 the molecular weight was calculated as 324,367 Da, which suggests four TreS (monomer) bound to one Pep2 subunit. Mixing TreS and Pep2 in a 1:1 molar ratio gave a molecular weight of 363,763 Da, which indicates a TreS tetramer with two bound monomers of Pep2. When TreS was analyzed on its own, it yielded a molecular mass of 257,019 Da, which demonstrates a tetramer, consistent with previous published data (Roy et al., 2013).



**Figure 4-18.** Sedimentation equilibrium analysis of the *M. smegmatis* TreS:Pep2 complex. A 1:1 molar ratio of *Msm*TreS and Pep2 mixture was analyzed at three different protein concentrations ([TreS]= 3.75, 2.5 and 1.25  $\mu$ M) at three rotation speeds (8,000, 9,000 and 10,000 rpm) and fitted to a single sedimenting species. Data points are shown as open symbols, and solid lines represent the best fit (top panel) and residuals (lower panel), respectively.

## 4.17 Summary

*Msm*TreS and *Msm*Pep2 were successfully overexpressed and purified (Figure 3-1 and 4-1). Both proteins were eluted as single peaks from SEC column at 144 and 188 mL, representing the tetrameric and monomeric states of these proteins in the solution, respectively. Using SEC, it was shown that TreS forms a complex with Pep2 and this complex formation is highly pH dependent, as increasing the pH from 6.0 to 9.0 diminishes the complex formation (Figure 4-3 and Figure 4-4). Isothermal titration calorimetry also, measured the highest and lowest affinities between TreS and Pep2 at pH 6.5 and 8.5, respectively (Table 4-2). Consistent with these data, the highest activity of Pep2 in the presence of TreS was observed only at pH 6.0 (Figure 4-7 and Table 4-3).

Solving the crystal structure of Pep2 in complex with TreS demonstrated a 4 + 4 configuration, in which two pairs of Pep2 monomers are placed on each apex of a diamond shape TreS tetramer (Figure 4-12). Additionally, the Pep2 rate-enhancement is due to conformational changes occurring upon binding to TreS, not substrate channeling as proposed earlier (Figure 7-1). Pep2 is composed of two lobes, N-terminal lobe and C-terminal lobe (Figure 4-13). The N-terminal lobe harbors a unique region ( $\beta 1 - \beta 3$ ), which only can be seen in mycobacteria, and is involved in structural stability of N-terminal lobe, as shown by removing this region (Pep2  $\Delta 70$  mutant) and probing the activity and the interaction of this mutant with TreS, and comparing the results with wild type Pep2.

SAXS and AUC suggested that complex formation between TreS and Pep2 in the solution occurs in a 4 + 2 configuration (Figure 4-17 and 4-18), in contrast to stoichiometry in the crystal lattice.

**5 Crystallization and interaction studies of**  
***Mycobacterium smegmatis* GlgE with**  
**Pep2 and GlgB**

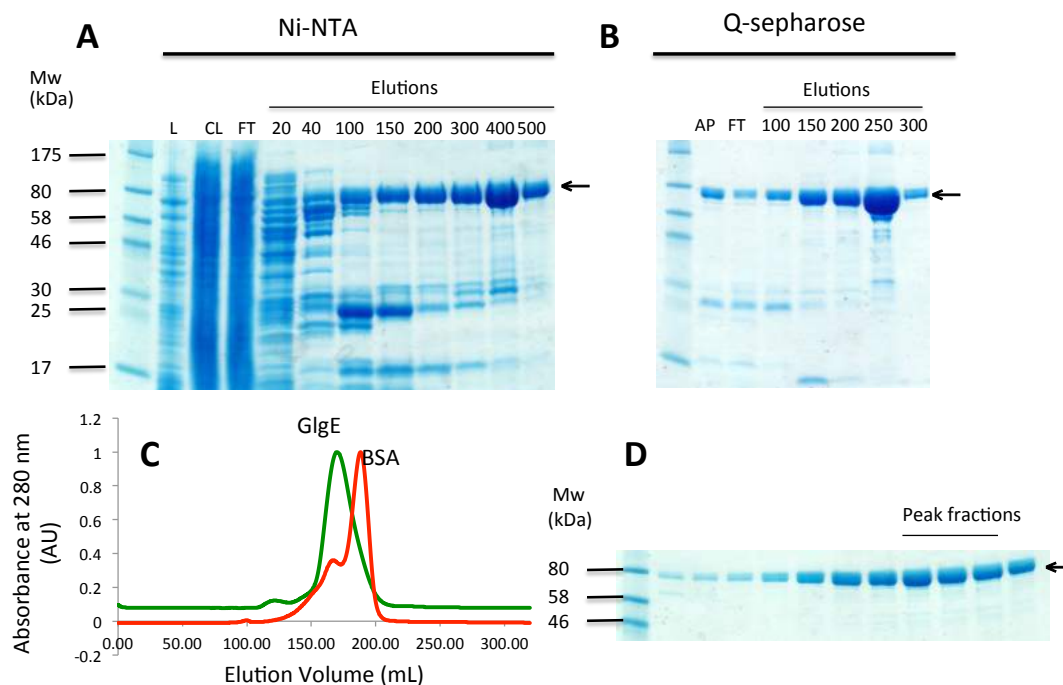


## 5.1 Crystallization of *Mtb*GlgE

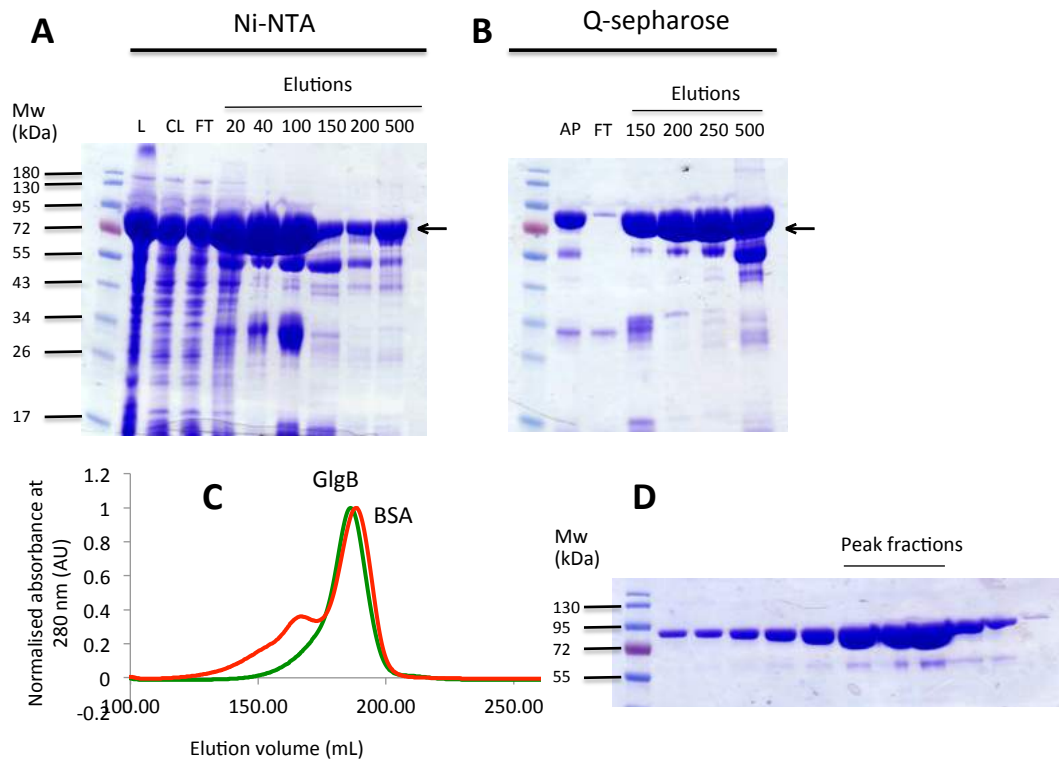
GlgE and GlgB catalyze two consecutive reactions from GlgE pathway, the transfer of maltose from maltose 1-phosphate (M1P) to maltooligosaccharides and introducing branches to the linear glucan at some position 6, respectively. Kalscheuer and colleagues (2010) demonstrated the essentiality of both enzymes for *Mtb* survival. They showed that inactivation of either GlgE or GlgB in *Mtb* leads to hyperaccumulation of M1P and subsequently self-poisoning of bacilli. Both enzymes could be potentially new drug targets, however, the presence of a human ortholog of GlgB makes it less attractive than GlgE. In this study I attempted to crystallize *Mtb*GlgE and explore its interactions with Pep2 and GlgB.

## 5.2 Protein expression and purification of *Mtb*GlgE and *Msm*GlgB

*Mtb*GlgE and *Msm*GlgB were purified from bacterial cell extracts using immobilized metal ion affinity chromatography (IMAC) and anion-exchange chromatography (Figure 5-1 and 5-2 A-B). Size exclusion chromatography was used to enhance the purity and to determine the oligomeric state of each protein in the solution (Figure 5-1 and 5-2 C-D). *Mtb*GlgE and *Msm*GlgB were eluted as single peaks from size exclusion chromatography at 168 mL (~120 kDa, estimated from column calibration) and 186 mL (~66 kDa), representing dimeric and monomeric state of these proteins in the solution, respectively, consistent with the literature (Syson et al., 2011 and Pal et al., 2010). Albumin was also used as a standard. Albumin elution profile contained two major peaks at 164 and 187 mL, representing the dimeric (133 kDa) and monomeric (66.5 kDa) state of this protein in the solution, respectively.



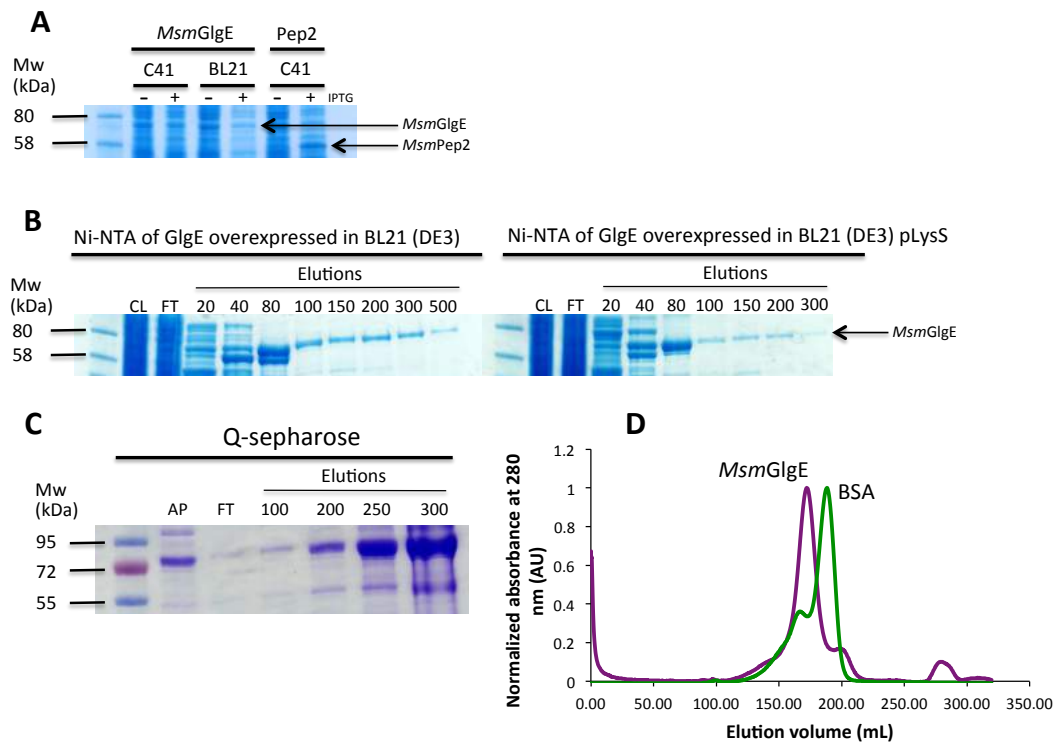
**Figure 5-1.** Purification of recombinant *MtbGlgE*. (A-B) SDS-PAGE analysis of the fractions from affinity (Ni-NTA) and anion-exchange (Q-Sepharose) chromatography. L: crude lysate, CL: clear lysate, FT: flow through, AP: applied sample. Ni-NTA elution fractions (20 mM Na<sub>2</sub>HPO<sub>4</sub> pH 7.2, 500 mM NaCl, 10 % v/v glycerol) supplemented with 20 to 500 mM imidazole and anion-exchange elution fractions (100 mM Bis-Tris propane pH 7.0, 10% v/v glycerol) supplemented with 100 to 300 mM NaCl as indicated above the gels. (C) The chromatogram of *MtbGlgE* (green) vs. albumin (red). (D) SDS-PAGE analysis of the peak fractions from size exclusion chromatography. Arrows indicate His<sub>6</sub> tagged *MtbGlgE* bands on the gel (~ 81 kDa with the His<sub>6</sub> tag).



**Figure 5-2.** Purification of recombinant *MsmGlgB*. (A-B) SDS-PAGE analysis of the fractions from affinity (Ni-NTA) and anion-exchange (Q-Sepharose) chromatography. L: crude lysate, CL: clear lysate, FT: flow through, AP: applied sample. Ni-NTA elution fractions (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 % v/v glycerol) supplemented with 20 to 500 mM imidazole and anion-exchange elution fractions (50 mM Tris-HCl pH 8.0, 10 % v/v glycerol) supplemented with 150 to 500 mM NaCl as indicated above the gels. (C) The chromatogram of *MsmGlgB* (green) vs. albumin (red). (D) SDS-PAGE analysis of the peak fractions from size exclusion chromatography. Arrows indicate His<sub>6</sub> tagged *MsmGlgB* bands on the gel (~ 85 kDa with the His<sub>6</sub> tag).

### 5.3 Expression and purification of *MsmGlgE*

First attempts to overexpress *MsmGlgE* proved the over-expression of this protein as a formidable task. In order to find the best host for producing *MsmGlgE*, the plasmid carrying *glgE* was transformed into three different strains of *E. coli*, BL21 (DE3), BL21 (DE3) pLysS and BL21 (DE3) C41, and induced each of them with 1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG). Visually testing the SDS-PAGE gels of uninduced and induced samples of BL21 (DE3) cultures transformed with *glgE* against those from BL21 (DE3) C41 cultures didn't show any significant difference (Figure 5-3A). Since the bands containing uninduced and induced samples of BL21 (DE3) pLysS cultures appeared as smeary on the SDS-PAGE gel and I was not able to compare them with other competent cells, I purified GlgE from these culture cells using IMAC and compared the intensity of bands from elution fractions with those from BL21 (DE3) cells on SDS-PAGE gel. Higher intensity of GlgE bands on BL21 (DE3) SDS-PAGE gel indicated that the amount of GlgE production in these cells should be relatively higher than BL21 (DE3) pLysS cells (Figure 5-3B). The *MsmGlgE* overexpressed in BL21 (DE3) was further purified using anion exchange chromatography and size exclusion chromatography (Figure 5-3 C-D). *MsmGlgE* was eluted as a major peak at 171 mL (~116 kDa, estimated from column calibration), representing the dimeric form of this protein in the solution, consistent with other literature (Syson et al., 2011).



**Figure 5-3.** Purification of recombinant *MsmGlgE* from bacterial cell extracts. (A) SDS-PAGE analysis of *MsmGlgE* overexpressed in two different competent cells, BL21 (DE3) and BL21 (DE3) C41. *MsmPep2* was also overexpressed in C41 and used as a positive control. (B) SDS-PAGE analysis of fractions from affinity (Ni-NTA) chromatography of *MsmGlgE* overexpressed in BL21 (DE3) and BL21 (DE3) pLysS. CL: clear lysate, FT: flow through. Ni-NTA elution fractions (20 mM Na<sub>2</sub>HPO<sub>4</sub> pH 7.2, 500 mM NaCl, 10 % v/v glycerol) supplemented with 20 to 500 mM imidazole. (C) SDS-PAGE analysis of anion-exchange elution fractions (100 mM Bis-Tris propane pH 7.0, 10% v/v glycerol) supplemented with 100 to 300 mM NaCl as indicated above the gel. (D) The chromatogram of *MsmGlgE* (purple) vs. albumin (green). Arrows indicate His<sub>6</sub> tagged *MsmGlgE* bands on the gel (~ 80 kDa with the His<sub>6</sub> tag).

## 5.4 Crystallization of *MtbGlgE*

*MtbGlgE* was concentrated to about 6.2 mg/mL and screened for crystallization in four different sparse matrix screens (for more information see section 2-8). However, it did not generate any crystals. Addition of maltose (as substrate) or cyclodextrin (as inhibitor) did not also promote the crystallization of *MtbGlgE*.

## 5.5 Probing stability of *GlgE* using thermal shift assay

In order to probe thermal stability of *MtbGlgE* as a function of buffer composition, a thermal shift assay was performed using a qPCR thermal cycler and monitored changes in melting temperature ( $T_m$ ) of this protein by monitoring fluorescence of the hydrophobic dye (Niesen, Berglund and Vedadi, 2007). The  $T_m$  value of *MtbGlgE* was measured as 18 °C and its changes ( $\Delta T_m$ ) were monitored throughout the experiment (Figure 5-4). An increase in  $T_m$  could suggest an elevation in structural stability of *GlgE*, whereas a decrease in  $T_m$  might represent a disordered conformation or a sign of misfolding. Large variation in stability was observed in different screens (Figure 5-4). Some buffers had a higher impact on *MtbGlgE* stability, mainly sodium acetate pH 5.4, MES pH 5.7 and 6.1, Hepes pH 7.0 and 8.0 with a  $\Delta T_m$  of about +8. The most destabilizing buffers are sodium citrate/citric acid pH 4.0 and sodium acetate pH 4.2 with a  $\Delta T_m$  of about -1 and some buffers with small effects on *GlgE* stability including sodium acetate pH 4.8, 250 mM NaCl and sodium citrate/citric acid pH 5.0, 250 mM NaCl. *MtbGlgE* was dialyzed against two buffers, Hepes pH 7.0 and 8.0, 250 mM NaCl and

10% glycerol and exposed to crystallization screens. Unfortunately, neither of these buffers promoted the crystallization of this protein.

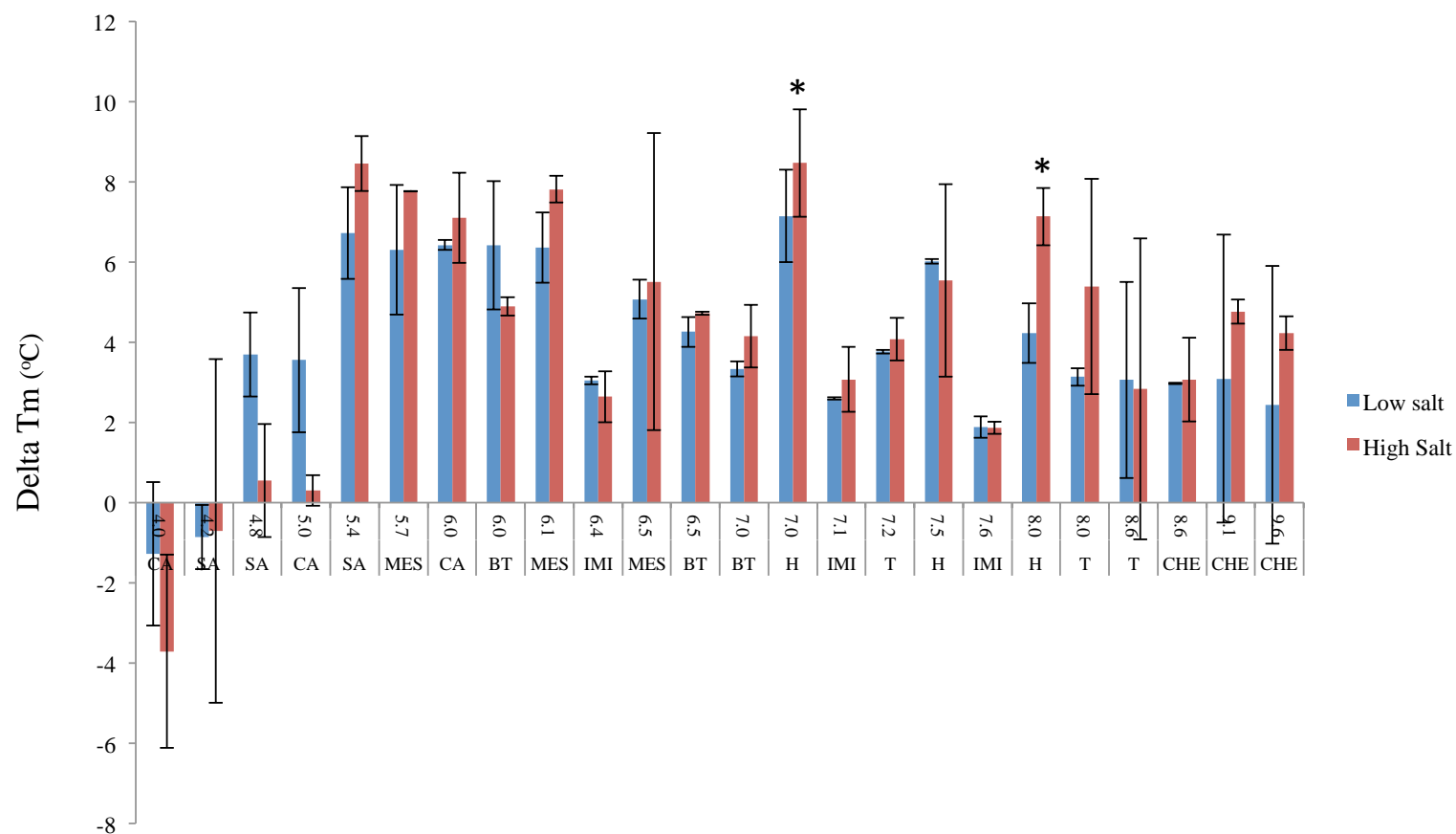
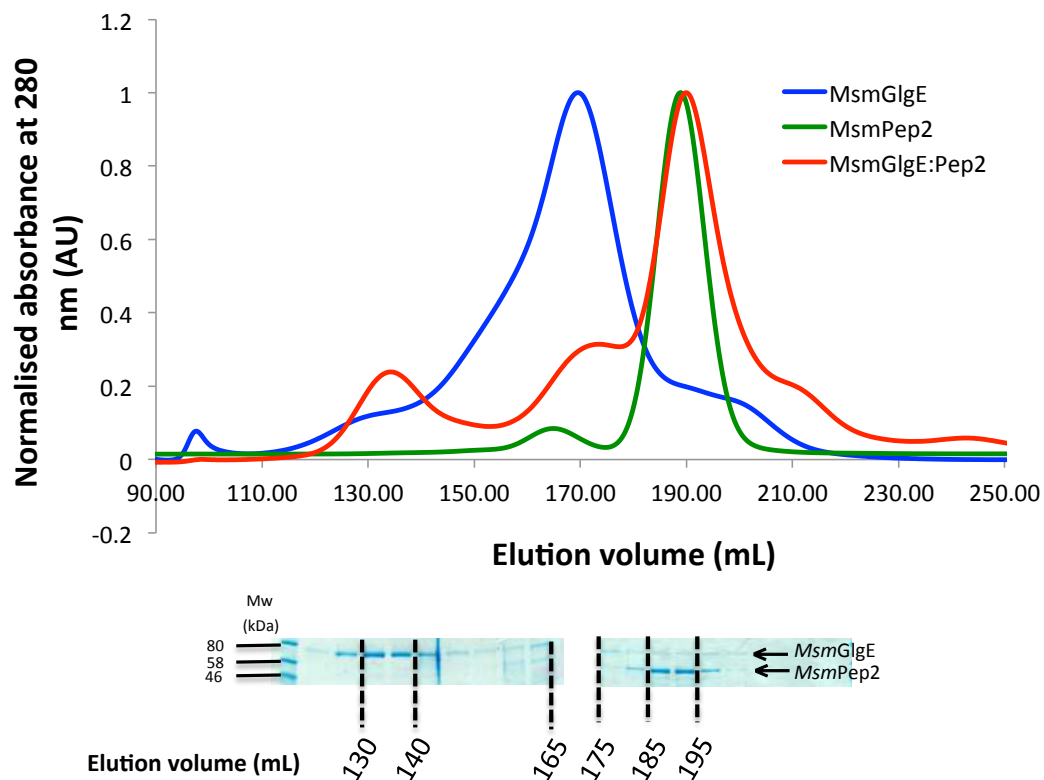


Figure 5-4. Exploring *MtbGlgE* stability versus buffer type, determined using thermal shift assay. Thermal stability was measured at 100 mM NaCl (■) and 250 mM NaCl (■). Stars above the bars indicate the buffers that were tested in the crystallization screenings. Error bars represent standard deviation. Abbreviations on X axis are explained in Table 2-4 of materials and methods.

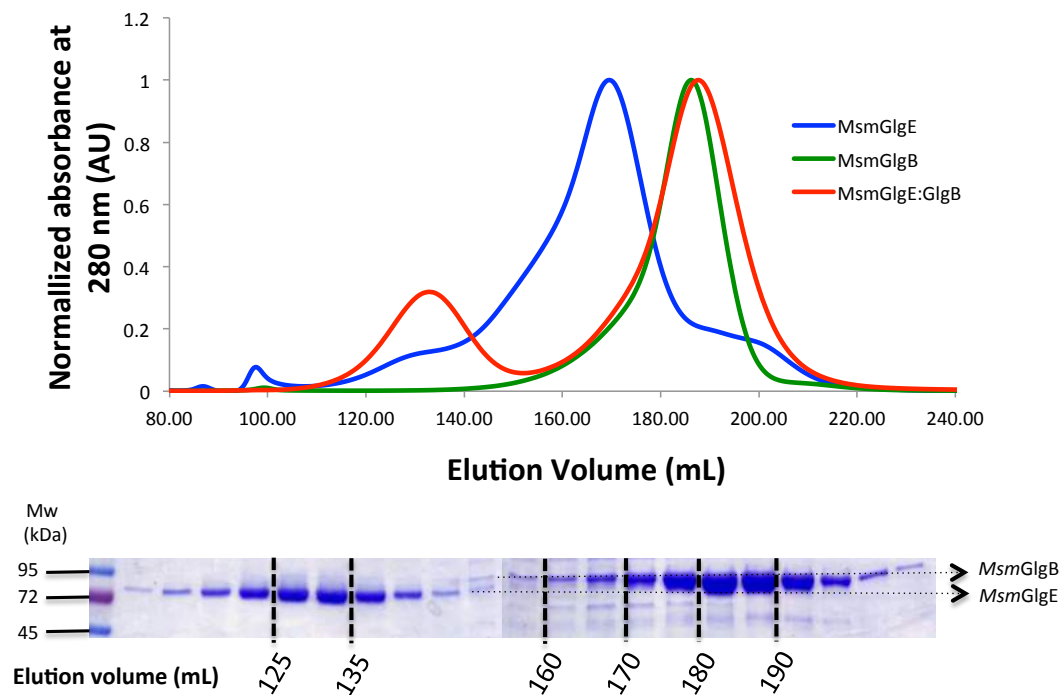


## 5.6 Testing the interaction of *MsmGlgE* with *MsmPep2* and *GlgB* using size exclusion chromatography

In order to probe the putative interactions between *MsmGlgE* and *MsmPep2*, both proteins were mixed together in 1:1 molar ratio (20  $\mu$ M of each proteins) and loaded into size exclusion chromatography (SEC) column. Three major peaks appeared at 189, 170 and 133 mL elution volumes, corresponding to the elution volumes of *MsmPep2*, *MsmGlgE* and higher oligomers of *GlgE*, respectively (Figure 5-5). If *Pep2* and *GlgE* were interacting in the form of a complex, we should see a peak at lower elution volume containing both proteins. But since each protein elutes separately, it implies that *Pep2* and *GlgE* do not form a complex. Similarly, *MsmGlgE* (60  $\mu$ M) was mixed with *MsmGlgB* (120  $\mu$ M) and loaded into the SEC column. Two major peaks appeared at 186 and 131 mL (Figure 5-6). The peak at 186 mL corresponds to *GlgB*, but the peak at 131 mL is not corresponding to elution volumes of neither *GlgE* nor *GlgB*. In order to determine the contents of this peak, the fractions pulled from SEC column were tested on SDS-PAGE gel. Fractions corresponding to the peak contained only *GlgE*, suggesting that adding *GlgB* changes the elution pattern of *GlgE* (from 170 mL to 131 mL). However, it does not provide enough evidence that *GlgE* and *GlgB* are forming a complex together. In order to support these results another technique, bacterial two-hybrid system, based on complementation between two subdomains, T25 and T18, of catalytic domain of *Bordetella pertussis* adenylate cyclase, was performed.



**Figure 5-5.** Exploring complex formation between *MsmGlgE* and *MsmPep2*. For more details see section 5.6.



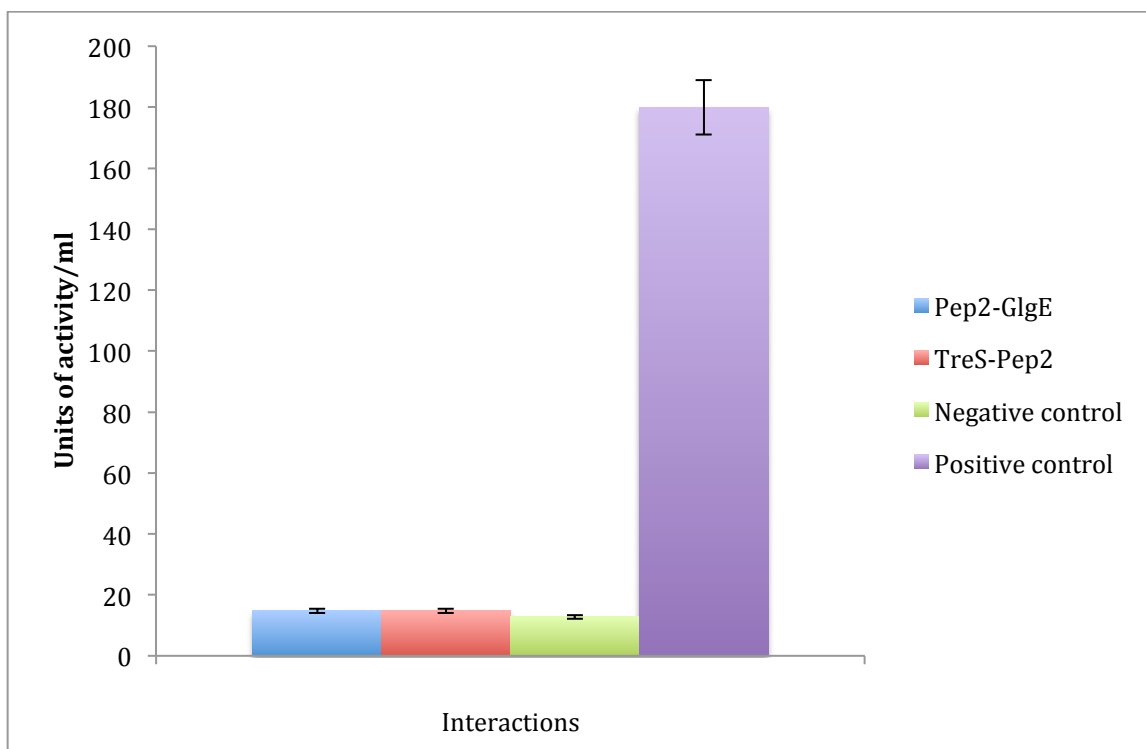
**Figure 5-6.** Exploring complex formation between *MsmGlgE* and *MsmGlgB*. Two protein bands appear at close proximity of each other on the gel. Arrows indicate each protein on the gel. The molecular weight of *MsmGlgE* and *MsmGlgB* (with the His<sub>6</sub> tag) is 80 and 85 kDa, respectively. For more details see section 5.6.

## 5.7 Bacterial Two-hybrid system

In spite of this fact that there is a functional link between TreS, Pep2, GlgE and GlgB, the interactions between these proteins barely have been studied. Roy and his colleagues showed that TreS and Pep2 form a hetero-octameric complex in *M. tuberculosis* (Roy et al., 2013) but the interaction between other members of the pathway still remains a mystery. In order to explore putative interactions, bacterial adenylate cyclase two-hybrid system (BACTH) was used (Karimova et al., 1998). This technique is based on complementation between two subdomains, T25 and T18, of catalytic domain of *Bordetella pertussis* adenylate cyclase. When T25 and T18 domains are together the enzyme is active and will synthesize cyclic adenosine 3,5-monophosphate (cAMP) from ATP. Calmodulin, a protein found only in eukaryotic organisms, acts as an intermediate in bringing two subdomains together. Newly produced cAMP interacts with catabolite activator protein (CAP) and cAMP-CAP complex elevates the transcription of lactose and maltose catabolic genes by binding to their promoters and acting as an inducer. Activation of these operons can be easily detected by measuring cAMP levels and  $\beta$ -galactosidase activities in liquid cultures. In practice two proteins of interest are fused to T25 and T18 subunits. If fused proteins interact together, they fulfill the role of calmodulin in bringing together T18 and T25 and restore the activity of adenylate cyclase.

Four enzymes involved in GlgE pathway, TreS, Pep2, GlgE and GlgB from *M. tuberculosis* were fused to T25 and T18 subdomains of adenylase cyclase on BACTH

vectors. Since the reporter strain lacks the endogenous adenylase cyclase, only interaction between two proteins fused to subdomains T25 and T18 can restore the activity of this enzyme. The interactions between two sets of proteins TreS and Pep2, Pep2 and GlgE were measured using  $\beta$ -galactosidase assay on overnight cultures of these proteins (Figure 5-7) (for more information see section 2.16). Negative control, two empty plasmids, and positive control, plasmids pKT25-zip and pUT18C-zip, were also included in the experiment. The  $\beta$ -galactosidase activities of negative and positive controls were measured as 12.8 and 180 units of activity per mL of overnight culture, respectively (Figure 5-7). Similar to negative control, the overnight cultures of both TreS:Pep2 and Pep2:GlgE combinations showed 14.8 units of  $\beta$ -galactosidase activity per mL, suggesting that these sets of proteins are not interacting together, contrary to published data on TreS:Pep2 interactions. It had been previously established that TreS and Pep2 from *M. tuberculosis* are forming a complex together in both in vitro and in vivo (Roy et al., 2013). These results discouraged me from further exploring the interactions of GlgE with the rest members of GlgE pathway. However, size exclusion chromatography also failed to show the complex formation between *Msm*GlgE and Pep2 (Figure 5-5).



**Figure 5-7.**  $\beta$ -galactosidase activity of overnight cultures prepared from BTH101 competent cells co-transformed with vectors carrying TreS and Pep2 (■), Pep2 and GlgE (■). Negative control (■) and positive control (■) also were included in this experiment. For more information see section 2.16. Error bars are representing standard deviation.

## 5.8 Summary

*Mtb*GlgE and *Msm*GlgB were purified from bacterial cell extracts (Figure 5-1 and Figure 5-2). *Mtb*GlgE and *Msm*GlgB were eluted as single peaks at 168 and 186 mL from SEC column, representing dimeric and monomeric states of these proteins in the solution, respectively (Figure 5-1C and Figure 5-2C).

The initial attempts to crystallize *Mtb*GlgE were not successful. In order to probe thermal stability of *Mtb*GlgE as a function of buffer composition, a thermal shift assay was performed. The most stabilizing buffers were determined as sodium acetate pH 5.4, MES pH 5.7 and 6.1, Hepes pH 7.0 and 8.0 (Figure 5-4) but none of them improved the crystallization of GlgE.

The interactions of *Msm*GlgE with *Msm*Pep2 and *Msm*GlgB were explored using SEC (Figure 5-5 and Figure 5-6) and BACTH systems (Figure 5-7). Together, these results suggest that the recombinant GlgE do not form a complex with either Pep2 or GlgB. However, it is obvious that adding *Msm*GlgB to *Msm*GlgE changes the elution pattern of GlgE towards lower elution volumes (Figure 5-6).

## **6 Crystallization and interaction studies of Rv3031 and Rv3032**

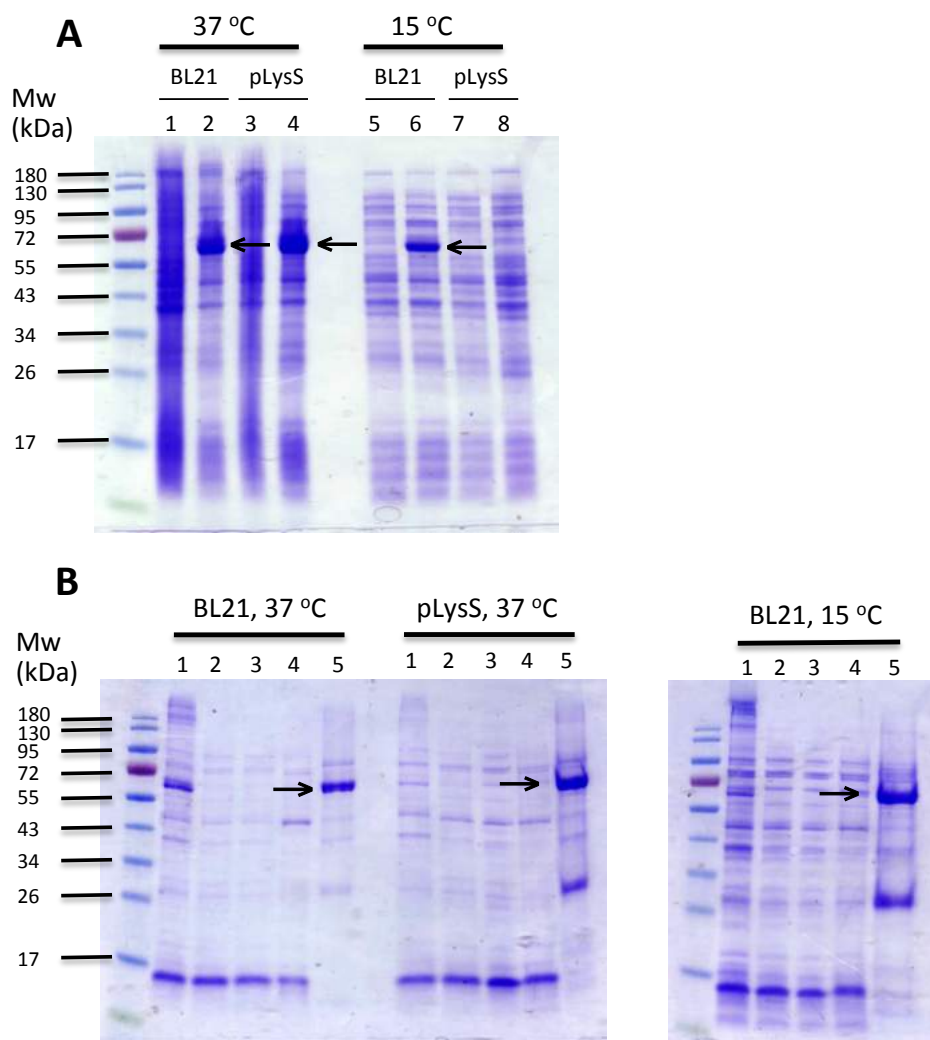


## **6.1 Crystallization of Rv3031 and Rv3032**

The third pathway involved in glucan biosynthesis in mycobacteria is Rv3032, composed of two key enzymes Rv3032 and Rv3031 (Mendes et al., 2010). Rv3032 is a (1→4)-glycosyltransferase that catalyzes the transfer of glucose from ADP-glucose or UDP-glucose to diglucosylglycerate (DGG) and Rv3031 serves as a  $\alpha$ -(1→6)-branching enzyme and introduces new branches into the growing glucan (Stadthagen et al., 2007). Neither of these enzymes are essential for mycobacterium survival, however, Rv3032 exhibits synthetic lethal interactions with both TreS and GlgA (Sambou et al., 2008 & Kalscheuer et al., 2010). Although, both enzymes, Rv3031 and Rv3032, have been known for several years and their roles in glucan synthesis has been investigated, but they are not still biochemically characterized. In addition, solving the crystal structure of both Rv3032 and GlgA might provide us with new insights into synthetic lethal interactions in mycobacteria.

## **6.2 Expression and purification of Rv3031**

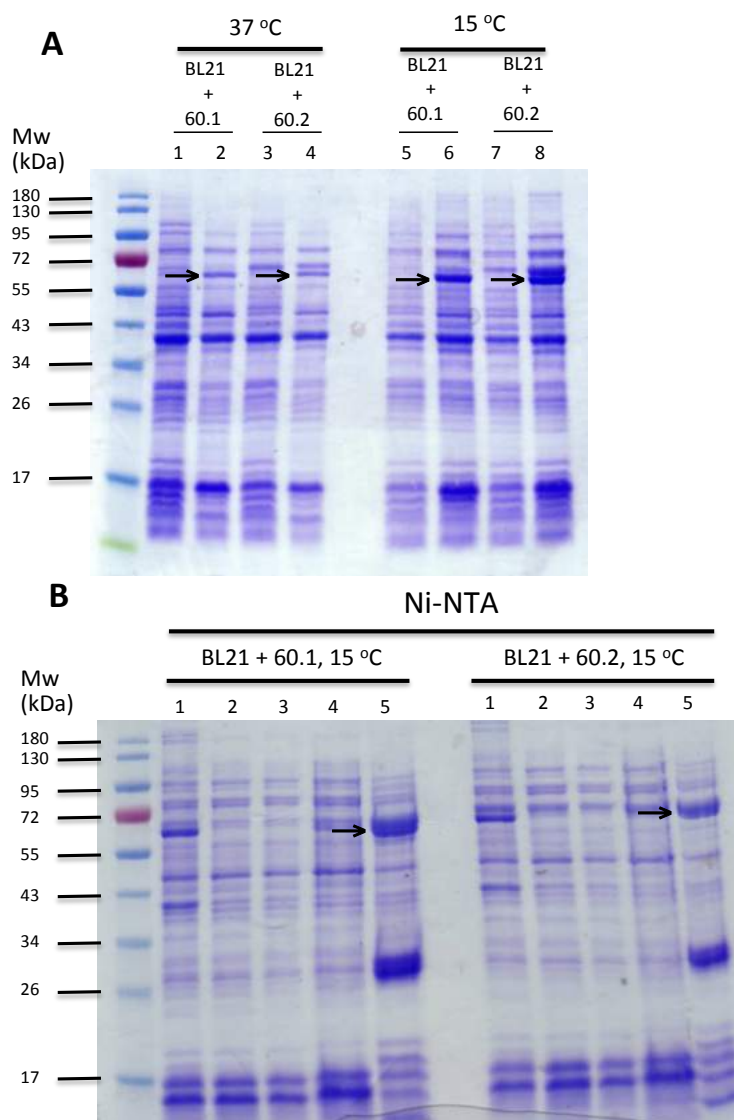
In order to find the optimum conditions to overexpress Rv3031, it was transformed into two different competent cells, BL21 (DE3) and BL21 (DE3) pLysS (Figure 6-1 A). Both types of competent cells were induced with 0.2 mM IPTG, followed by incubation either at 37 °C for 3 hours or overnight at 15 °C. Visually checking the SDS-PAGE gels of uninduced and induced samples from all four cultures indicated the overexpression of recombinant Rv3031 in both competent cells at 37 °C as well as BL21 (DE3) overnight at 15 °C but not in BL21 (DE3) pLysS at 15 °C (Figure 6-1 A). Between these two competent cells I chose BL21 (DE3) for further experimentation.



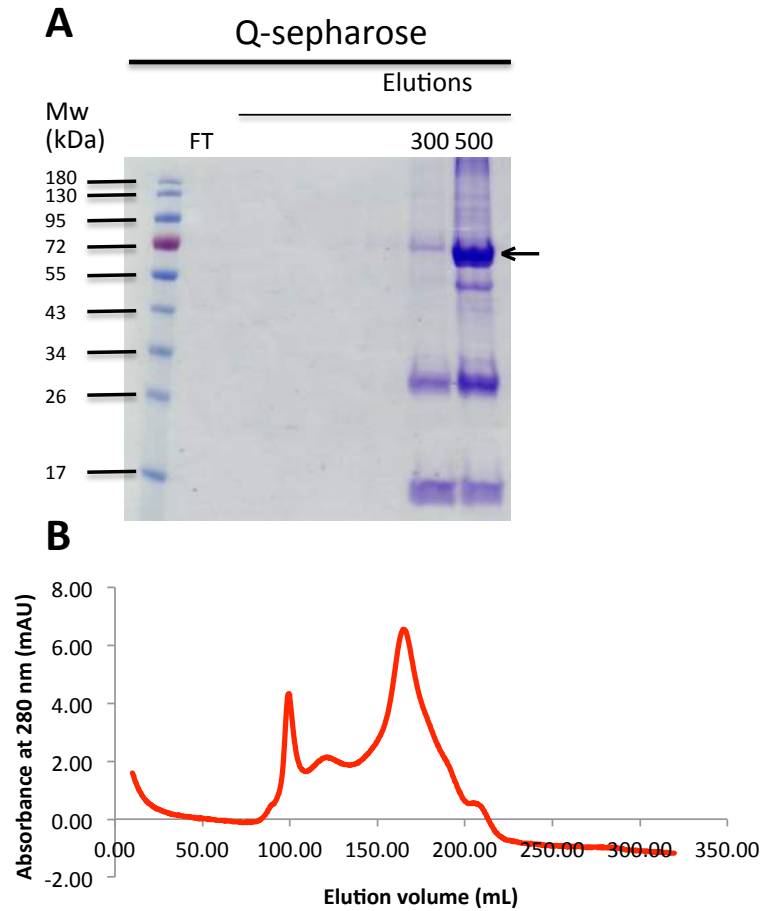
**Figure 6-1.** SDS-PAGE analysis of recombinant Rv3031 overexpression and solubility. (A) Testing the overexpression of Rv3031 using BL21 (DE3) and BL21 (DE3) pLysS competent cells at 37 °C for three hours or overnight at 15 °C. Odd and even numbers represent uninduced and induced samples, respectively. (B) SDS-PAGE analysis of the fractions from affinity (Ni-NTA) column, Lane 1: crude lysate; Lane 2: clear lysate; Lane 3: flow through and Lane 4-5: samples eluted with 20 and 500 mM imidazole, respectively. The fraction containing the protein was concentrated more than 20 fold. Arrows indicate His<sub>6</sub> tagged Rv3031 bands on the gel (~ 60 kDa with His<sub>6</sub> tag).

In order to improve the solubility of Rv3031, this protein was coexpressed with *M. tuberculosis* Cpn60.1 and *E. coli* GroES or *M. tuberculosis* Cpn60.2 and *E. coli* GroES at 37 °C for three hours or overnight at 15 °C (Figure 6-2A). Cultures grown overnight at 15 °C produced relatively more proteins than cultures grown at 37 °C for three hours. Cells collected from overnight cultures were lysed and Rv3031 was purified using Ni-NTA chromatography for further analysis (Figure 6-2B). In order to be able to see the protein bands on the gels, each fraction was 30 folds concentrated. However, coexpression of Rv3031 with chaperons also didn't improve the solubility of this protein significantly. The highest amount of protein was obtained only when 3 liters of bacterial cultures was used and cultures were induced with lower amount of IPTG (0.2 mM) (Figure 6-3A).

Size exclusion chromatography (SEC) was used to enhance the purity of Rv3031 and also to determine the oligomeric state of this protein. Rv3031 was eluted as three peaks at 98, 119 and 163 mL elution volumes, representing aggregation, higher oligomeric and dimeric state of this protein in the solution, respectively (Figure 6-3B).



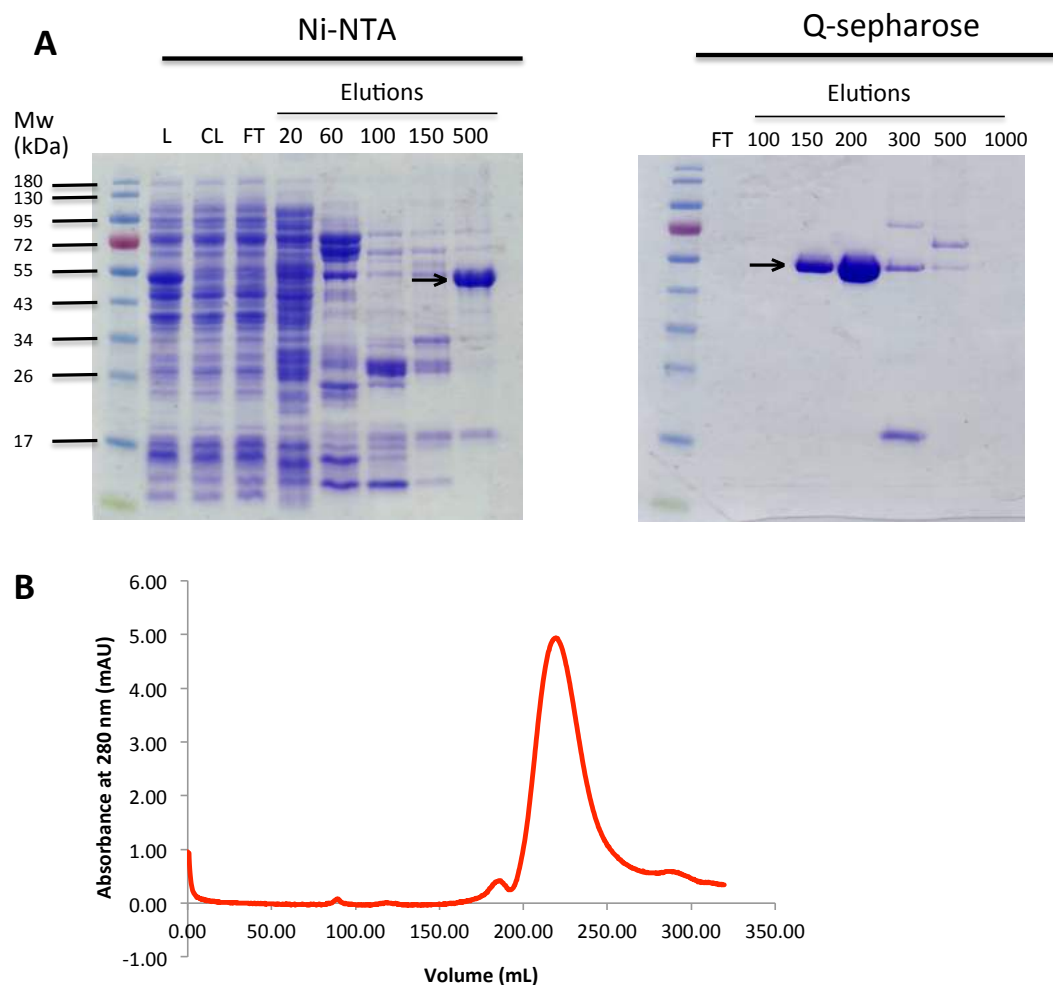
**Figure 6-2.** SDS-PAGE analysis of recombinant Rv3031 overexpression and solubility in the presence of chaperones. (A) Testing the overexpression of Rv3031 in the presence of 60.1 GroES and 60.2 GroES chaperones at 37 °C for three hours or overnight at 15 °C. Odd and even numbers represent uninduced and induced samples, respectively. (B) SDS-PAGE analysis of the fractions from affinity (Ni-NTA) column, Lane 1: crude lysate; Lane 2: clear lysate; Lane 3: flow through and Lane 4-5: samples eluted with 20 and 500 mM imidazole, respectively. Arrows indicate His<sub>6</sub> tagged Rv3031 bands on the gel (~ 60 kDa with His<sub>6</sub> tag). In order to be able to see the bands on the gel, each fraction was ~30 folds more concentrated.



**Figure 6-3.** Purification of recombinant Rv3031 from three liters of bacterial cultures. (A) SDS-PAGE analysis of anion-exchange elution fractions (25 mM Tris-HCl pH 8.2, 10% v/v glycerol) supplemented with 0 to 500 mM NaCl. FT: flow through. (B) The size exclusion chromatogram of Rv3031 from Sephacryl S300 HR column. Arrow indicates His<sub>6</sub> tagged Rv3031 band on the gel (~ 60 kDa with His<sub>6</sub> tag).

### **6.3 Expression and purification of Rv3032**

Rv3032 was purified using immobilized metal ion affinity chromatography (IMAC) and anion-exchange chromatography (Figures 6-4A). Using size exclusion chromatography (SEC) the purity of Rv3032 was enhanced and the oligomeric state of protein was determined (Figure 6-4B). Rv3032 was eluted as a major peak from SEC column at 218 mL, suggesting that this protein behaves as a monomer in the solution.



**Figure 6-4.** Purification of recombinant Rv3032. (A) SDS-PAGE analysis of the fractions from affinity (Ni-NTA) and anion-exchange (Q-Sepharose) chromatography. L: crude lysate, CL: clear lysate, FT: flow through. (B) The size exclusion chromatogram of Rv3032 from Sephacryl S300 HR column. Arrows indicate 6xHis-tagged Rv3032 bands on the gel (~ 47 kDa with the His<sub>6</sub> tag).

## 6.4 Domain structure of Rv3031 and Rv3032

According to Pfam (pfam.xfam.org), Rv3031 is composed of two domains, GH57 and DUF1957. Glycoside hydrolase family 57, also known as GH57, is a group of enzymes involved in hydrolyzing glycosidic bonds between two carbohydrates. The members of this family have a range of different activities such as:  $\alpha$ -galactosidase,  $\alpha$ -amylase,  $\alpha$ -glucanotransferase and branching enzyme. On the other hand, DUF1957 is a domain with unknown activity. This domain has been identified in several branching enzymes including TT1467 from *Thermus thermophiles* (PDB entry 1UFA) and TK1436 from hyperthermophilic archaeon *Thermococcus kodakaraensis* (PDB entry 3N8T). In TK1436, this domain is involved in the formation of active-site cleft (Santos et al., 2011).

Similarly, Rv3032 is made of two domains, glycosyl transferase 4-like domain and glycosyl transferases group 1. Glucosyltransferases are known to catalyze the formation of glycosidic bonds by using sugar phosphates as glycosyl donors (<https://www.cazypedia.org/index.php/Glycosyltransferases>).

## 6.5 Summary

It was shown that the highest amount of recombinant Rv3031 and Rv3032 could be obtained only if the genes encoding these proteins are transformed in BL21 (DE3) competent cells and grown overnight at low temperatures (15 °C). However, it is advisable to use 3 - 5 liters of culture instead of one liter and the amount of inducing IPTG be reduced to 0.2 mM or even lower (some literatures suggest 10  $\mu$ M IPTG). SEC



chromatogram suggest that Rv3031 and Rv3032 behave mainly as dimers and monomer in the solution (Figure 6-3B and Figure 6-4C).

## **7 Discussion**

It has been shown that capsule of mycobacteria, the most outer layer of cell envelope, is involved in modulating the host immune system (Geurtsen, et al., 2009 & Sani et al., 2010). Thus, understanding the mechanisms of capsule biosynthesis can help us to better understand the biology of this pathogen, its interactions with the host cells and ultimately aid the development of new tuberculosis therapies. The main constituent of the capsule,  $\alpha$ -glucan, is synthesized in three interlinked pathways (Chandra, Chater and Bornemann, 2011). In the recently discovered GlgE pathway, trehalose is converted to maltose in the presence of TreS, Pep2 phosphorylates the latter to maltose-1-phosphate using ATP, GlgE transfers maltose from maltose 1-phosphate to maltooligosaccharides and GlgB introduces branches to the linear glucan at some position 6. In a recent study by Roy et al, it was shown that TreS and Pep2 form an octameric complex in *Mtb* and they solved the structure of TreS. The purpose of this study was to solve the structure of GlgE, a genetically validated novel anti-tuberculosis drug target, Pep2 on its own or in complex with TreS and determine the affinity of TreS:Pep2 association.

The crystal structure of GlgE from *Streptomyces coelicolor* (an orthologous of *M. tuberculosis*) was described for the first time by Syson and colleagues (Syson et al., 2011). The *S. coelicolor* GlgE structure (~51% sequence similarity with *Mtb*GlgE) describes the catalytic and kinetic properties of the enzyme, however, in order to be able to develop an inhibitor explicitly, it is essential to obtain the structure of GlgE from *M. tuberculosis*. *Mtb*GlgE was successfully purified to ~95% purity (single band on the SDS-PAGE gel) and exposed to a wide range of crystallization screens (Figure 5-1). Unfortunately, it did not generate any crystals. One possible solution to this problem was

promoting the crystallization of GlgE by improving the stability and solubility of this protein in the solution (Niesen, Berglund and Vedadi, 2007). Therefore, a thermal shift assay was performed to find the most suitable buffer and pH for enhancing the stability of *MtbGlgE*. According to thermal shift assay sodium acetate pH 5.4, MES pH 5.7 and 6.1, Hepes pH 7.0 and 8.0 could be the most promising buffers for *MtbGlgE* stabilization, though; these buffers also didn't promote the crystallization of *MtbGlgE*.

The crystal structure of GlgE from *Mtb* co-crystallized with maltose and acceptor molecule, maltohexaose, was recently described (Lindenberger et al., 2015). The *MtbGlgE* exhibits a very similar structure to *Streptomyces coelicolor* (*Sco*) GlgE with an R.M.S. displacement value of 2.5 Å for all Ca positions. However, SAXS data indicate a difference in orientation of the monomers within the homodimers of *Mtb* and *Sco* GlgE. This difference in homodimers can be explained as a result of forming a disulfide bridge between monomers in *MtbGlgE*.

M1P binding site in *MtbGlgE* is highly conserved and has a similar structure and amino acid sequence to *ScoGlgE* (Lindenberger et al., 2015). It is composed of two subunits: the -1 subunit where the reducing end of maltose sits and the -2 subunit where the non-reducing end of maltose binds. The main difference between M1P binding sites of two enzymes is observed in -2 subunit where one amino acid is changed. While *MtbGlgE* contains a serine at position 303, *ScoGlgE* contains a valine in this position (V279). *MtbGlgE* S303 forms a hydrogen bond with O5 atom of the substrate, while V279 in *ScoGlgE* forms a van der Waals interaction with 2 carbon atoms of the substrate.

Lindenberger and colleagues also found a novel maltooligosaccharide binding site in *MtbGlgE* different from previously described cyclodextrin binding site in *ScoGlgE*. The newly discovered binding site is located in Domains A and C of *MtbGlgE* and can harbor an acceptor as large as a maltohexaose through forming hydrogen bonds with its atoms. Cyclodextrin in *ScoGlgE* is identified in hydrophobic region of homodimer interface between Domain A of one subunit and Domain N of the second subunit. Cyclodextrin binding site is in close proximity of maltohexaose binding site and inhibits the activity of *ScoGlgE* by blocking the latter site and hindering the access of glucan substrate to enzyme active site. However, in *MtbGlgE* R427 and G84 residues are replaced by proline residues, resulting a clash between proline residues side chains with cyclodextrin. This prevents the binding of inhibitor to the binding site and therefore, maintaining the activity of *MtbGlgE*. This study provided further insight into the structure and activity of *MtbGlgE* and paving the way to design more efficient inhibitors of this enzyme.

The second area of interest was to explore the putative interactions of GlgE with other members of the pathway. Given that the accumulation of maltose-1-phosphate (produced by the catalytic activity of Pep2) is toxic to the mycobacterial cells, one can argue that the cells need to ensure that the intermediate (M1P) cannot accumulate. Using two different techniques, size exclusion chromatography and bacterial two hybrid systems (BACTH), I explored the putative interaction of Pep2:GlgE and GlgE:GlgB. Results of these studies suggest that the recombinant GlgE do not form a complex with either Pep2 or GlgB (Figure 5-5 and 5-6). However, the interaction between these pairs of proteins might be too weak to be detected by size exclusion chromatography and BACTH, which known to

produce artifacts, or additional factors such as TreS or maltose might be required to stimulate the complex formation, more sensitive techniques can be examined to further investigate these results. This is evident when probing the activity of *MsmPep2* and *MsmTreS* (positive control) using bacterial two-hybrid system. Although the complex formation between these two proteins has been confirmed in vitro and in vivo using a range of biophysical techniques, the level of  $\beta$ -galactosidase production between cells carrying the complementing T18 and T25 fused to TreS and Pep2 and negative cells is the same, indicating there is no complex formation between *MsmTreS* and *MsmPep2* (Contrary to published data) (Figure 5-7). Moreover, according to the crystal structure of *MtbPep2*, Li et al. (2014) proposed that Pep2 could serve as a docking site for GlgE to form a super complex with TreS. However, they also failed to show this complex formation using pull-down and Biacore assays (Li et al., 2014). Interestingly, adding *MsmGlgB* to *MsmGlgE* changed the elution pattern of GlgE towards lower elution volumes (Figure 5-6), which might suggest changing the oligomeric state of this protein in the solution.

Roy and his colleagues described the structure of *MtbTreS* recently and confirmed the complex formation between *MtbTreS* and *MtbPep2* using size exclusion chromatography, and analytical ultracentrifugation in sedimentation equilibrium mode. However, they did not solve the structure of complex nor determine the affinity of complex formation between TreS and Pep2. Initial attempts to crystallize Pep2 from *Mtb* were not successful. Both size exclusion chromatography and analytical ultracentrifugation (AUC) results showed that *MtbPep2* behaves as a heterodisperse protein (forming monomers,

dimers and higher oligomers) in the solution (Figure 3-2), which makes the crystallization less likely. A small angle scattering experiment performed at ESRF's bioSAXS beamline also, showed that this protein is prone to aggregation in solution (data not shown). Thus, I tried to crystallize its homologue from *M. smegmatis*. In contrast to *MtbPep2*, *MsmPep2* appeared as a monomer in the solution (Figure 3-2) and formed sizable crystals in initial crystallization trials. However, the crystals diffracted to only 7.5 Å and only the unit cells could be determined. My attempts to improve the diffraction quality of *MsmPep2* crystals through systematic screening of conditions, including for instance the use of additives, were not successful. One reason could be the fact that this protein was co-purified with trace amounts of proteases, leading to degradation, with fragments appearing as lower bands on SDS-PAGE following ion exchange chromatography (Figure 3-1). These lower bands are also, seen on SDS-PAGEs of TreS. Western blotting of *MsmPep2* and *MsmTreS* with an anti-His<sub>6</sub> tag antibody showed that degradation happens at the C-terminal end. Degradation of TreS from *M. smegmatis* was previously reported by Pan and colleagues (2008). They monitored *MsmTreS* for 43 days and found that the molecular mass was reduced from 68 kDa to 58 kDa, losing about 10 kDa. I tried to examine the degradation of *MsmPep2* in different periods of time and *MsmPep2* showed to be very unstable on ice. Another possibility to improve the diffraction of *MsmPep2* crystals was swapping the affinity tag from the N- to the C-terminus. C-terminally tagged *MsmPep2* was successfully overexpressed, purified (Figure 3-5) and used for crystallization trials but this time it didn't generate any crystals at all. Applying other techniques such as limited proteolysis or thermal shift assay to improve the crystallization of *MsmPep2* were not successful too.

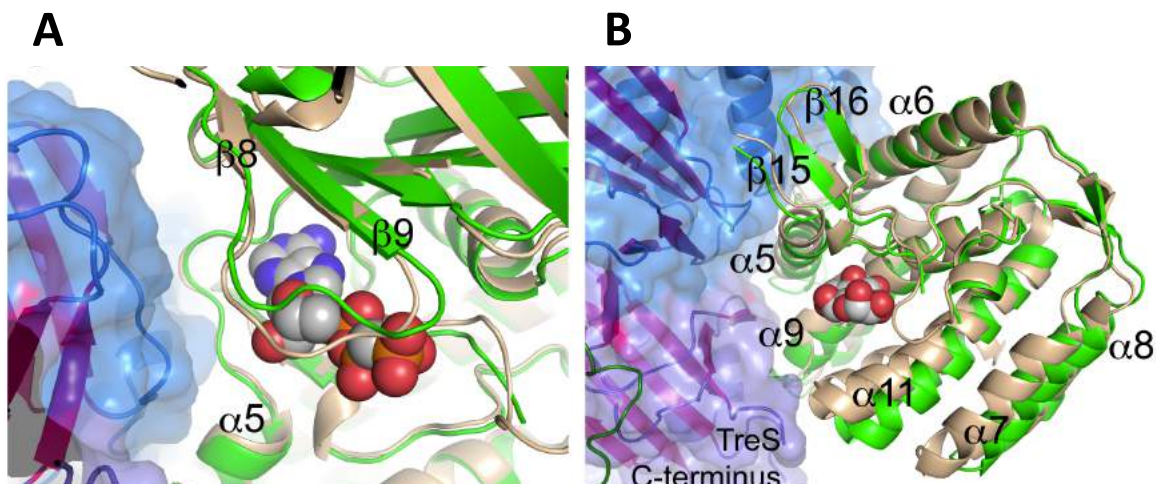
Using size exclusion chromatography I showed that TreS and Pep2 from *M. smegmatis*, similar to *Mtb* orthologues, form a complex together in the solution (Figure 4-3). Addition of Pep2 to TreS shifted the TreS tetramer peak (from 144 mL) towards lower elution volumes (Figure 4-3 and 4-4). Interestingly, it was observed that the association of *Msm*TreS and Pep2 is highly pH dependent, as the complex peak in the SEC chromatogram shifted towards lower elution volumes at the pH lowered. The highest peak was observed at pH 6.0 (134 mL) and by increasing the pH the height of peak reduced (142 mL at pH 9.0), suggesting a decrease in molecular mass (Figure 4-4). These results were confirmed by isothermal titration calorimetry (ITC) (Figure 4-6 and Table 4-2). The strongest and lowest affinities were measured at pH 6.5 and 8.5, respectively (with a  $K_d$  of 3.5  $\mu$ M and 34.8  $\mu$ M). Probing the activity of *Msm*Pep2 in the presence of *Msm*TreS at pH 6.0 - 7.5 indicated that the Pep2 rate-enhancement previously observed in *Mtb* (Roy et al., 2013) occurs only at pH 6.0, consistent with SEC and ITC results that measured the highest affinity in this pH. Based on this fact that the stronger association between TreS and Pep2 occurs at mildly acidic pH, it is not very surprising to obtain the best diffracting crystals at reservoir pH values between 6.7 and 7.0 (see below).

Given that TreS and Pep2 form a complex and that crystallization of Pep2 failed to yield crystals of sufficient quality, I decided to attempt crystallization of the TreS:Pep2 complex. While co-purifying the two proteins failed to produce crystals, purifying each protein separately and then mixing them at molar ratios of 1:1 and 2:1 (TreS:Pep2) just before setting up the crystal trays was the only way to generate *Msm*TreS:Pep2 crystals. Co-purification of TreS:Pep2 together or co-eluting from size exclusion chromatography



hinders the crystallization. Although *Msm*TreS:Pep2 complex crystallized readily but improving the diffraction from 8.5 Å to 3.6 Å was achieved only by extensive screenings and systematic variation of crystallization conditions. The initial crystals appeared at 0.2 M MgCl<sub>2</sub>, 0.1 M Tris pH 7.0 and 10% (w/v) PEG 8K. I realized that pH variation has a significant impact on crystal formation. Varying pH in 0.15 units was critical to obtain well- diffracting crystals of the TreS:Pep2 complex.

The thermodynamic equilibrium favors trehalose production over maltose production (Zhang et al., 2011). Thus, in order to drive the pathway towards glucan biosynthesis, substrate channeling might be a factor that can facilitate this. Moreover, in the genome of a significant number of bacteria *treS* and *pep2* are fused together (Chandra, Chater and Bornemann, 2011). Therefore, I was interested to test hypothesis. However, it became evident from crystal structure of *Msm*TreS:Pep2 complex that there is no substrate channels between the active site of TreS and Pep2 and the rate enhancement of Pep2 is due to the conformational changes in Pep2 induced by binding to TreS. Superimposing *Msm*Pep2 from TreS:Pep2 complex with the structures from *Mtb* (PDB entry 4O7P) and *M. vanbaalenii* (PDB entry 4U98) indicated that the conformational change for ATP binding site is limited to  $\beta$ 8- $\beta$ 9 loop, which does not contact with TreS (Figure 7-1 A). But maltose-binding site is probably more affected by binding to TreS (Figure 7-1 B). Binding to TreS obstructs the bending effect of helices  $\alpha$ 11- $\alpha$ 12, an effect, which is apparent in the conformation of Pep2 in the absence of TreS and upon binding of maltose. Secondly, contacts between the  $\alpha$ 9-  $\alpha$ 10 loop and TreS affects the conformation of helix  $\alpha$ 9 as a whole.



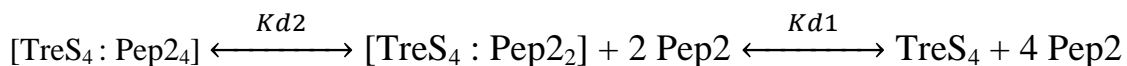
**Figure 7-1.** Comparison of *MsmPep2* from TreS:Pep2 complex with Pep2 homologue structures and structural rational for Pep2 rate-enhancement. (A), (B) Superposition of chain J of Pep2 (green) in the TreS:Pep2 complex with AppCp- and maltose-bound structures (beige) of the Pep2 orthologues from *M. vanbaalenii* (PDB entry 4U98) and *M. tuberculosis* (PDB entry 4O7P), respectively.

Prior to solving the crystal structure of TreS:Pep2 complex, our attempts to predict the tertiary structure of Pep2 using the HHpred threading server ([toolkit.tuebingen.mpg.de/hhpred](http://toolkit.tuebingen.mpg.de/hhpred)) failed to provide a structural model for the folding of the first 100 amino acids of the Pep2 sequence, while high-confidence models for the remainder of the sequence were obtained. This observation suggested that either the N-terminal sequence encodes a unique structure or that this sequence is intrinsically disordered. In order to test the hypothesis of intrinsic disorder, I generated N-terminal truncation mutants of *MsmPep2* in steps of 10 residues, removing up to 70 residues. Following structure determination of the TreS:Pep2 complex, it became evident that this region is folded but structural alignment of Pep2 with its closest structural neighbors (apart from the mycobacterial orthologues) demonstrated that some parts of this region,

$\beta$ 1-  $\beta$ 3, (Figure 4-13) are unique to mycobacterial Pep2 and are missing from the structural neighbors identified by distance matrix alignment (DALI (Holm and Rosenstrom, 2010)) (Figure 4-13). Since the function of this unique  $\beta$ -sheet was not apparent from the crystal structure, the Pep2 ( $\Delta$ 70) mutant was used to explore the role of this region of Pep2 in binding to TreS and the activity of Pep2. Using size exclusion chromatography I confirmed that mutant Pep2 ( $\Delta$ 70) was still able to bind to TreS, though with lower affinity, (Figure 4-14) and the activity ( $V_{\max}$ ) of truncation mutant reduced by  $\sim$ 10 folds compare to the wild type protein (Figure 4-15 and Table 4-5). These changes in affinity and activity can be described by destabilizing the N-terminal lobe, which contains the ATP binding site, due to truncating Pep2. Therefore, the unique  $\beta$ -sheet region is likely important for structural stability of the N-terminal lobe.

The crystal structure of TreS:Pep2 indicated that four Pep2 monomers bind to a TreS tetramer (4 + 4 configuration) (Figure 4-12). However, exploring the stoichiometry of the TreS and Pep2 association in the solution using analytical ultra centrifugation (AUC) in sedimentation equilibrium mode and small angle X-ray scattering (SAXS) suggested that only 2 subunits of Pep2 bind to a TreS tetramer (4 + 2 configuration) (Figures 4-17 and 4-18). This discrepancy in stoichiometry between the crystalized form and the behavior in the solution is likely due to crystal lattice contacts, in which the solvent is excluded and new packing interactions can affect the assembly state. In the solution state, packing interactions evidently play no role. Close inspection of the ITC data demonstrated that there is a small, but consistent deviation between the single-binding site model and the data points. Fitting a two-binding site model removes this deviation, suggesting that the

two models have  $K_d$  values of 0.5 and 18  $\mu\text{M}$ , respectively (Table 4-2). The stoichiometry for site one and two is 0.45 and 0.46, respectively (Table 4-2). Adding the stoichiometric ratios of the two sites together generate the total stoichiometry of  $n = 0.91$ , suggesting a TreS:Pep2 ratio of near unity. Based on this data the following model can be proposed:



In this model, two copies of Pep2 bind to a TreS tetramer with the higher affinity ( $K_{d1} = 0.5 \mu\text{M}$ ), while two additional copies of Pep2 can bind with lower affinity ( $K_{d2} = 18 \mu\text{M}$ ), resulting the dominance of 4 + 2 configuration in the solution. The SAXS derived molecular envelope suggests that two Pep2 monomers bind to opposite apical positions on the TreS tetramer (Figure 4-17). In this case, site 1 represents the apical positions on both sides of TreS, while site 2 is the second open site at each apex.

Solving the structure of TreS:Pep2 and determining the stoichiometry of complex formation in the solution can be used to explain why bacterial two hybrid system failed to show the interaction between these two proteins. It is well established that TreS behaves as a tetramer in the solution, with C-terminal  $\beta$ -sandwich domains participating in dimer interface. Fusing T18 or T25 domains of adenylate cyclase to C-terminal domain of TreS can hinder this dimer formation, while assembling TreS tetramer is the first step in complex formation with Pep2. Moreover, Pep2 monomers bind to opposite apical positions of the TreS tetramer with both lobes of Pep2 participating in binding interface (Figure 4-16). Likewise, fusing either of two domains of adenylate cyclase to Pep2 lobes

might affect the complex formation between TreS and Pep2. These data all together, suggest that BACTH may not be the best approach to probe the interaction between TreS and Pep2, since fusing external domains to these proteins probably influences the folding, subunits interface and interaction of these two proteins together.

Roy and colleagues confirmed the complex formation between *Mtb*TreS and Pep2 in vivo (Roy et al., 2013). They concluded that complex formation is an efficient strategy to direct the flow of  $\alpha$ -glucan pathway towards glucan synthesis, despite the fact that thermodynamic equilibrium favors trehalose production over maltose production. This change in flow direction prevents the accumulation of toxic pathway intermediate maltose 1-phosphate. Accumulation of large amounts of maltose 1-phosphate can lead to inhibition of respiration, induction of the stringent response and DNA damage and ultimately death of *Mtb* (Kalscheuer et al., 2010). Using complex formation as an approach to regulate pathway activity has been previously reported in mycobacteria. Complex formation between chorismate mutase (CM) and 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAH7PS) is a direct example of regulating metabolic pathways through complex formation between two distinct proteins in *Mtb* (Blackmore et al., 2015). They showed that both enzymes are active but complex formation enhances the catalytic activity of both enzymes, controls the flux of pathway and subsequently directs the pathway intermediate, chorismate, into phenyl alanine/tyrosine or tryptophan biosynthesis.

## Future work

During this study we were able to determine the role of complex formation between TreS and Pep2 on the activity of each enzyme but the importance of this complex formation for GlgE pathway is not still very clear. Also, it was shown that the interaction between TreS and Pep2 is highly pH dependent, but what makes this interaction pH-dependent remains to be determined. Despite the fact that there are three pathways involved in  $\alpha$ -glucan production, perturbing in two pathways (either TreS and Rv3032 or Rv3032 and GlgA) can lead to death of *Mtb* (synthetic lethal interactions). Revealing the reasons behind synthetic lethality can shed light on understanding the mechanisms of *Mtb* survival within the host or exploiting the derived information for designing the new anti-TB drugs. Finally, one of the key questions regarding the capsule biosynthesis is to understand what pathway is specifically responsible for  $\alpha$ -glucan production and how these materials are transported to the outside of cell wall after being synthesized in the cytoplasm. These are only some of the questions still remained unexplained and future studies are needed to find the answer of these questions.

## 8 References

Andries, K., Verhasselt, P., Guillemont, J., Gohlmann, H.W.H., Neefs, J.-M., Winkler, H., Gestel, J.V., Timmerman, P., Zhu, M., Lee, E., Williams, P., Chaffoy, D.d., Huitric, E., Hoffner, S., Cambau, E., Truffot-Pernot, C., Lounis, N. and Jarlier, V. (2005) 'A diarylquinoline drug active on the ATP synthase of *Mycobacterium tuberculosis*', *Science*, vol. 307, pp. 223-227.

Basel, H.H. (1998) 'History of tuberculosis', *Respiration*, vol. 65, pp. 5-15.

Battesti, A. and Bouveret, (2012) 'The bacterial two-hybrid system based on adenylate cyclase reconstitution in *Escherichia coli*', *Methods*, vol. 58, pp. 325-334.

Birch, H.L., Alderwick, L.J., Bhatt, A., Rittmann, D., Krumbach, K., Singh, A., Bai, Y., Lowary, T.L., Eggeling, L. and Besra, G.S. (2008) 'Biosynthesis of mycobacterial arabinogalactan: identification of a novel  $\alpha(1\rightarrow3)$  arabinofuranosyltransferase', *Molecular Microbiology*, vol. 69, pp. 1191-1206.

Blackmore, N.J., Nazmi, R., Hutton, R.D., Webby, M.N., Baker, E.N., Jameson, G.B. and Parker, E.J. (2015) 'Complex formation between two biosynthetic enzymes modifies the allosteric regulatory properties of both: an example of molecular symbiosis', *The Journal of Biological Chemistry*.

Brennan, P.J. (2003) 'Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*', *Tuberculosis*, vol. 83, pp. 91-97.

Brennan, P.J., Young, D.B. and Robertson, B.D. (ed.) (2008) *Handbook of anti-tuberculosis agents*, Elsevier.

Brightbill, H.D., Libraty, H., Krutzik, S.R., Yang, R.-B., Belisle, J.T., Bleharski, J.R., Maitland, M., Norgard, M.V., Plevy, S.E., Smale, S.T., Brennan, P.J., Bloom, B.R., Godowski, P.J. and Modlin, R.L. (1999) 'Host defence mechanisms triggered by microbial lipoproteins through toll-like receptors', *Science*, vol. 285, pp. 732-736.

Caner , S., Nguyen, N., Aguda , , Zhang , R., Pan, Y.T., Withers , S.G. and Brayer , G.D. (2013) 'The structure of the *Mycobacterium smegmatis* trehalose synthase (TreS) reveals an unusual active site configuration and acarbose binding mode', *Glycobiology*, pp. 1-31.

Chandra, G., Chater , K.F. and Bornemann, S. (2011) 'Unexpected and widespread connections between bacterial glycogen and trehalose metabolism', *Microbiology*, pp. 1565-1572.

Chapman, G.B., Hanks , J.H. and Wallace, J.H. (1959) 'An electron microscope study of the disposition and fine structure of *Mycobacterium lepraemurium* in mouse spleen', *Journal of Bacteriology*, vol. 77, pp. 205-211.

Cole, J.L., Lary, J.W., Moody, T. and Laue, T. (2008) 'Analytical Ultracentrifugation: Sedimentation Velocity and Sedimentation Equilibrium', *Methods Cell Biology*, vol. 84, pp. 143-179.

Cosma, C.L., Sherman, D.R. and Ramakrishnan, L. (2003) 'The secret lives of the pathogenic mycobacteria', *Annual Review of Microbiology*, vol. 57, pp. 641-676.

Cywes , C., Hoppe, H.C., Daffe , M. and Ehlers, M.R.W. (1997) 'Nonopsonic binding of *Mycobacterium tuberculosis* to Complement Receptor Type 3 Is mediated by capsular polysaccharides and is strain dependent', *Infection And Immunity*, vol. 65, pp. 4258-4266.

Daffe , M. and Etienne , G. (1999) 'The capsule of *Mycobacterium tuberculosis* and its implications for pathogenicity', *Tubercle and Lung Disease* , pp. 153-169.

Daniel , T.M. (2006) 'The history of tuberculosis', *Respiratory Medicine*, vol. 100, pp. 1862-1870.

Dinadayala, P., Sambou, T., Daffe, M. and Lemassu, A. (2008) 'Comparative structural analyses of the  $\alpha$ -glucan and glycogen from *Mycobacterium bovis*', *Glycobiology*, pp. 502-508.



Dye, C. (2009) 'Doomsday postponed? Preventing and reversing epidemics of drug-resistant tuberculosis', *Nature Reviews Microbiology*, pp. 81-87.

Ehlers , M.R.W. and Daffe, M. (1998) 'Interactions between *Mycobacterium tuberculosis* and host cells: are mycobacterial sugars the key?', *Trends In Microbiology*, vol. 6, pp. 328-335.

Elbain, A.D., Pan, Y.T., Pastuszak, I. and Carroll, D. (2003) 'New insights on trehalose: a multifunctional molecule ', *Glycobiology*, vol. 13, pp. 17-27.

Elbein, A.D., Pastuszak, I., Tackett, A.J., Wilson, T. and Pan, Y.T. (2010) 'Last step in the conversion of trehalose to glycogen', *The Journal of Biological Chemistry*, pp. 9803-9812.

Ericsson, U.B., Hallberg, M., DeTitta, G.T., Dekker, N. and Nordlund, P. (2006) 'Thermofluor-based high-throughput stability optimization of proteins for structural studies', *Analytical Biochemistry*, vol. 357, pp. 289-298.

Ernst, J.D. (1998) 'Macrophage receptors for *Mycobacterium tuberculosis*', *infection and immunity*, vol. 66, pp. 1277-1281.

Finn, R.D., Bateman, A., Clements , J., Cogill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L.L., Tate, J. and Punta, M. (2014) 'Pfam: the protein families database', *Nucleic Acid Research*, vol. 42, pp. 222-230.

Fraga, J., Maranha, A., Mendes, V., Barbosa Pereira, P.J., Empadinhas, N. and Macedo-Ribeiro, S. (2015) 'Structure of mycobacterial maltokinase, the missing link in the essential GlgE-pathway', *Scientific Reports*, vol. 26, p. 8026.

Franceschini, A., Szklarczyk, D., Frankild, S., Huhn, M., Simonovic , M., Roth, A., Lin, J., Minguez , P., Bork, P., von Mering, C. and Jensen , L.J. (2013) 'STRING v9.1: protein-protein interaction networks, with increased coverage and integration', *Nucleic Acids Research*, pp. 808-815.

Franke, D. and Svergun, D.I. (2009) 'Rapid ab initio shape determination by simulated annealing using a single phase dummy atom model', *Journal of Applied Crystallography*, vol. 42, pp. 342-346.

Freyer, M.W. and Lewis, E.A. (2008) 'Isothermal Titration Calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions', *Methods Cell Biology*, vol. 84, pp. 79-113.

Gatfield, J. and Pieters, J. (2000) 'Essential role for cholesterol in entry of mycobacteria into macrophages', *Science*, vol. 288, pp. 1647-1650.

Geurtsen, J., Chedammi, S., Mesters, J., Cot, M., Driessen, N.N., Sambou, T., Kakutani, R., Ummels, R., Maaskant, J., Takata, H., Baba, O., Terashima, T., Bovin, N., Vandenbroucke-Grauls, C., Nigou, J., Puzo, G., Lemassu, A., Daffé, M. and Appelmelk, B.J. (2009) 'Identification of mycobacterial  $\alpha$ -glucan as a novel ligand for DC-SIGN: involvement of mycobacterial capsular polysaccharides in host immune modulation', *The Journal of Immunology*, vol. 183, pp. 5221-5231.

Glickman, M.S., Cox, J.S. and Jacobs, W.R. (2000) 'A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of *Mycobacterium tuberculosis*', *Molecular Cell*, vol. 5, pp. 717-727.

Herzog, H. (1998) 'History of tuberculosis', *Respiration*, vol. 65, pp. 5-15.

Holm, L. and Rosenstrom, P. (2010) 'Dali server: conservation mapping in 3D', *Nucleic Acids Research*, vol. 38, pp. W545-W549.

Jacques, D.A. and Trewhella, J. (2010) 'Small-angle scattering for structural biology—Expanding the frontier while avoiding the pitfalls', *Protein Science*, vol. 19, pp. 642-657.

Jose' Barbosa Pereira, P., Empadinhas, , Albuquerque, , Sa'-Moura, , da Costa, M.S. and Macedo-Ribeiro, (2008) '*Mycobacterium tuberculosis* Glucosyl-3- Phosphoglycerate Synthase: Structure of a Key Enzyme in Methylglucose Lipopolysaccharide Biosynthesis', *PLoS One*, vol. 3, p. e3748.

- Kalscheuer, R., Syson, K., Veeraraghavan, U., Weinrick, B., Biermann, K.E., Sacchettini, J.C., Besra, G., Bornemann, S. and Jacobs Jr, W.R. (2010) 'Self-poisoning of *Mycobacterium tuberculosis* by targeting GlgE in an  $\alpha$ -glucan pathway', *Nature Chemical Biology*, pp. 376-384.
- Karimova, G., Pidoux, J., Ullmann, A. and Ladant, D. (1998) 'A bacterial two-hybrid system based on a reconstituted signal transduction pathway', *PNAS*, pp. 5752-5756.
- Kaufmann, S.H.E. (2001) 'How can immunology contribute to the control of tuberculosis?', *Nature Reviews, Immunology*, pp. 20-30.
- Kaur, , Pham, , Larrouy-Maumus, , Riviere, , Vissa, , Guerin, E., Puzo, , Brennan, J. and Jackson, (2009) 'Initiation of Methylglucose Lipopolysaccharide Biosynthesis in *Mycobacteria*', *PLoS One*, vol. 4, pp. 1-7.
- Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H. and Svergun, D.I. (2003) 'PRIMUS: a Windows PC-based system for small-angle scattering data analysis', *Journal of Applied Crystallography*, vol. 36, pp. 1277-1282.
- Koul, A., Arnoult, E., Lounis, N., Guillemont, J. and Andries, K. (2011) 'The challenge of new drug discovery for tuberculosis', *Nature*, vol. 469, pp. 483-490.
- Lamichhane, G. (2011) 'Novel targets in *M. tuberculosis*: search for new drugs', *Trends in Molecular Medicine*, vol. 17, pp. 25-33.
- Li, J., Guan, X., Shaw, N., Chen, W., Dong, Y., Xu, X., Li, X. and Rao, Z. (2014) 'Homotypic dimerization of a maltose kinase for molecular scaffolding', *Scientific Reports*, vol. 4, pp. 1-8.
- Lindenberger, J.J., Veleti, S.K., Wilson, B.N., Sucheck, S.J. and Ronning, R. (2015) 'Crystal structure of *Mycobacterium tuberculosis* GlgE and complexes with non-covalent inhibitors', *Scientific Reports*, vol. 5, p. 12830.

Mahapatra, S., Crick, D.C., McNeil, M.R. and Brennan, P.J. (2008) 'Unique structural features of the peptidoglycan of *Mycobacterium leprae*', *Journal of Bacteriology*, vol. 190, pp. 655-661.

Makarov, V., Manina, G., Mikusova, K., Möllmann, U., Ryabova, O., Saint-Joanis, B., Dhar, N., Pasca, M.R., Buroni, S., Lucarelli, A.P., Milano, A., De Rossi, E., Belanova, M., Bobovska, A., Dianiskova, P., Kordulakova, J., Sala, C., Fullam, E., Schneider, P. and McKinney, J.D. (2009) 'Benzothiazinones kill *Mycobacterium tuberculosis* by blocking arabinan synthesis', *Science*, vol. 324, pp. 801-804.

Matsumoto, M., Hashizume, H., Tomishige, T., Kawasaki, M., Tsubouchi, H., Sasaki, H., Shimokawa, Y. and Komatsu, M. (2006) 'OPC-67683, a Nitro-Dihydro-Imidazoazole derivative with promising action against tuberculosis in vitro and in mice', *PLOS Medicine*, vol. 3, pp. 2131-2144.

McCarter, J.D. and Withers, S.G. (1994) 'Mechanisms of enzymatic glycoside hydrolysis', *Current Opinion in Structural Biology*, vol. 4, pp. 885-892.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) 'Phaser crystallographic software', *Journal of Applied Crystallography*, vol. 40, pp. 658-674.

Mendes, V., Maranha, A., Lamosa, P., Costa, M., and Empadinhas, N. (2010) 'Biochemical characterization of the maltokinase from *Mycobacterium bovis* BCG', *BMC Biochemistry*, pp. 1-10.

Mendes, V., Maranha, A., Alarico, S., da Costa, M.S. and Empadinhas, N. (2011) '*Mycobacterium tuberculosis* Rv2419c, the missing glucosyl-3-phosphoglycerate phosphatase for the second step in methylglucose lipopolysaccharide biosynthesis', *Scientific Reports*, vol. 1, pp. 1-8.

Miah , F., Koliwer-Brandl, H., Rejzek, M., Field, R.A., Kalscheuer , R. and Bornemann, S. (2013) 'Flux through trehalose synthase flows from trehalose to the alpha anomer of maltose in mycobacteria', *Chemistry and Biology* , pp. 487-493.

Niesen, F.H., Berglund, H. and Vedadi, M. (2007) 'The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability', *Nature Protocols*, vol. 2, pp. 2212-2221.

Pal, K., Kumar , S., Sharma, S., Garg , S.K., Alam, M.S., Xu, H.E., Agrawal, P. and Swaminathan, K. (2010) 'Crystal structure of full-length *Mycobacterium tuberculosis* H37Rv glycogen branching enzyme ', *The Journal of Biological Chemistry*, pp. 20897-20903.

Pan, Y.-T., Carroll, J.D., Asano , N., Pastuszak, I., Edavana , V.K. and Elbein, A.D. (2008) 'Trehalose synthase converts glycogen to trehalose ', *FEBS*, pp. 3408-3420.

Park, S.-H. and Bendelac, A. (2000) 'CD1-restricted T-cell responses and microbial infection', *Nature*, vol. 406, pp. 788-792.

Pierce, M.M., Raman, C.S. and Nall, B.T. (1999) 'Isothermal Titration Calorimetry of protein–protein Interactions', *Methods*, vol. 19, pp. 213-221.

Pieters, J. (2008) '*Mycobacterium tuberculosis* and the macrophage: maintaining a balance', *Cell Host & Microbe*, vol. 3, pp. 399-407.

Public Health England (2014) 'Tuberculosis in the UK: 2014 report'.

Roy, , Usha, , Kermani, , Scott, J., Hyde, E.I., Besra, S., Alderwick, J. and Fütterer , (2013) 'Synthesis of  $\alpha$ -glucan in mycobacteria involves a hetero-octameric complex of trehalose synthase TreS and maltokinase Pep2', *American Chemical Biology*, vol. 8, pp. 2245-2255.

Russell, D.G. (2007) 'Who puts the tubercle in tuberculosis?', *Nature Microbiology Reviews*, pp. 39-47.

Saito, K., Kase, T., Takahashi, E., Takahashi, E. and Horinouchi, S. (1998) 'Purification and characterization of a trehalose synthase from the basidiomycete *Grifola frondosa*', *Appl. Environ. Microbiol.*, vol. 64, pp. 4340-4345.

Sambou, T., Dinadayala, P., Stadthagen, G., Barilone, N., Bordat, Y., Constant, P., Levillain, F., Neyrolles, O., Gicquel, B., Lemassu, A., Daffé, M. and Jackson, M. (2008) 'Capsular glucan and intracellular glycogen of *Mycobacterium tuberculosis*: biosynthesis and impact on the persistence in mice', *Molecular Microbiology*, vol. 70, pp. 762-774.

Sani, M., Houben, E.N.G., Geurtsen, J., Pierson, J., Punder, K., van Zon, M., Wever, B., Piersma, S., Jimenez, C.R., Daffe, M., Appelmelk, B., Bitter, W., van der Wel, N. and Peters, P.J. (2010) 'Direct visualization by cryo-EM of the mycobacterial capsular layer: a labile structure containing ESX-1-secreted proteins', *PLOS Pathogens*, pp. 1-10.

Santos, C.R., Tonoli, C.C.C., Trindade, D.M., Betzel, C., Takata, H., Kuriki, T., Kanai, T., Imanaka, T., Arni, R.K. and Murakami, M.T. (2011) 'Structural basis for branching-enzyme activity of glycoside hydrolase family 57: structure and stability studies of a novel branching enzyme from the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1', *Proteins*, vol. 79, no. 547-557.

Sareen, D., Steffek, M., Newton, L.G. and Fahey, R.C. (2002) 'ATP-dependent L-cysteine:1D-myo-inositol 2- amino-2-deoxy-alpha-D-glucopyranoside ligase, mycothiol biosynthesis enzyme MshC, is related to class I cysteinyl-tRNA synthetases', *Biochemistry*, pp. 6885-6890.

Sasseti, C.M., Boyd, D.H. and Rubin, J. (2003) 'Genes required for mycobacterial growth defined by high density mutagenesis', *Molecular Microbiology*, vol. 48, pp. 77-84.

Senaratne, R., Mobasher, H., Papavinasasundaram, K., Jenner, P., Lea, E.J.A. and Draper, P. (1998) 'Expression of a gene for a porin-like protein of the OmpA family from *Mycobacterium tuberculosis* H37Rv', *Journal of Bacteriology*, pp. 3541-3547.

Stadthagen, G., Sambou, T., Guerin, M., Barilone, N., Boudou, F.r., Kordula´kova´, J., Charles, P., Alzari, P., Lemassu, A., Daffe´, M., Puzo, G., Gicquel, B., Rivie`re, M. and Jackson, M. (2007) 'Genetic basis for the biosynthesis of methylglucose lipopolysaccharides in *Mycobacterium tuberculosis*', *The Journal Of Biological Chemistry*, vol. 282, pp. 27270–27276.

Stoop, E.J.M., Bitter, W. and van der Sar, A.M. (2012) 'Tubercle bacilli rely on a type VII army for pathogenicity', *Trends in Microbiology*, vol. 20, pp. 477-484.

Stover, C.K., Warrenner, P., VanDevanter, D.R., Sherman, D.R., Arain, T.M., Langhorne, M.H., Anderson, S.W., Towell, J.A., Yuan, Y., McMurray, D.N., Krelswirth, B.N., Barry, C.E. and Baker, W.R. (2000) 'A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis', *Nature*, vol. 405, pp. 962-966.

Svergun, D.I., Barberato, C. and Koch, M.H.J. (1995) 'Evaluation of the solution scattering from macromolecules with known atomic structure and fitting to experimental data', *Journal of Applied Crystallography*, vol. 28, pp. 768-773.

Syson, K., Stevenson, C.E.M., Rejzek, M., Fairhurst, S.A., Nair , A., Bruton, C.J., Field, R.A., Chater, K.F., Lawson, D.M. and Bornemann, S. (2011) 'Structure of Streptomyces maltosyltransferase GlgE, a homologue of a genetically validated anti-tuberculosis target', *The Journal of Biological Chemistry*, pp. 38298-38310.

Tahlan, K., Wilson, R., Kastrinsky, D.B., Arora, K., Nair, V., Fischer, E., Barnes, S.W., Walker, J.R., Alland, D., Barry , C.E. and Boshoff, H.I. (2012) 'SQ109 targets MmpL3, a membrane transporter of trehalose monomycolate involved in mycolic acid donation to the cell wall core of *Mycobacterium tuberculosis*', *Antimicrobial Agents and Chemotherapy*, vol. 56, pp. 1797-1809.

Takayama, K., Wang, C. and Besra, S. (2005) 'Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*', *Clinical Microbiology Reviews*, vol. 18, pp. 81-101.

Tam, P.-H. and Lowary, T.L. (2009) 'Recent advances in mycobacterial cell wall glycan biosynthesis', *Current Opinion in Chemical Biology*, vol. 13, pp. 618-625.

Teo, J.W.P., Thayalan, P., Beer, D., Yap, A.S.L., Nanjundappa, M., Ngew, X., Duraiswamy, J., Liung, S., Dartois, V., Schreiber, M., Hasan, S., Cynamon, M., Ryder, N.S., Yang, X., Weidmann, B., Bracken, K., Dick, T. and Mukherjee, K. (2006) 'Peptide deformylase inhibitors as potent antimycobacterial agents', *Antimicrobial Agents And Chemotherapy*, vol. 50, pp. 3665-3673.

Volkov, V.V. and Svergun, D.I. (2003) 'The program suite DAMAVER is a set of programs to align ab initio models, select the most typical one and build an averaged model', *Journal of Applied Crystallography*, vol. 36, pp. 860-864.

Weerd, R.v.d., Berbis, M.A., Sparrius, M., Maaskant, J.J., Boot, M., Paauw, N.J., Vries, N.d., Boon, L., Baba, O., Canada, F.J., Geurtsen, J., Jimenez-Barbero, J. and Appelmelk, B.J. (2015) 'A murine monoclonal antibody to glycogen: characterization of epitope-fine specificity by Saturation Transfer Difference (STD) NMR spectroscopy and its use in mycobacterial capsular  $\alpha$ -glucan research', *ChemBioChem*, vol. 16, pp. 977-989.

Zhang, R., Pan, Y.T., He, S., Lam, M., Brayer, G.D., Elbein, A.D. and Withers, S.G. (2011) 'Mechanistic analysis of trehalose synthase from *Mycobacterium smegmatis*', *The Journal of Biological Chemistry*, pp. 35601-35609.

Zumla, A., Nahid, P. and Cole, S.T. (2013) 'Advances in the development of new tuberculosis drugs and treatment regimens', *Nature Reviews*, pp. 388-404.



## **9 Supplements**

## Media used in this study:

Terrific Broth (TB) for 900 mL

Trypton            12 g

Yeast Extract    24 g

Glycerol           0.4%

TB salt for 100 mL

$\text{KH}_2\text{PO}_4$            2.3 g

$\text{K}_2\text{HPO}_4$           12.5 g

Both TB and TB salt should be autoclaved separately and mixed together just before inoculation.

MB7H9 media for 900 mL

4.7 g    Difco MB7H9

0.5%    Glycerol

Autoclave at 121 °C for 10 min and add 100 mL of Albumin Dextrose Catalase (ADC) to make it to 1 L.

**Table S9-1.** All approved drugs to treat tuberculosis (Brennan, Young and Robertson, 2008).

Drug name	Drug target/mechanism	Drug resistance mechanism	In vitro activity
Clarithromycin (CLA)	Inhibits protein biosynthesis through binding to the 50S ribosomal subunit	Methylation of the 23S rRNA	Active against <i>M. tuberculosis</i> (H37Rv) and <i>M. avium</i> with MIC 8 µg/ml at pH 7.4 and MIC <sub>90</sub> 8 µg/ml, respectively
Clofazimine (CLOF)	Mode of action is not identified	Not very well known	Active against <i>M. tuberculosis</i> (H37Rv) with MIC 0.1 µg/ml
Gatifloxacin (GATI)	Fluoroquinolone, see MOXI		MIC against <i>M. tuberculosis</i> (H37Rv) 0.25 µg/ml
Levofloxacin (LEV)	Second generation quinolone, see MOXI	See MOXI	MIC <i>M. tuberculosis</i> (H37Rv) 0.5 µg/ml
Linezolid (LIN)	Inhibits protein synthesis through binding to 23S rRNA, preventing proper binding of formyl-methionine tRNA	Mutations in 23S rRNA	MIC <i>M. tuberculosis</i> (H37Rv) 0.25 µg/ml
LL-3858	Not known		MIC <i>M. tuberculosis</i> (H37Rv) 0.12-0.025 µg/ml
Moxifloxacin (MOXI)	As a quinolone inhibits ATP-dependent enzymes topoisomerase II (DNA gyrase) and topoisomerase IV encoded by <i>gyrA</i> , <i>gyrB</i> ; and <i>parC</i> and <i>parE</i> , respectively	Mutations in <i>gyrA</i> or <i>gyrB</i>	MIC against <i>M. tuberculosis</i> (H37Rv) 0.5 µg/ml

Prothionamide (PRO)	See ethionamide (ETA)	Cross resistance between PRO and ETA	MIC against <i>M. tuberculosis</i> (H37Rv) 0.5 µg/ml
Rifabutin (RIFAB)	See Rifampin (RIF)	Mutations in <i>rpoB</i> gene	MIC against <i>M. tuberculosis</i> (H37Rv) <0.015 µg/ml
Rifalazil (RIFAL)	See Rifampin (RIF)	Mutations in <i>rpoB</i> gene	MIC against <i>M. tuberculosis</i> (H37Rv) <0.015 µg/ml
Rifampin (RIF)	Inhibits the RNA polymerase β-subunit encoded by <i>rpoB</i> gene	Mutations in <i>rpoB</i> gene	MIC against <i>M. tuberculosis</i> (H37Rv) 0.1-0.39 µg/ml
Rifapentine (RIFAP)	See Rifampin (RIF)	Mutations in <i>rpoB</i> gene	MIC against <i>M. tuberculosis</i> (H37Rv) 0.031 µg/ml
Thioridazine	Inhibiting calmodulin or NADH2-menaquinone-oxidoreductase (Ndh2)		MIC against <i>M. tuberculosis</i> ATCC 27294 10 µg/ml
TMC-207	Targets c subunit of ATP synthase encoded by <i>atpE</i> gene	Mutations in <i>atpE</i> gene	MIC against <i>M. tuberculosis</i> (H37Rv) 0.06 µg/ml (0.03-0.12 µg/ml)

**Table S9-2.** List of new anti-TB drugs.

Drug name	Drug target/mechanism	Drug resistance mechanism	In vitro activity	Reference
Bedaquiline (R207910)	Targets ATP synthase	Point mutations in <i>atpE</i> gene	Inhibits both drug-sensitive and drug-resistant <i>M. tuberculosis</i> in vitro (MIC 0.06 µg/mL)	Andries et al., 2005
OPC-67683 (Delamanid)	Inhibits mycolic acid biosynthesis	Mutation in <i>Rv3547</i> , the gene responsible for activating the drug	Inhibits both drug-sensitive and drug-resistant <i>M. tuberculosis</i> in vitro (MIC 0.0006-0.024 µg/mL)	Matsumoto et al., 2006
PA-824	Inhibits mycolic acid and protein biosynthesis	Mutation in <i>Rv3547</i> , the gene responsible for activating the drug Point mutations in <i>Rv0407</i> encoding F420-dependent glucose 6-phosphate dehydrogenase (Fgd)	MIC against <i>M. tuberculosis</i> (H37Rv) 0.15-0.3 µg/ml	Stover et al., 2000
SQ109	Inhibits TDM synthesis by inhibiting MmpL3, the transmembrane transporter responsible for transporting precursor of TDM	Mutations in <i>mmpL3</i> , the gene encoding MmpL3 transmembrane transporter		Tahlan et al., 2012

**Table S9-3.** List of recently identified targets in *M. tuberculosis*. The targets, their roles within the cells, their inhibition impacts on bacilli and inhibitor (if any) are shown.

Target	Activity	Impacts on <i>M. tuberculosis</i>	Inhibitor	Reference
<b>GlgE</b>	Adds maltose 1-P to the growing 1,4 glucan chain	Deletion of <i>glgE</i> gene leads to toxic accumulation of M1-P Synthetic lethality with Rv3032		Kalscheuer et al., 2010
<b><i>pcaA</i></b>	Cyclopropane synthase involved in the formation of mycolic acids	Mutant bacteria fail to persist within mice	Diocetylamine	Glickman, Cox and Jacobs, 2000
<b>DprE1/DprE2</b>	Converts decaprenylphosphoryl ribose into decaprenylphosphoryl arabinose	Inhibits the production of arabinogalactan	Benzothiazinone Dinitrobenzamides	Makarov et al., 2009
<b>MshC</b>	Encodes an ATP-dependent mycothiol ligase that catalyses the last step in the biosynthesis of mycothiol		Dequalinium chloride	Sareen et al., 2002
<b>Peptide deformylase (Def)</b>	Deformylates the first methionine at N terminus		LBK611	Teo et al., 2006

**Table S9-4. X-ray diffraction data and refinement statistics.**

<b>X-ray diffraction data</b>	
Crystal	Tres:Pep2
Beamline	ID29, ESRF
Wavelength (Å)	0.97717
Space group	$P4_3$
Cell parameters (Å)	315.68, 315.68, 124.95
TreS:Pep2 complexes per asymmetric unit	2
Resolution range (Å)	49.91 - 3.60
High resolution shell (Å)	3.66 - 3.60
Rmerge (%) <sup>1)</sup>	17.0 (154.2)
Total observations, unique reflections	1090211, 142362
I/σ(I) <sup>1)</sup>	9.1 (1.1)
Completeness (%) <sup>1)</sup>	99.8 (98.1)
Multiplicity <sup>1)</sup>	7.7 (7.0)
CC <sub>1/2</sub> <sup>1), 2)</sup>	0.997 (0.343)
<b>Refinement</b>	
Resolution range	49.9 - 3.6
Unique reflections	135356
R <sub>cryst</sub> , R <sub>free</sub> (%)	26.1, 28.1
No of non-H atoms	56072
RMSD bonds (Å)	0.006
RMSD angles (°)	1.0
B-factors	
Wilson (Å <sup>2</sup> )	126
Average overall (Å <sup>2</sup> )	111
RMSD B-factors (Å <sup>2</sup> )	1.36
Ramachandran statistics <sup>3)</sup>	
Favoured regions (%)	94.9
Allowed regions (%)	4.5
Disallowed (%)	0.6
Clash score	4.6
Rotamer outliers (%)	1.1