

**Molecular and biochemical characterisation of
glycosyltransferases involved in cell wall assembly
of *Corynebacterineae***

By

MONIKA JANKUTE

A thesis submitted to the University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Biosciences

College of Life and Environmental Sciences

University of Birmingham

September 2014

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This unpublished thesis/dissertation is copyright of the author and/or third parties.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.

Abstract

Mycobacterium tuberculosis, the etiological agent of tuberculosis, remains the leading cause of mortality from a single infectious organism. The persistence of this human pathogen is associated with its distinctive lipid rich cell wall structure that is highly impermeable to hydrophilic drugs. This highly complex and unique structure is crucial for the growth, viability and virulence of *M. tuberculosis*, thus representing an attractive target for vaccine and drug development. In this study, we have demonstrated that enzymes involved in *Corynebacterium glutamicum* cell wall assembly and precursor formation build complicated multi-protein complexes. Specifically, we have identified 24 putative interactions *in vivo* between 12 proteins responsible for AG biosynthesis. Additionally, we have demonstrated that AftB arabinosyltransferase have a dual functionality and plays a role in both AG and LAM biosynthesis in *M. smegmatis*. We have also demonstrated that Emb arabinosyltransferase, the target of ethambutol, possesses $\alpha(1\rightarrow5)$ activity in *C. glutamicum*. Finally, we examined the overexpression and biophysical characterisation of membrane AftA solubilised in a styrene maleic acid polymer, demonstrating that detergent free extraction of membrane proteins from corynebacteria is possible. These findings provide a useful recourse for understanding the biosynthesis and function of the vital cell wall (lipo)polysaccharides of *Corynebacterineae*, as well as providing new therapeutic targets for drug design against the pathogenic species of *M. tuberculosis*, *Mycobacterium marinum* and *Corynebacterium diphtheriae*.

Declaration

The work presented in this thesis has been carried out under the supervision of Prof. G. S. Besra and Dr. L. J. Alderwick in the School of Biosciences at the University of Birmingham, UK, B15 2TT during the period September 2010 to September 2014. The work in this thesis is original except where acknowledged by reference. The thesis resulted in several publications, two with Dr. S. Grover as the co-author and one as the first author that may bear similarities to chapters 1, 2 and 4. These similar sections included in my thesis are my original work with no significant contribution to the text by Dr. Shipra Grover, Prof. G. S. Besra and Dr. Luke Alderwick. No part of the work is being, or has been submitted for a degree, diploma or any other qualification at any other University.

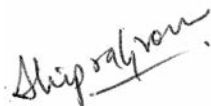
Prof. G. S. Besra



Dr. Luke Alderwick



Dr. Shipra Grover



Acknowledgments

I would like to thank my supervisors, Prof. G. S. Besra and Dr. L. J. Alderwick for their support, guidance and enthusiasm over the duration of my PhD. I must also thank Dr. A. Bhatt for always readily offering his expertise and advice at every point of my PhD.

I would like to thank Dr. L. Eggeling and his laboratory from Jülich Research Center, Germany for construction of mutant strains in *Corynebacterium glutamicum* that appears in this thesis and Prof. A. Dell from Department of Life Sciences, Imperial College London for performing few of the analytical experiments. Also, I would like to thank Prof. T. Dafforn and Dr. M. Jamshad for their help using lipid nanoparticles and Dr. I. Cadby for the protein thermostability assays. Very special thanks to Prof. T. L. Lowary from University of Alberta, Canada for providing synthetic acceptors and Prof. B. J. Appelmeik from VU University Medical Center, Amsterdam for providing monoclonal antibodies.

Finally, I would like to thank my parents, my brother and my fiancé for always having faith in me and providing a boost of confidence when I most needed it. Also, my special thanks to everyone in the Besra Lab, you know who you are, for your constant support, advice and friendship.

This thesis is dedicated to my parents.

Table of contents

University of Birmingham Research Archive.....	i
Abstract.....	ii
Declaration.....	iii
Acknowledgments.....	iv
Dedication.....	v
Table of Contents.....	vi
List of Figures.....	xii
List of Tables.....	xvii
List of Abbreviations.....	xviii
Published work associated with the thesis.....	xxiv
1 General Introduction.....	1
1.1 Mycobacteria and related species.....	2
1.2 History of tuberculosis.....	4
1.3 Global epidemiology of tuberculosis.....	7
1.4 Clinical manifestations.....	10
1.5 Diagnosing tuberculosis.....	10
1.6 Pathogenesis of tuberculosis.....	12
1.7 Tuberculosis drug treatments.....	15
1.8 Mycobacterial cell wall.....	17
1.8.1 Peptidoglycan.....	19
1.8.2 Arabinogalactan.....	21
1.8.2.1 Structural features of arabinogalactan.....	21
1.8.2.2 Precursor formation.....	24
1.8.2.3 Biosynthesis of arabinogalactan.....	30
1.8.2.4 Decoration of arabinogalactan and its attachment to peptidoglycan and mycolic acids.....	33

1.8.3	Lipoarabinomannan.....	35
1.8.3.1	Structure of phosphatidyl- <i>myo</i> -inositol mannosides, lipomannan and lipoarabinomannan.....	36
1.8.3.2	Biosynthesis of phosphatidyl- <i>myo</i> -inositol mannosides, lipomannan and lipoarabinomannan.....	39
1.8.4	Ethambutol inhibits arabinogalactan and lipoarabinomannan biosynthesis	44
1.8.5	Mycolic acids.....	48
1.9	Surrogate systems in <i>M. tuberculosis</i> research.....	50
1.10	Project aims.....	53
2	Elucidation of a protein-protein interaction network involved in <i>Corynebacterium glutamicum</i> cell wall biosynthesis as determined by bacterial two-hybrid analysis.....	55
2.1	Introduction.....	56
2.2	Materials and methods.....	62
2.2.1	Construction of vectors overexpressing GlfT2 and AftA.....	62
2.2.2	Pull-down detection of putative GlfT2 and AftA binding proteins	62
2.2.3	Mass spectrometry and analysis of putative GlfT2 and AftA binding proteins	63
2.2.4	Plasmid construction for BACTH analysis.....	63
2.2.5	Bacterial two-hybrid system.....	66
2.2.6	β -Galactosidase assay.....	68
2.2.7	Statistical analysis.....	69
2.3	Results.....	69
2.3.1	Network analysis of AG biosynthetic proteins.....	69
2.3.2	Pull-down assay and detection of putative GlfT2 binding proteins from <i>M. smegmatis</i>	71
2.3.3	Pull-down assay and detection of putative AftA binding proteins from <i>C. glutamicum</i>	75
2.3.4	Bacterial two-hybrid analysis of AG proteins.....	80
2.3.5	Self-association of <i>C. glutamicum</i> cell wall biosynthesis proteins.....	82
2.3.6	<i>In vivo</i> interaction network among AG proteins.....	84
2.3.7	Verification of recombinant protein expression.....	85

2.4	Discussion.....	88
3	Molecular and biochemical characterisation of AftA and AftB arabinosyltransferases from <i>Corynebacterineae</i>.....	92
3.1	Introduction.....	93
3.2	Materials and methods.....	98
3.2.1	Construction of merodiploid strains	98
3.2.2	Construction of <i>aftA</i> and <i>aftB</i> deletion mutants in <i>M. smegmatis</i>	99
3.2.3	Plasmid construction for overexpression of full-length AftA and AftB from <i>M. tuberculosis</i> and <i>C. glutamicum</i>	100
3.2.4	Plasmid construction for overexpression of C-terminal domain of AftA and AftB from <i>M. tuberculosis</i>	101
3.2.5	Genomic DNA extraction from <i>M. smegmatis</i>	101
3.2.6	Southern blot analysis.....	102
3.2.7	Conditional depletion of AftA and AftB.....	102
3.2.8	Protein production.....	103
3.2.9	Purification of <i>M. tuberculosis</i> AftA ^{CT}	103
3.2.10	Differential scanning fluorimetry with AftA ^{CT}	104
3.2.11	AftA ^{CT} crystallisation trials.....	105
3.2.12	Styrene maleic acid solubilisation.....	105
3.2.13	Purification of <i>C. glutamicum</i> AftA within SMALP.....	106
3.2.14	Circular dichroism spectroscopy.....	106
3.2.15	Thermostability analysis of <i>C. glutamicum</i> AftA within SMALP.....	107
3.2.16	Tryptophan fluorescence binding assay.....	107
3.2.17	Analytical ultracentrifugation.....	107
3.3	Results.....	108
3.3.1	Molecular and biochemical characterisation of AftA.....	108
3.3.1.1	Evidence suggesting that AftA from <i>M. smegmatis</i> is an essential gene.....	108
3.3.1.2	Overexpression and purification of recombinant AftA ^{CT}	111
3.3.1.3	Differential scanning fluorimetry profile of AftA ^{CT}	115
3.3.1.4	Crystallisation trials of <i>M. tuberculosis</i> AftA ^{CT}	118
3.3.1.5	Overexpression studies of recombinant AftA from <i>M. tuberculosis</i> and <i>C. glutamicum</i>	119

3.3.1.6	Complementation of <i>C. glutamicum</i> Δ <i>aftA</i> mutant with <i>C. glutamicum</i> AftA to verify its enzyme function.....	121
3.3.1.7	.Purification of SMALP encapsulated AftA from <i>C. glutamicum</i>	124
3.3.1.8	Biophysical characterisation of SMALP encapsulated AftA from <i>C. glutamicum</i>	125
3.3.1.9	Ligand binding assays of AftA from <i>C. glutamicum</i>	126
3.3.2	Molecular and biochemical characterisation of AftB.....	128
3.3.2.1	Evidence suggesting that AftB from <i>M. smegmatis</i> is an essential gene.....	128
3.3.2.2	Lipid characterisation of the <i>M. smegmatis</i> Δ <i>aftB</i> ::pMV306- <i>aftB</i> conditional mutant.....	130
3.3.2.3	Lipoglycan analysis of <i>M. smegmatis</i> Δ <i>aftB</i> ::pMV306- <i>aftB</i> conditional mutant.....	133
3.3.2.4	Role of AftB in the synthesis of hexa-arabinan motif of LAM.....	135
3.3.2.5	Overexpression studies of recombinant AftB from <i>M. tuberculosis</i> and <i>C. glutamicum</i>	136
3.4	Discussion.....	139
4	Molecular and biochemical characterisation of Emb arabinosyltransferase from <i>Corynebacterium glutamicum</i>.....	144
4.1	Introduction.....	145
4.2	Materials and Methods.....	148
4.2.1	Construction of plasmids and strains.....	148
4.2.2	Extraction and purification of arabinogalactan from <i>C. glutamicum</i>	150
4.2.3	NMR spectroscopic analysis of AG from <i>C. glutamicum</i> and <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i>	150
4.2.4	Extraction and purification of lipoglycans from <i>C. glutamicum</i> and <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i>	151
4.2.5	Preparation of <i>C. glutamicum</i> membranes.....	151
4.2.6	Preparation of <i>C. glutamicum</i> P60 cell wall material.....	152
4.2.7	[¹⁴ C]-pRpp synthesis.....	152
4.2.8	Arabinofuranosyltransferase activity using membrane preparations of <i>C. glutamicum</i> , <i>C. glutamicum</i> Δ <i>aftA</i> , <i>C. glutamicum</i> Δ <i>emb</i> , <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i> and <i>C. glutamicum</i> Δ <i>aftB</i> Δ <i>aftD</i>	153
4.2.9	Analysis of arabinofuranosyltransferase reaction products prepared from <i>C. glutamicum</i> and <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i> membranes.....	154

4.3	Results.....	155
4.3.1	Lipid characterisation of <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i> mutant.....	155
4.3.2	Structural characterisation of <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i>	158
4.3.3	Lipoglycan analysis of <i>C. glutamicum</i> and <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i>	162
4.3.4	<i>In vitro</i> arabinofuranosyltransferase activity with extracts of <i>C. glutamicum</i> , <i>C. glutamicum</i> Δ <i>aftA</i> , <i>C. glutamicum</i> Δ <i>emb</i> , <i>C. glutamicum</i> Δ <i>aftB</i> Δ <i>aftD</i> and <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i>	163
4.3.5	Characterisation of arabinosyltransferase reactions products.....	166
4.4	Discussion.....	171
5	Conclusions and Future work.....	177
6	General Materials and Methods.....	183
6.1	Chemicals and reagents.....	184
6.2	Culture media preparations.....	184
6.2.1	Luria-Bertani broth.....	184
6.2.2	Luria-Bertani agar.....	184
6.2.3	Brain Heart Infusion broth.....	184
6.2.4	Brain Heart Infusion agar.....	185
6.2.5	Tryptic soy broth.....	185
6.2.6	Tryptic soy agar.....	185
6.2.7	7H9 basal agar	185
6.2.8	7H9 soft agar.....	186
6.2.9	MacConkey agar.....	186
6.2.10	CGXII minimal media.....	186
6.2.11	M63 minimal media.....	187
6.2.12	M63 minimal agar.....	187
6.2.13	Sauton's minimal media.....	187
6.2.14	Minimal media for <i>M. smegmatis</i>	188
6.2.15	Minimal agar for <i>M. smegmatis</i>	188
6.3	Bacterial strains and growth conditions.....	188
6.3.1	General bacterial strains and growth conditions used in this thesis.....	188
6.3.2	Bacterial strains and growth conditions used for BACTH.....	189

6.4	Antibiotics and supplements.....	190
6.5	Molecular biology techniques.....	190
6.5.1	Polymerase chain reaction.....	190
6.5.2	DNA gel electrophoresis.....	191
6.5.3	DNA purification.....	192
6.5.4	DNA digestion.....	192
6.5.5	DNA ligation.....	192
6.5.6	Plasmid DNA extraction.....	193
6.5.7	Preparation of competent <i>E. coli</i> cells.....	193
6.5.8	Preparation of competent <i>C. glutamicum</i> cells.....	194
6.5.9	Preparation of competent <i>M. smegmatis</i> cells.....	194
6.5.10	Transformation of <i>E. coli</i> competent cells.....	195
6.5.11	Transformation of <i>C. glutamicum</i> competent cells.....	195
6.5.12	Electroporation of <i>M. smegmatis</i> competent cells.....	195
6.5.13	Membrane fraction preparation.....	196
6.5.14	Extraction and analysis of cell wall associated lipids.....	196
6.5.15	Extraction and analysis of cell wall bound lipids.....	197
6.5.16	Purification of mAGP complex.....	197
6.5.17	Acid hydrolysis and alditol acetate derivatisation.....	198
6.5.18	Gas chromatography analysis.....	198
6.5.19	Extraction and purification of lipoglycans.....	198
6.5.20	Acid hydrolysis of LAM for sugar analysis.....	199
6.6	Protein biochemistry techniques.....	200
6.6.1	SDS-PAGE analysis.....	200
6.6.2	Western blot analysis.....	200
6.6.3	Protein expression and purification.....	200
	Appendix.....	202
	References.....	216

List of Figures

Figure 1.1	The genus <i>Mycobacterium</i> with examples of slow-growing and fast-growing bacilli.....	3
Figure 1.2	Estimated numbers of TB cases worldwide in 2012.....	8
Figure 1.3	Estimated HIV prevalence in new TB cases worldwide in 2012.....	9
Figure 1.4	Different outcomes of <i>M. tuberculosis</i> infection.....	14
Figure 1.5	Cell envelope of <i>M. tuberculosis</i>	18
Figure 1.6	Biosynthesis of peptidoglycan in <i>M. tuberculosis</i>	21
Figure 1.7	The general structure of arabinogalactan present in <i>M. tuberculosis</i>	23
Figure 1.8	The biosynthesis of sugar donors required for mycobacterial arabinogalactan biosynthesis.....	29
Figure 1.9	Schematic representation of mycobacterial arabinogalactan biosynthesis.....	33
Figure 1.10	Structural features of acylated phosphatidyl- <i>myo</i> -inositol hexamannoside, lipomannan and lipoarabinomannan.....	38
Figure 1.11	Biosynthesis of phosphatidyl- <i>myo</i> -inositol mannosides, lipomannan and lipoarabinomannan in <i>M. tuberculosis</i>	44
Figure 1.12	<i>M. tuberculosis</i> mycolic acid biosynthesis.....	50
Figure 2.1	Generalised scheme of a pull-down assay for detection of putative protein-protein interactions.....	60
Figure 2.2	Organisation of <i>Bordetella pertussis</i> adenylate cyclase and principle of detecting protein-protein interactions using BACTH.....	61
Figure 2.3	Network of <i>C. glutamicum</i> proteins found to be important for cell wall assembly as determined by STRING analysis.....	70
Figure 2.4	SDS-PAGE and Western blot analysis of <i>M. smegmatis</i> GlfT2 over-expression, purification, and pull-down assay.....	73

Figure 2.5	SDS-PAGE and Western blot analysis of <i>C. glutamicum</i> AftA over-expression, purification, and a pull-down assay.....	77
Figure 2.6	Single expression of hybrid proteins fused with T18 fragment as determined by Western blot analysis.....	86
Figure 2.7	Expression of GlfT2 fusions as determined by Western blot analysis.....	88
Figure 2.8	An interaction network of <i>C. glutamicum</i> proteins involved in AG biosynthesis.....	90
Figure 3.1	Transmembrane protein encapsulated in SMALP disc.....	97
Figure 3.2	Generation of conditional <i>M. smegmatis</i> Δ aftA::pMV306-aftA mutant....	110
Figure 3.3	Growth of <i>M. smegmatis</i> Δ aftA pMV306-aftA with and without acetamide.....	111
Figure 3.4	SDS-PAGE and Western blot analysis of AftA C-terminus from <i>M. tuberculosis</i> overexpressed in <i>E. coli</i> BL21 (DE3) and C43 (DE3) cells.....	113
Figure 3.5	SDS-PAGE and Western blot analysis of AftA C-terminus from <i>M. tuberculosis</i> overexpressed in <i>E. coli</i> Rosetta (DE3) pLysS and Tuner (DE3) cells.....	114
Figure 3.6	SDS-PAGE analysis of <i>M. tuberculosis</i> AftA ^{CT} purification.....	115
Figure 3.7	Fluorescence of SYPRO Orange acquired by differential scanning fluorimetry in the presence of AftA ^{CT} domain as a function of temperature.....	117
Figure 3.8	The effects of different buffers and pH expressed as Δ Tm for AftA ^{CT} domain.....	118
Figure 3.9	SDS-PAGE and Western blot analysis of full-length AftA from <i>C. glutamicum</i> overexpressed in <i>C. glutamicum</i> ATCC 13032 cells.....	120
Figure 3.10	SDS-PAGE and Western blot analysis of <i>C. glutamicum</i> Δ aftA complemented with either pMSX-aftA or empty pMSX plasmid.....	122
Figure 3.11	Glycosyl compositional analysis of cell walls of <i>C. glutamicum</i> , <i>C. glutamicum</i> Δ aftA, and <i>C. glutamicum</i> Δ aftA::pMSX-aftA.....	123

Figure 3.12	SDS-PAGE and Western blot analysis of SMALP encapsulated AftA purification.....	124
Figure 3.13	Biophysical characterisation of SMALP encapsulated AftA from <i>C. glutamicum</i>	126
Figure 3.14	Binding kinetics of SMALP encapsulated AftA from <i>C. glutamicum</i> to different substrates.....	127
Figure 3.15	Generation of conditional <i>M. smegmatis</i> Δ <i>aftB</i> ::pMV306- <i>aftB</i> mutant....	129
Figure 3.16	Growth of <i>M. smegmatis</i> Δ <i>aftB</i> ::pMV306- <i>aftB</i> in liquid minimal medium with and without acetamide.....	130
Figure 3.17	Analysis of cell wall lipids from <i>M. smegmatis</i> , merodiploid and <i>M. smegmatis</i> Δ <i>aftB</i> ::pMV306- <i>aftB</i> strains grown with and without acetamide.....	132
Figure 3.18	SDS-PAGE analysis of lipoglycans extracted from <i>M. smegmatis</i> , merodiploid and <i>M. smegmatis</i> Δ <i>aftB</i> ::pMV306- <i>aftB</i> strains grown with and without acetamide.....	134
Figure 3.19	SDS-PAGE immunoblot of mycobacteria probed with anti-arabinan monoclonal antibody F30-5.....	135
Figure 3.20	SDS-PAGE and Western blot analysis of AftB C-terminus from <i>M. tuberculosis</i> overexpressed in <i>E. coli</i> BL21 (DE3) and C43 (DE3) cells.....	137
Figure 3.21	SDS-PAGE and Western blot analysis of AftB C-terminus from <i>M. tuberculosis</i> overexpressed in <i>E. coli</i> Rosetta (DE3) pLysS and Tuner (DE3) cells.....	138
Figure 4.1	Construction strategy of <i>C. glutamicum</i> Δ <i>aftB</i> Δ <i>aftD</i> and <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i> double mutants.....	149
Figure 4.2	Analysis of cell wall lipids from <i>C. glutamicum</i> and <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i>	157
Figure 4.3	Glycosyl compositional analysis of cell walls of <i>C. glutamicum</i> and <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i>	160
Figure 4.4	Two-dimensional NMR spectra of base solubilised AG from <i>C. glutamicum</i> and <i>C. glutamicum</i> Δ <i>aftA</i> Δ <i>emb</i>	161

Figure 4.5	SDS-PAGE analysis of lipoglycans extracted from <i>C. glutamicum</i> and <i>C. glutamicum</i> Δ aftA Δ emb.....	162
Figure 4.6	Chemical structures of neoglycolipid acceptors used in this study.....	163
Figure 4.7	Arabinofuranosyltransferase activity assay utilising MJ-13-77 and MJ-14-01 neoglycolipid acceptors and membranes prepared from <i>C. glutamicum</i> , <i>C. glutamicum</i> Δ aftA, <i>C. glutamicum</i> Δ emb, <i>C. glutamicum</i> Δ aftB Δ aftD and <i>C. glutamicum</i> Δ aftA Δ emb.....	165
Figure 4.8	MALDI-TOF MS analysis of acceptors and enzymatic products.....	168
Figure 4.9	MALDI-TOF MS/MS analysis of per- <i>O</i> -methylated MJ-13-77 and MJ-14-01 acceptors.....	169
Figure 4.10	MALDI-TOF MS/MS analysis of per- <i>O</i> -methylated product A and product.....	170
Figure 4.11	MALDI-TOF MS/MS analysis of per- <i>O</i> -methylated product C.....	171
Appendix. Figure 1	BACTH analysis of self-interactions of WecA, WbbL, GlfT1, and GlfT2 from <i>C. glutamicum</i>	203
Appendix. Figure 2	BACTH analysis of self-interactions of AftA, AftB and AftC from <i>C. glutamicum</i>	204
Appendix. Figure 3	BACTH analysis of self-interactions of DprE1 and DprE2 from <i>C. glutamicum</i>	205
Appendix. Figure 4	BACTH analysis of interactions between WecA-AftB and WecA-AftC from <i>C. glutamicum</i>	206
Appendix. Figure 5	BACTH analysis of interactions between WecA-UbiA and WecA-Emb from <i>C. glutamicum</i>	207
Appendix. Figure 6	BACTH analysis of interactions between WbbL-UbiA and GlfT1-AftB from <i>C. glutamicum</i>	208
Appendix. Figure 7	BACTH analysis of interactions between GlfT1-DprE2 and AftA-AftB from <i>C. glutamicum</i>	209

Appendix. Figure 8 BACTH analysis of interactions between AftA-AftC and AftA-UbiA from <i>C. glutamicum</i>	210
Appendix. Figure 9 BACTH analysis of interactions between AftA-DprE2 and AftA-Emb from <i>C. glutamicum</i>	211
Appendix. Figure 10 BACTH analysis of interactions between AftB-AftC and AftB-UbiA from <i>C. glutamicum</i>	212
Appendix. Figure 11 BACTH analysis of interactions between AftB-DprE2 and AftB-Emb from <i>C. glutamicum</i>	213
Appendix. Figure 12 BACTH analysis of interactions between AftC-UbiA and AftC-DprE2 from <i>C. glutamicum</i>	214
Appendix. Figure 13 BACTH analysis of interactions between DprE1-DprE2 from <i>C. glutamicum</i>	215

List of Tables

Table 1.1	TB drugs and their targets.....	16
Table 2.1	Oligonucleotides used in this study.....	64
Table 2.2	Bacterial strains and plasmids used in this study.....	67
Table 2.3	List of proteins that potentially form multi-protein complexes with GlfT2 from <i>M. smegmatis</i> identified in this study.....	74
Table 2.4	List of proteins that potentially form multi-protein complexes with AftA from <i>C. glutamicum</i> identified in this study.....	78
Table 2.5	Predicted topology and function of <i>C. glutamicum</i> proteins described in this study.....	81
Table 2.6	Protein-protein interactions between <i>C. glutamicum</i> AG biosynthetic proteins determined by BACTH.....	84
Table 3.1	Primers used in this thesis.....	100
Table 6.1	Antibiotics and supplements used in this thesis.....	190
Table 6.2	PCR mixtures used in this thesis.....	191
Table 6.3	PCR cycling conditions.....	191
Table 6.4	General reaction mixture of DNA digestion used in this thesis.....	192
Table 6.5	General reaction mixture of DNA ligation used in this thesis.....	193

List of Abbreviations

ACP	acyl carrier protein
ADP	adenosine diphosphate
AMK	amikacin
AMP	adenosine monophosphate
AG	arabinogalactan
AIDS	acquired immuno-deficiency syndrome
<i>Araf</i>	arabinofuranose
<i>ArafT</i>	arabinofuranosyltransferase
AraLAM	un-capped LAM
ATP	adenosine triphosphate
BACTH	bacterial two hybrid
BCA	bicinchoninic protein assay
BCG	bacillus Calmette-Guérin
BHI	brain heart infusion
BHIS	brain heart infusion sorbitol
BN	blue native
BSA	bovine serum albumin
BTZ	benzothiazinone
C	cytosine
CAP	catabolite activator protein
CESTET	conditional expression-specialized transduction essentiality test
CDP-DAG	cytidine diphosphate-diacylglycerol

Ci	curie
CMAMES	cell wall bound corynomycolic acid methyl esters
CoA	coenzyme A
CPM	counts per minute
DAG	diacylglycerol
DAP	diaminopimelic acid
DAT	diacyl trehalose
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNB	dinitrobenzamide
DOTS	observed treatment-short course
DPA	decaprenyl-phosphate-D-arabinose
DPG	diphosphatidyl glycerol
DPM	decaprenyl-phospho-mannose
DPPR	decaprenylphosphoryl-5- phosphoribose
DPR	decaprenyl-phosphate-D-ribose
DTT	dithiothreitol
DXD	aspartic acid residue motif in glycosyltransferases
EDTA	ethylenediaminetetraacetic acid
EMB	ethambutol
ERDR	ethambutol resistance-determining region
ETH	ethionamide
FAME	fatty acid methyl esters
FAS	fatty acid synthase

FQN	fluoroquinolone
g	grams
G	guanine
Galf	galactofuranose
GalfT	galactofuranosyltransferase
GalN	galactosamine
GC-MS	gas chromatography-mass spectrometry
GDP-	Man _p guanosine diphospho-mannose <i>pyranose</i>
GlcNAc	<i>N</i> -acetylglucosamine
GCMC	glucose monocorynomycolate
GMM	glucose monomycolate
His ₆	polyhistidine tag
HIV	human immuno-deficiency virus
IPTG	isopropylthio-β-D-galactoside
INH	isoniazid
KAN	kanamycin
kDa	kilodalton
l	litre
LAM	lipoarabinomannan
LB	Luria-Bertani
LC	liquid chromatography
LM	lipomannan
LU	linkage unit
mAGP	mycolyl-arabinogalactan-peptidoglycan

M	molar
MAC	<i>Mycobacterium avium</i> complex
mAGP	mycolyl-arabinogalactan-peptidoglycan complex
MALDI-TOF	matrix assisted laser desorption ionisation-time of flight
MAMES	cell wall bound mycolic acid methyl esters
ManLAM	mannosyl capped LAM
Man _p	mannopyranose
MDR	multi-drug resistant
mg	milligram
MIC	minimum inhibitory concentration
ml	millilitre
mM	millimolar
MOPS	4-morpholine propane sulfonic acid
MPA	molybdophosphoric acid
MPI	mannosyl-phosphatidyl-myo-inositol
MS	mass spectrometry
MTBC	<i>Mycobacterium tuberculosis</i> complex
MurNAc	<i>N</i> -acetylmuramic acid
MurNGlyc	<i>N</i> -glycolylmuramic acid
nm	nanometres
NAD	nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
NADP	nicotinamide adenine dinucleotide phosphate
NADPH	reduced nicotinamide adenine dinucleotide phosphate

NMR	nuclear magnetic resonance
OD	optical density
ONPG	ortho-nitrophenyl- β -galactoside
P	phosphate
PAGE	polyacrylamide gel electrophoresis
PAS	<i>p</i> -aminosalicylic acid
PCR	polymerase chain reaction
PG	peptidoglycan
PI	phosphatidyl-myo-inositol
PILAM	lipoarabinomannan with phosphoinositide caps
PIM	phosphatidyl-myo-inositol mannoside
PPM	polyprenyl monophosphate
pRpp	5-phosphoribofuranose pyrophosphate
PZA	pyrazinamide
RIF	rifampicin
SDS	sodium dodecyl sulfate
SL	sulfolipid
STR	streptomycin
t	terminal
TB	tuberculosis
TBAH	<i>tetra</i> -butylammonium hydroxide
TDCM	trehalose dicorynomycolate
TDM	trehalose dimycolate
TDR	totally drug resistant

TFA	trifluoroacetic acid
TLC	thin-layer chromatography
TLM	thiolactomycin
TMCM	trehalose monocorynomycolate
TMM	trehalose monomycolate
TRC	triclosan
UDP-Galf	uridine diphospho-galactofuranose
UDP-GlcA	UDP-D-glucuronic acid
WHO	World Health Organisation
XDR	extensively drug resistant
X-gal	5-bromo-4-chloro-3-indolyl- β -D- galactopyranoside
$^{\circ}\text{C}$	degrees centigrade
v/v	volume/volume
w/v	weight per volume
mg	microgram
ml	microlitre
mM	micromolar

Published work associated with this thesis

Jankute M.*, Byng C.V., Alderwick L.J., and Besra G.S. 2014. Elucidation of a protein-protein interaction network involved in *Corynebacterium glutamicum* cell wall biosynthesis as determined by bacterial two-hybrid analysis. *Glycoconjugate Journal*. 31(6-7):475-83.

Jankute M.*, Grover S.*, Birch H.L., Besra G.S. 2014. Genetics of mycobacterial arabinogalactan and lipoarabinomannan assembly. *Microbiology Spectrum* 2(4):MGM2-0013-2013.

Jankute M.*, Grover S.*, Birch H.L. and Besra G.S. 2014. Genetics of mycobacterial arabinogalactan and lipoarabinomannan synthesis. *Molecular Genetics of Mycobacteria*, second edition, ASM Press.

Jankute M.*, Grover S.*, Rana A.K. and Besra G.S. 2012. Arabinogalactan and lipoarabinomannan biosynthesis: structure, biogenesis and their potential as drug targets. *Future Microbiology*. 7(1):129-47.

Chapter 1

General Introduction

1 General Introduction

1.1 Mycobacteria and related species

Mycobacterium tuberculosis, a causative agent of human tuberculosis (TB), remains one of the most successful bacterial pathogens. *M. tuberculosis* belongs to the genus *Mycobacterium*, the single genus within the family of *Mycobacteriaceae*, within the sub-order of *Corynebacterineae*, and the order of Actinomycetales. The Actinomycetales contains a diverse range of bacterial genera including *Corynebacterium*, *Nocardia*, *Rhodococcus* and *Streptomyces* (Brennan and Nikaido, 1995). The generic name *Mycobacterium* was established in 1896 by Lehmann and Neumann (Lehmann and Neumann, 1896) that proposed the genus *Mycobacterium* to include pathogenic *M. tuberculosis* and *Mycobacterium leprae* and placed it in the current mycobacterial taxonomy.

The genus *Mycobacterium* is a collection of non-motile, acid-fast, rod-shaped, Gram-positive bacteria with high guanine and cytosine content (58-69%; Clark-Curtiss, 1990) in their DNA. Classification to *Mycobacterium* is highly associated with the complex cell envelope and its components, specifically, the presence of characteristic β -hydroxy- α -alkyl branched long-chain fatty acids, termed mycolic acids. These acids contain 60-90 carbons and are cleaved to C22 to C26 fatty acid methyl esters by pyrolysis (Levy-Frebault and Portaels, 1992). This distinctive lipid-rich cell wall structure, thicker than in many other bacteria, acts as a permeability barrier and therefore contributes to the difficulty in treating mycobacterial diseases (Brennan and Nikaido, 1995, Nguyen and Thompson, 2006).

At present, over a 120 recognised species meet the criteria for inclusion into the genus *Mycobacterium* (Brown-Elliott *et al.*, 2002). Species within this genus are traditionally divided into slow-growing and fast-growing mycobacteria (Figure 1). This is based on time for visible colonies to appear under optimal conditions of nutrients and temperature from diluted suspensions. The fast-growing species appear within seven days, while the slow-growing species require seven days or more to yield visible colonies under equivalent conditions. Characteristically, fast-growing species are mostly harmless saprophytes, whereas the slow-growing species include pathogens of both human and veterinary significance, such as *M. tuberculosis*, *Mycobacterium leprae*, the causative agent of leprosy, and *Mycobacterium bovis* the causative agent of TB in cattle.

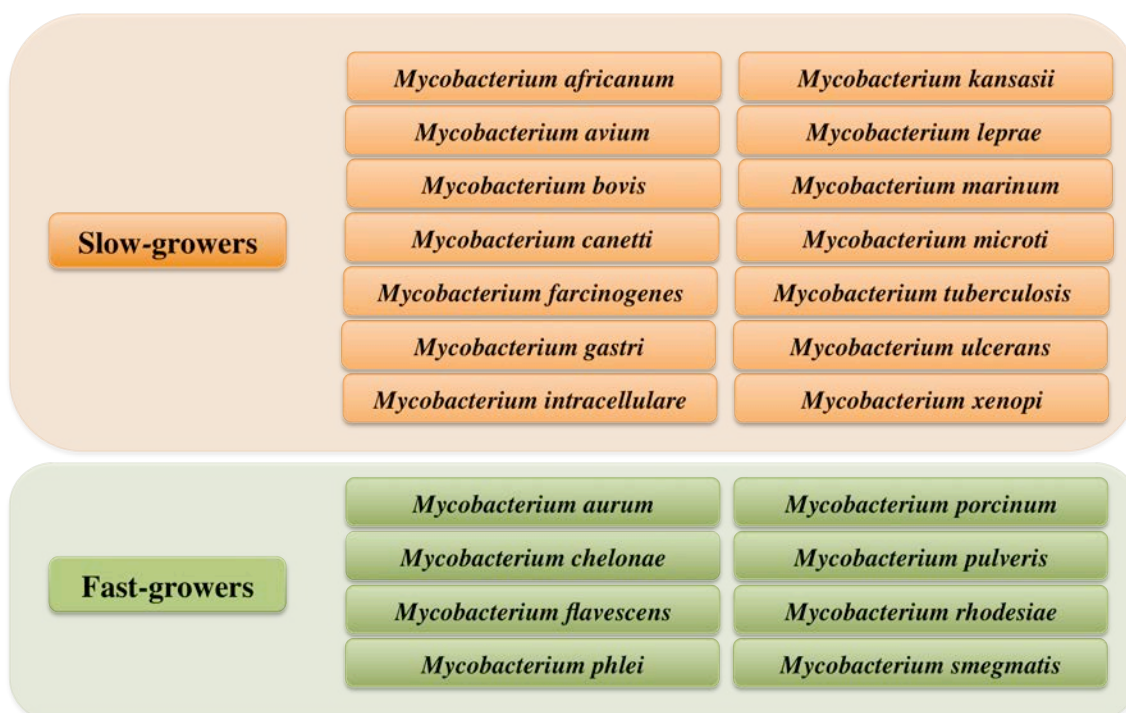


Figure 1 The genus *Mycobacterium* with examples of slow-growing and fast-growing bacilli.

The causative agents of mammalian TB are grouped into the *Mycobacterium tuberculosis* complex (MTBC) and include *M. tuberculosis*, *M. bovis*, *Mycobacterium africanum*, *Mycobacterium canetti*, *Mycobacterium microti*, *Mycobacterium pinnipedii* and *Mycobacterium caprae*. These closely related *Mycobacterium* species share more than 99.9 % genome sequence similarity (Brosch *et al.*, 2002). A notable exception is *M. canetti* that possess a smooth colony morphology and contains recombinant segments of DNA in their chromosomes (van Soolingen *et al.*, 1997). In addition, another complex termed the *Mycobacterium avium* complex (MAC) containing *Mycobacterium kansasii*, *Mycobacterium fortuitum* and *Mycobacterium avium* affect immunocompromised individuals, such as patients with human immunodeficiency virus (HIV) (Jarlier and Nikaido, 1994)

1.2 History of tuberculosis

TB is an ancient disease with a long and intricate history. It is believed that the genus *Mycobacterium* could have originated around 150 million years ago (Hayman, 1984, Daniel, 2006). Using extensive genetic analysis, Gutierrez *et al.* (2005) have demonstrated that human tubercle isolates from East Africa predate the modern MTBC members and, therefore, may correspond to progenitor species from which the MTBC strains have emerged. This ancestral species have been estimated to be nearly 3 million years old (Gutierrez *et al.*, 2005). Although members of MTBC have a distinct host preference and different phenotypic characteristics, they show high genetic homogeneity at the nucleotide level, as well as a clonal population structure (Smith *et al.*, 2009). Thus, suggesting a recent evolutionary bottleneck that might have occurred around 15,000-

35,000 years ago, resulting in a single common ancestor that later diversified into modern MTBC strains (Gutierrez *et al.*, 2005, Smith *et al.*, 2009).

Archaeological evidence documents cases of TB in various civilisations across the world. TB was present in Eastern Mediterranean around 9,250-8,160 years ago as suggested by ancient DNA analysis and lipid profiles of skeletal remains (Hershkovitz *et al.*, 2008). Genetic and molecular evidence of several ancient mummies confirms the presence and prevalence of TB in ancient Egypt around 5,000 years ago (Nerlich *et al.*, 1997, Crubezy *et al.*, 1998, Zink *et al.*, 2003). In addition, skeletal abnormalities typical to TB were documented in written texts of ancient China dating back as far as 2,300 years ago (Daniel, 2006). Finally, genetic studies of mummified bodies from Central America revealed the presence of *M. tuberculosis* infections before European colonisation (Salo *et al.*, 1994, Sotomayor *et al.*, 2004).

In 460 BC, ancient Greek physician Hippocrates described TB, known as Phthisis, as a wide spread disease accompanied by coughs, fever and consumption (Pease, 1940). Although the common believe was that TB was a hereditary disease, Aristotle (384-322 BC) defined scrofula (swelling of the lymph nodes) as an indication for TB and claimed the disease to be infectious (Garisson, 1913). Thereafter, Claudius Galenus (131-201 AD), a prominent physician in the Roman empire, described the classic symptoms of phthisis, including chronic coughs, blood-tinged sputum, fever, loss of appetite and weakness (Herzog, 1998).

TB persisted throughout centuries until the early 17th century, when the TB epidemic arose in Europe causing severe mortalities for over two centuries. This epidemic was later

termed as the Great White Plague to describe the enormity of challenges caused by TB. In 1679, Franciscus Sylvius de la Boe, a Dutch physician and scientist, published *Opera Medica*, where he first used the term “tubercle” to characterise nodules found in the lungs of patients with TB. He also indicated that “tubercles” could be observed in other organs that developed into cavities and ulcers (Keers, 1978). It was not until the 1720, when an English physician Benjamin Marten asserted TB to be highly contagious. In his manuscript he stated that consumption might be caused by miniature living creatures - “animalcules” - that lived in the lungs and caused lesions associated with TB (Doetsch, 1978).

In 1882, German bacteriologist Robert Koch gave a lecture entitled *Die Aetiologie der Tuberculose*, where he presented his discovery of *M. tuberculosis* bacilli as the etiological agent of TB (Koch, 1882). Koch demonstrated that bacilli extracted from the diseased animals or patients and subsequently injected into the guinea pigs resulted in the same disease. Employing his new staining technique, he showed that cultures from previously infected hosts and guinea pigs both contained mycobacteria. For his input into *M. tuberculosis* research, he was awarded a Nobel Prize.

French bacteriologists, Albert Calmette and Camille Guerin started their work on TB vaccine development in the early 1910's. Applying Edward Jenner's discovery of smallpox vaccine as a paradigm, they attenuated a strain of *M. bovis* to a stage where the strain was incapable of causing TB in the animals (Calmette and Guerin, 1924, Calmette, 1928). *M. bovis* BCG vaccine remains the only licensed vaccine against TB, regardless of its highly variable efficacy in many clinical trials (Doherty and Andersen, 2005). The identification and discovery of streptomycin (STR) in 1943 led to successful treatments

of life-threatening diseases including TB. Further discovery of drugs against TB followed and included *p*-aminosalicylic acid (PAS), isoniazid (INH), rifampicin (RIF), and ethambutol (EMB) (Zhang, 2005). Finally, Cole *et al.* (1998) deciphered a complete genome sequence of *M. tuberculosis* H37Rv in 1998 marking a significant milestone in combating TB.

1.3 Global epidemiology of tuberculosis

TB remains a major causes of mortality and morbidity from a single infectious organism. In 2012, an estimated 8.6 million people across the world developed TB and 1.3 million died from the disease (WHO, 2012). Globally, the highest numbers of TB cases are reported in low and middle income regions: South-East Asia (29 %), Africa (27 %) and Western Pacific (19 %) (Figure 2). Importantly, India and China alone account for 38 % of all new TB cases (WHO, 2012). Moreover, TB remains a global health problem with individuals living with HIV infection. It is estimated that at least 30 % of HIV patients worldwide have latent TB and, as a result, are 30 times more likely to develop active TB in their lifetime (WHO, 2012). In 2012, approximately 1.1 million people with HIV infection contracted TB and 0.3 million died from co-infection (WHO, 2012). The 75% of HIV/TB co-infection cases were registered in the regions of Africa (Figure 3). Although the rate of new TB cases is slowly declining, the number of TB associated deaths remains extremely high considering that TB is a treatable and preventable disease.

A number of factors have contributed to the current TB incidence worldwide. The HIV/AIDS pandemic, high population density, the variable efficacy of the BCG vaccine, poor patient compliance with current drug regimes, and malnutrition in developing

countries, which have led to the subsequent resurgence of the epidemic. These factors have also contributed to the emergence of drug resistant TB, which, according to World Health Organisation (WHO) is a major current health concern.

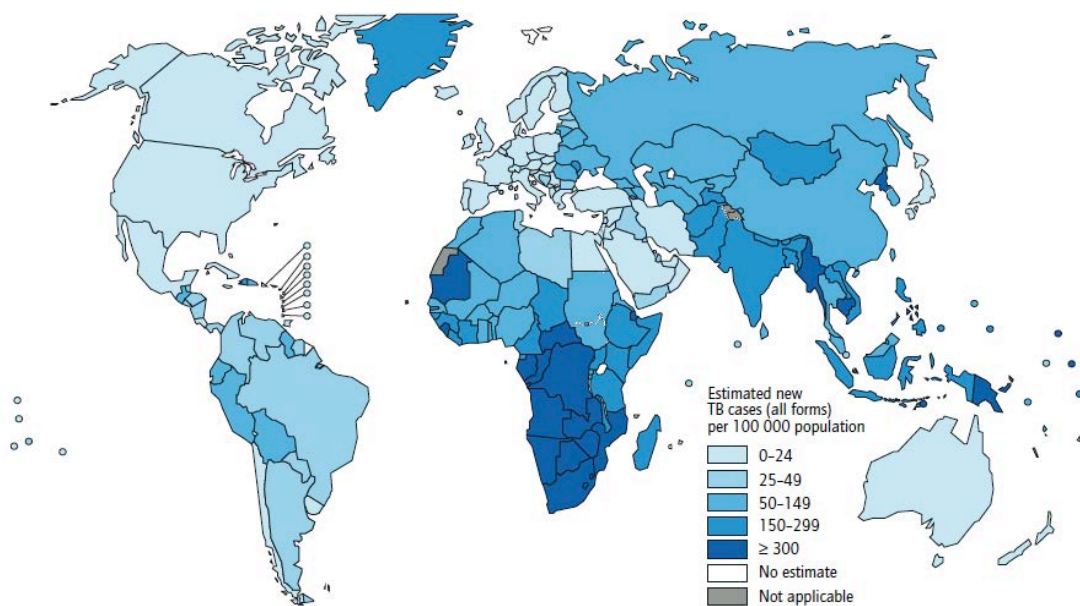


Figure 2 Estimated numbers of TB cases worldwide in 2012 (WHO, 2012).

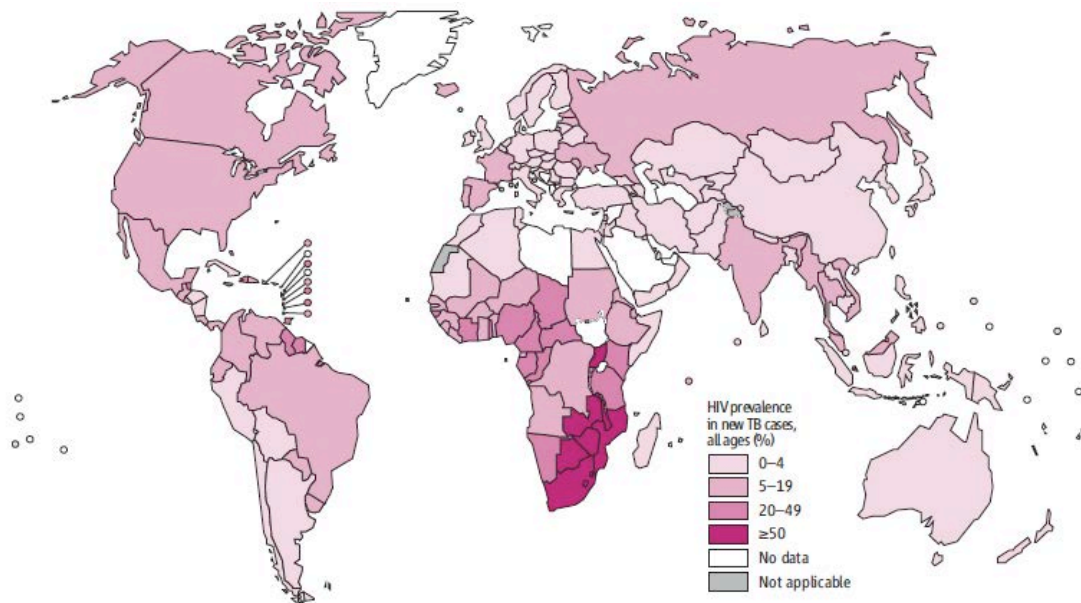


Figure 3 Estimated HIV prevalence in new TB cases worldwide in 2012 (WHO, 2012).

In 2012 alone, an estimated 450,000 cases of multi-drug resistant (MDR)-TB emerged worldwide and was documented in nearly all countries surveyed (WHO, 2012). Over half of these cases were registered in regions of Asia: China, India, and Russia. MDR-TB is defined as TB caused by *M. tuberculosis* strains that acquired resistance to at least INH and RIF, both of which are standard first-line drugs (Yew and Leung, 2006). Notably, 92 countries to date have confirmed at least one case of extensively drug resistant (XDR)-TB (WHO, 2012). XDR-TB is a result of MDR-TB strains acquiring resistance to second-line drugs, including amikacin (AMK), kanamycin (KAN), capreomycin (CAP), and fluoroquinolone (FQN) (WHO, 2012). Finally, a newly documented form of TB, totally drug resistant (TDR)-TB, has been recorded in Italy (Migliori *et al.*, 2007), Iran (Velayati *et al.*, 2009), India (Udwadia *et al.*, 2012, Udwadia and Vendoti, 2013), and South Africa (Klopper *et al.*, 2013, Velayati *et al.*, 2013). These strains are reported to be resistant to

all available first- and second-line drugs. Emergence of these drug resistance clinical isolates has prompted the need for new drugs and new drug targets.

1.4 Clinical manifestations

M. tuberculosis infection most frequently affects the lungs and is termed pulmonary TB. The symptoms of pulmonary disease include fever, severe chronic cough, chest pain, night sweats, weight loss, and general weakness. However, infection may also develop in other parts of the body affecting bones, joints, lymph nodes, pleural cavity, nervous system, liver, or spleen causing extra-pulmonary TB (Sharma and Mohan, 2013). The diagnosis of this type of TB can be more difficult, since patients are lacking a severe cough that is so commonly associated with TB infections. Another type of TB, miliary-TB, is spread throughout the body as a direct result of *M. tuberculosis* bacilli travelling from lungs and infecting multiple organs, such as the liver, kidneys or spleen (Sharma *et al.*, 2005). The hallmark of miliary-TB is a unique pattern observed on radiographs of numerous tiny spots disseminated throughout organs. Approximately 20 % extra-pulmonary TB cases are identified as miliary-TB (Ray *et al.*, 2013). Patients that have this type of TB experience a diverse range of symptoms varying from fever and enlarged lymph nodes to multiple organ failure. Untreated miliary-TB proves almost always lethal to sufferers.

1.5 Diagnosing tuberculosis

Accurate and quick diagnosis is crucial in combating TB. Various methods have been developed to date and are employed in conjunction to detect TB (Pai, 2013). The tuberculin skin test is the most widely used method to test for *M. tuberculosis* infection.

A small dose of fluid termed tuberculin (purified protein derivative) is injected intradermally and read 48 to 72 hours later. An individual with previous TB exposure would develop an immune reaction to bacterial proteins and, as a result, form an induration. However, the skin test does not discriminate between active TB and latent TB infection. In addition, the skin test is broadly accepted to lack specificity, especially in individuals residing in areas of high TB prevalence or individuals that have received the BCG vaccine. Generally, a chest radiograph is also used in conjunction with the skin test to assess individuals with suspected active TB. A radiograph would detect any chest abnormalities, such as lesions, caused by TB. However, since it has relatively poor sensitivity, similar atypical features may originate from different diseases, and follow up tests are usually performed before the final diagnosis is confirmed. The WHO has strongly promoted the sputum smear test that utilises Ziehl-Neelsen staining technique as the preferred method for diagnosing *M. tuberculosis* infection in developing countries. Even though it is an easy, quick and cost-effective method to detect acid-fast bacilli, it does not necessarily confirm the diagnosis of TB. Therefore, cell culturing is performed subsequently to confirm the diagnosis. Over many years, *M. tuberculosis* culturing from primary samples has become the gold standard for confirmation of active TB. The limitation is, however, the long growth time of *M. tuberculosis* leading to a delayed diagnosis.

Relatively recently developed methods, nucleic acid amplification test and blood assay, have also been employed in detecting active TB (Dorman, 2010). Nucleic acid amplification test, such as polymerase chain reaction, can amplify even the tiniest amounts of genetic material and differentiate between the MTBC strains. In addition, it is

also capable in detecting most common drug resistant *M. tuberculosis* strains, including those resistant to RIF and INH. A high sensitivity interferon gamma release assay using a single blood specimen may contribute additional evidence to TB diagnosis. Although the test is estimated to be more accurate than the skin test in diagnosing active TB, it is also expensive and, thus, inaccessible to general public from low and middle income countries.

1.6 Pathogenesis of tuberculosis

M. tuberculosis is an airborne disease transmitted from an active TB sufferer through aerosol droplets generated by coughing and sneezing. Once bacilli are inhaled into the respiratory tract, the majority become trapped in the upper airways together with the larger particles. However, smaller droplets containing *M. tuberculosis* bacilli can reach the lung alveoli, where they are ingested by alveolar macrophages and dendritic cells (Russell, 2007). Three different outcomes have been identified at this stage of TB development: host immune system clears the infection; infection quickly progresses into an active TB (10 %); or infection is contained within the host with no symptoms present (90 %) (Kaufmann, 2001). Years of evolution allowed mycobacteria to acquire effective mechanisms that deceive initial host microbicidal defences, thus preventing the elimination of *M. tuberculosis*. The immune system unable to eliminate pathogenic bacilli protects the organism by isolating them. In case of containment, infected macrophages attract monocytes to the site of infection, where the lesion starts to develop (Russell, 2007). Engulfed mycobacteria shed a variety of antigens, including lipids, glycolipids and lipoproteins that are packaged into vesicles and delivered to bystander cells. Dendritic cells possess a high number of cell surface molecules, notably CD1, which can bind to

the lipid antigens and present them to T-cells (Russell, 2001). At the lymph nodes, dendritic cells present these mycobacterial antigens to different types of T cells resulting in a release of inflammatory cytokines and development of additional lesions at the site of infection (Russell, 2001). This robust pro-inflammatory response leads to formation of the granuloma, an aggregation of infected macrophages surrounded by layers of lymphocytes and fibroblasts. Since access to oxygen and nutrients are limited inside the granuloma, the bacilli go into the state of dormancy (Russell, 2007). Such structures can persist for years in immunocompetent individuals, thus preventing the growth and dispersion of mycobacteria within the host. Rupture of the granuloma is frequently associated with the weakening of the immune system, notably HIV infection. *M. tuberculosis* and HIV co-infection hasten the progression of one another, accelerating the deterioration of host immunological functions (Kaufmann and McMichael, 2005).

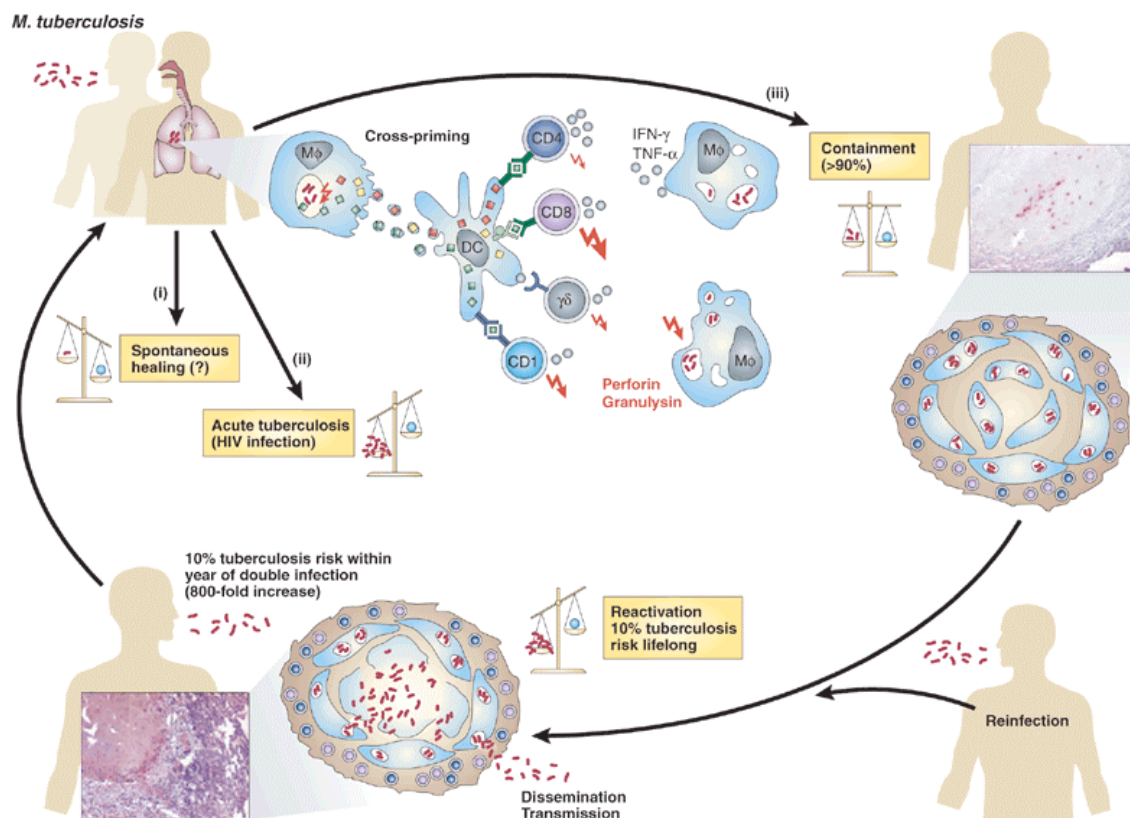
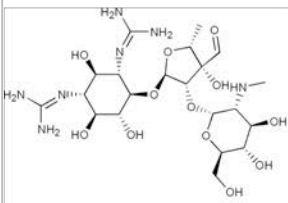
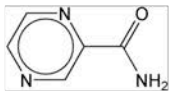
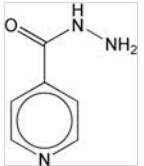
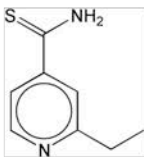
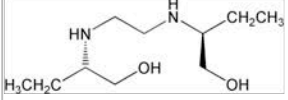
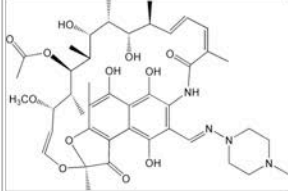


Figure 4 Different outcomes of *M. tuberculosis* infection. There are three potential outcomes of *M. tuberculosis* infection in the human host. (i) Spontaneous healing, although it is estimated to be extremely rare. (ii) Disease progression to acute active TB, frequently observed in immunodeficient individuals. (iii) In majority of cases, containment of mycobacteria within the granuloma and, therefore, prevention of active TB. Briefly, *M. tuberculosis* bacilli are engulfed by alveolar macrophages and dendritic cells. The latter, present mycobacterial antigens to different T-cell populations including CD4⁺, CD8⁺, γδ, and CD1⁺. Subsequently, T-cells enhance the antibacterial activity of macrophages by releasing cytokines such as interferon-γ that typically result in either arrest or clearance of the infection. In the case of containment, disease usually progresses later due to reactivation, when rupture of the granuloma leads to dissemination of *M. tuberculosis* bacilli throughout the host resulting in active TB. Reprinted by permission from Macmillan Publishers Ltd: *Nature Medicine*, Kaufmann & McMichael (2005).

1.7 Tuberculosis drug treatments

The present most effective TB drug regimen recommended by WHO is Directly Observed Treatment-Short Course (DOTS) introduced in 1991. This TB strategy, lasting at least six months, reaches approximately 95 % success rate and is divided into two phases: initial phase and continuation phase (WHO, 2012). Initial phase aims to kill actively growing and semi-dormant *M. tuberculosis* bacilli through administration of INH, RIF, pyrazinamide (PZA), and EMB for two months. Subsequent administration of INH and RIF for the following four months eliminates the remaining dormant bacilli in the continuation phase. If the isolate is resistant to one or more drugs, the treatment regimen is changed and extended. Emergence of MDR-TB prompted the need for a new drug strategy, specifically adjusted to treat drug resistant cases. Therefore, the DOTS-Plus regimen was introduced by WHO and is currently adopted by many countries around the world as a national drug resistant TB control policy (WHO, 2004). This treatment strategy includes administration of second-line drugs that are generally more toxic and less effective than the first-line drugs. The treatment contains two or more antibiotics to which the *M. tuberculosis* strain is susceptible in addition to the primary drug given for at least six months. The duration of DOTS-Plus regimen is eighteen to twenty-four months. There is no regimen in place for XDR-TB cases, since different *M. tuberculosis* isolates can be resistant to various first- and second-line drugs, making treatment options limited.

Table 1 TB drugs and their targets.

Drug	Drug mechanism of action	Genes involved in resistance	Function of gene
Streptomycin (STR) (1943) 	Inhibition of protein synthesis	<i>rpsL</i> <i>rrs</i>	S12 ribosomal protein 16S rRNA
Pyrazinamide (PZA) (1952) 	Acidification of cytoplasm and de-energising the membrane	<i>pncA</i>	Nicotinamidase
Isoniazid (INH) (1952) 	Inhibition of mycolic acid biosynthesis and other effects on DNA, lipids, carbohydrates and NAD metabolism	<i>katG</i> <i>inhA</i> <i>ndh</i> <i>oxyR-aphC</i>	Catalase-peroxidase Enoyl-ACP reductase NADH dehydrogenase II Alkyl hydroperoxidase
Ethionamide (ETH) (1956) 	Inhibition of mycolic acid biosynthesis	<i>ethA</i> <i>inhA</i>	Flavin mono-oxygenase Enoyl-ACP reductase
Ethambutol (EMB) (1961) 	Inhibition of cell wall arabinogalactan biosynthesis	<i>embCAB</i>	Arabinosyltransferases
Rifampicin (RIF) (1966) 	Inhibition of transcription	<i>rpoB</i>	RNA polymerase

1.8 Mycobacterial cell wall

Despite belonging to the Gram-positive branch of classification, *Mycobacterium* species possess a cell envelope that distinguishes them from both Gram-positive and Gram-negative organisms. Its cell wall has an unusually high content of lipids that act as a permeability barrier and hence contributes to resistance to common antibiotics and chemotherapeutic agents. The cell wall core of *M. tuberculosis* is composed of peptidoglycan (PG), covalently attached to arabinogalactan (AG), which is further esterified by mycolic acids – a model first proposed by Minnikin and later supported by McNeil and Brennan (Daffe *et al.*, 1990, McNeil *et al.*, 1990, McNeil *et al.*, 1991, Besra *et al.*, 1995, Dover *et al.*, 2004, Kaur *et al.*, 2009, Minnikin 1982, Minnikin *et al.*, 2002). The outer membrane segment is comprised of solvent extractable lipids believed to either intercalate with the cell wall bound mycolates or form a more clearly defined interlayer (Liu *et al.*, 1995, Hoffmann *et al.*, 2008). Other important components of the mycobacterial cell wall include the phosphatidyl-*myo*-inositol mannosides (PIMs), and their more glycosylated products lipomannan (LM) and lipoarabinomannan (LAM) (Mishra *et al.*, 2011a). These glycoconjugates display a broad range of immunomodulatory activities associated with TB pathogenesis (Mishra *et al.*, 2011a). The cell wall structure is crucial for the growth, viability and virulence of *M. tuberculosis* and, therefore, is often the target of effective chemotherapeutic agents against TB.

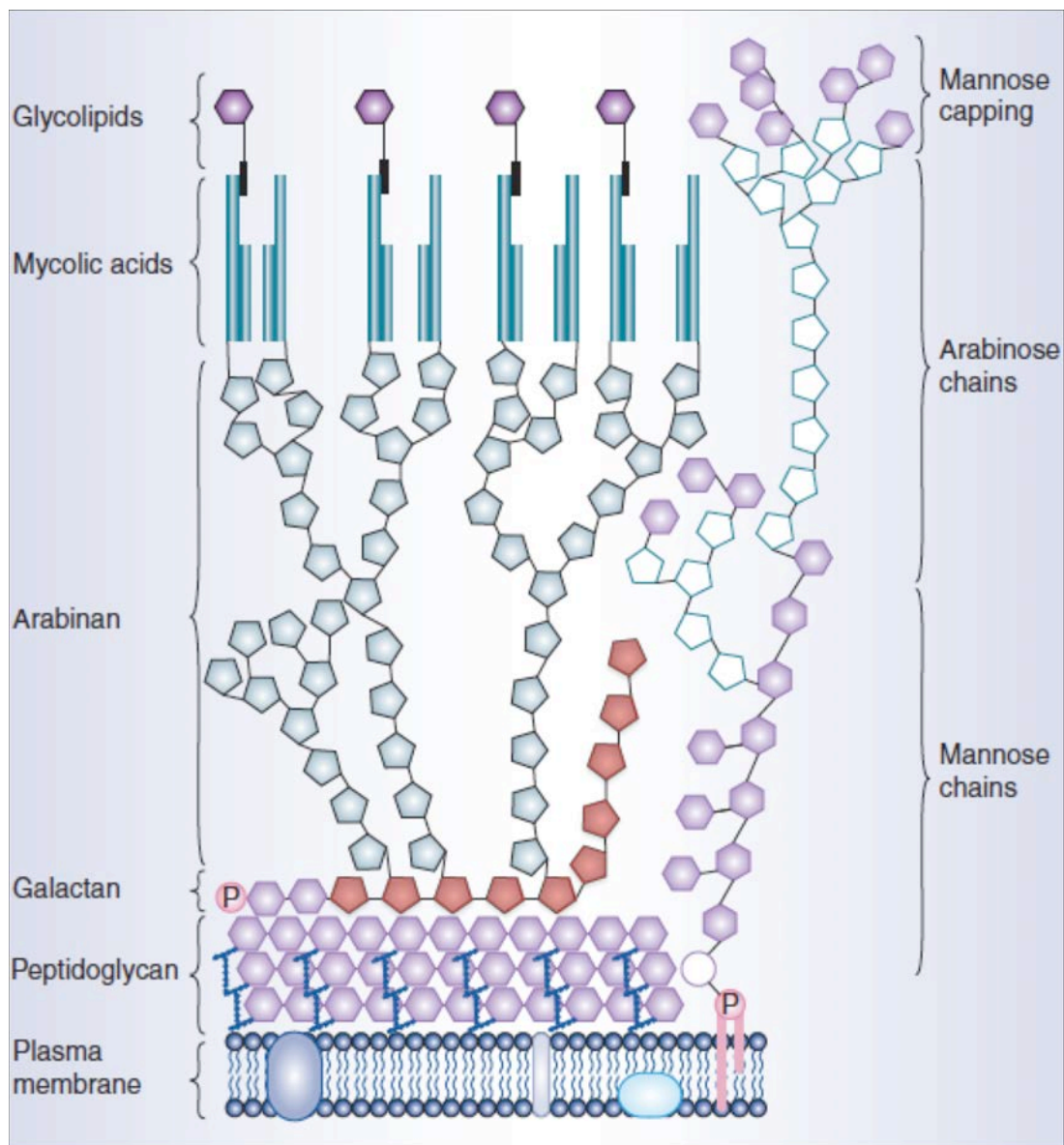


Figure 5 Cell envelope of *Mycobacterium tuberculosis*. The cell wall core consists of peptidoglycan connected to arabinogalactan and covalently attached hydrophobic mycolic acids, which are intercalated with glycolipids and a diverse repertoire of complex lipids. Lipoarabinomannan is linked to the plasma membrane *via* the lipid portion of the phosphatidylinositol anchor and serves as a modulin with immunoregulatory effects Reprinted by permission from Future Medicine Ltd: *Future Microbiology*, Jankute *et al.*, (2012).

1.8.1 Peptidoglycan

PG is a complex macromolecular structure located directly above the plasma membrane and found in nearly all eubacteria (van Heijenoort, 2001a, van Heijenoort, 2001b, van Heijenoort, 2007). The mesh-like arrangement of PG provides strength to the bacterial cell wall, thus enabling it to counteract the osmotic pressure while retaining cell integrity and shape. Although mycobacterial PG has not been intensively studied, it is expected to be comparable to that of *E. coli*, as both of them fall into the A1 γ class of PG (Schleifer and Kandler, 1972). *M. tuberculosis* PG is comprised of glycan chains composed of alternating *N*-acetylglucosamine (GlcNAc) and *N*-glycolylmuramic acid (MurNGlyc) residues, linked together *via* a β -1,4 linkage. A side chain of L-alanine-D-glutamate-D-isoglutaminsyl-*meso*-diaminopimelyl-D-alanine-D-alanine peptide substitutes the carboxylic acid group of each MurNGlyc residue (Petit *et al.*, 1969). Notably, PG of *E. coli* contains *N*-acetylmuramic acid (MurNAc) residues rather than MurNGlyc, suggesting that the *N*-acetyl group has been oxidised to *N*-glycolyl to form MurNGlyc (Azuma *et al.*, 1970). This distinguishing feature of *M. tuberculosis* PG may provide additional hydrogen bonding and, therefore, leading to a more robust mesh-like structure (Brennan and Nikaido, 1995). Moreover, the overall degree of cross-linking in mycobacteria is approximately 70-80 %, whereas in *E. coli* it is 30-50 % (Vollmer and Holtje, 2004). The majority of cross-links of the adjacent peptides are between the carboxyl group of a terminal D-alanine and the amino group of *meso*-diaminopimelic acid (DAP). The remaining cross-links are between the carboxyl group of one DAP unit and the amino group of another DAP. Finally, several MurNGlyc residues are used as an attachment site for the AG, where C-6 of MurNGlyc forms a phosphodiester bond with the

α -D-GlcNAc residue of the AG linker unit (Wietzerbin *et al.*, 1974, Lederer *et al.*, 1975, McNeil *et al.*, 1990).

PG biosynthesis begins with MurA (*Rv1315*) transferring an enoylpyruvate residue from phosphoenoylpyruvate to UDP-GlcNAc (Brown *et al.*, 1995, Xu *et al.*, 2014). Subsequently, MurB (*Rv0482*) reduces this enoylpyruvate moiety, yielding UDP-MurNAc (Pucci *et al.*, 1992, Kumar *et al.*, 2011). The second monosaccharide donor, UDP-MurNGly, is generated utilising UDP-GlcNAc in a single reaction catalysed by NamH (*Rv3818*) (Raymond *et al.*, 2005). The side chain peptide of MurNGlyc is synthesised by MurC (*Rv2152c*) (Liger *et al.*, 1995, Mahapatra *et al.*, 2000), MurD (*Rv2155c*) (Mengin-Lecreulx *et al.*, 1989), MurE (*Rv2158c*) (Maruyama *et al.*, 1988, Munshi *et al.*, 2013), and MurF (*Rv2157c*) (Maruyama *et al.*, 1988, Munshi *et al.*, 2013) ligases adding L-alanine, D-glutamic acid, DAP, D-alanine, D-alanine, respectively. The translocase MurX (*Rv2156c*), previously designated MraY, then transfers the newly formed UDP-MurNAc with the side chain to undecaprenyl phosphate, yielding Lipid I (Bouhss *et al.*, 1999, Bouhss *et al.*, 2004). Subsequently, MurG transfers GlcNAc residue from UDP-GlcNAc to Lipid I, resulting in Lipid II (Ikeda *et al.*, 1990). The yet unknown flippase translocates Lipid II to the periplasmic side for the final assembly of PG.

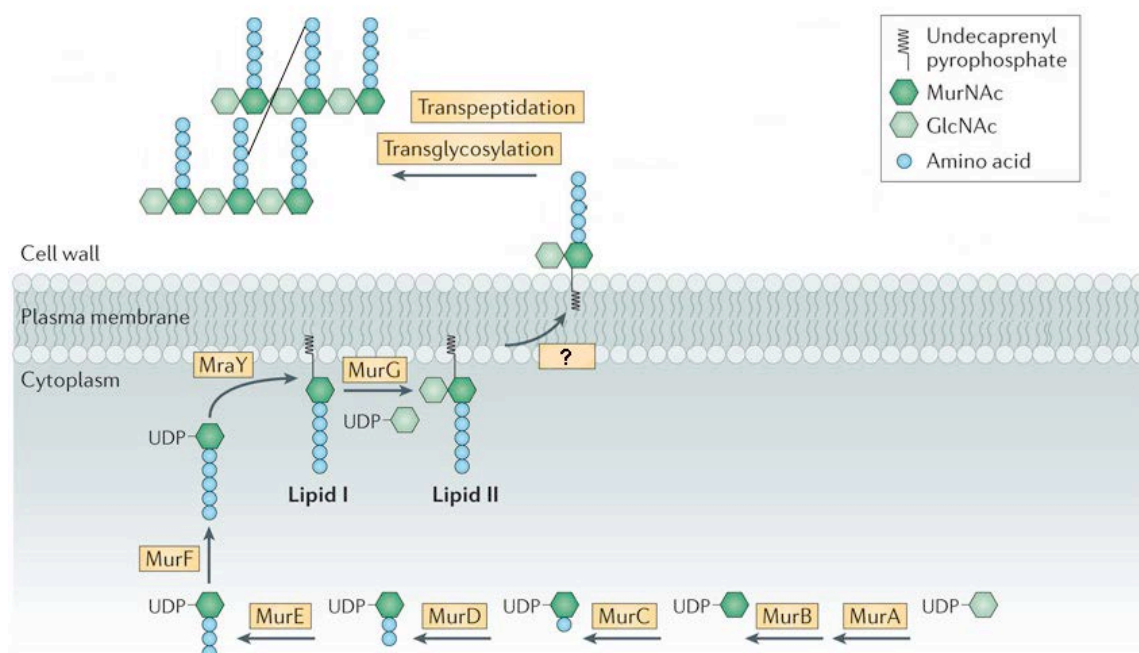


Figure 6 Biosynthesis of peptidoglycan in *M. tuberculosis*. The synthesis of UDP-GlcNAc is initiated in the cytoplasm and mediated by MurA and MurB enzymes. MurC, MurD, MurE, and MurF catalyse further addition of L-alanine, D-glutamic acid, meso-diaminopimelic acid, and D-alanine-D-alanine peptide to UDP-GlcNAc, respectively. This intermediate is transferred to undecaprenyl phosphate lipid carrier by MurX (previously designated as MraY) leading to formation of lipid I. Finally, MurG attaches the GlcNAc residue to generate lipid II and the entire molecule is translocated across the plasma membrane, where it is utilised in the biosynthesis of PG. Adapted by permission from Macmillan Publishers Ltd: *Nature Reviews Microbiology*, Pinho *et al.* (2013).

1.8.2 Arabinogalactan

1.8.2.1 Structural features of arabinogalactan

AG is a key structural component of the mycolyl-arabinogalactan-peptidoglycan (mAGP) complex that constitutes approximately 35% of the mycobacterial cell wall mass. This heteropolysaccharide plays an important role in covalently anchoring the mycolic acid layer to the inner PG layer, and is unique in that all of its sugars are present in the furanose ring form (McNeil *et al.*, 1987). Moreover, in contradiction to most bacterial

polysaccharides, AG is comprised of several distinct structural motifs instead of repeating units (Daffe *et al.*, 1990, McNeil *et al.*, 1990, Besra *et al.*, 1995). The whole mycolyl-AG structure is attached to PG *via* a specific linker unit and thus, it is believed to be the most vulnerable part of the complex. Characterisation of per-*O*-alkylated oligosaccharide alditols together with gas chromatography-mass spectrometry, fast atom bombardment-mass spectrometry and nuclear magnetic resonance (NMR) analysis elucidated the structural details of AG.

The galactan domain is composed of approximately 30 alternating $\beta(1\rightarrow5)$ and $\beta(1\rightarrow6)$ galactofuranosyl (Gal f) residues connected in a linear fashion. At the reducing end of AG, the galactan chain is linked to the C-6 position of selected *N*-glycolylmuramic acid residues of PG *via* an α -L-Rhap-(1 \rightarrow 3)- α -D-GlcNAc-1-phosphate linker unit (McNeil *et al.*, 1990). Three similar D-arabinan chains comprising roughly 30 arabinofuranosyl (Ara f) residues each are attached to the C-5 of specific $\beta(1\rightarrow6)$ linked-Gal f residues (Besra *et al.*, 1995). Since the AG structure is essential to *M. tuberculosis*, many gene deletion studies investigating AG have been performed in the *Corynebacterium* genus, where aspects of AG biosynthesis are non-essential. Knockout mutants in *C. glutamicum* together with mass spectrometry determined that the arabinan chains of AG are attached to the 8th, 10th, and 12th residue of the linear galactan chain (Alderwick *et al.*, 2005) (Figure 7). Previous work demonstrated that the arabinan domain is present as a highly branched network synthesised on a backbone of $\alpha(1\rightarrow5)$ linked sugars with branching introduced by the presence of 3,5- α -D-Ara f residues (Daffe *et al.*, 1990). The non-reducing termini of arabinan structure consists of a distinct hexa-arabinofuranosyl motif [β -D-Ara f -(1 \rightarrow 2)- α -D-Ara f]₂-3,5- α -D-Ara f -(1 \rightarrow 5)- α -D-Ara f] (McNeil *et al.*, 1994).

Analysis of per-*O*-methylated mAGP and per-*O*-alkylated oligoglycosyl alditols determined that position 5 of both terminal β -D-Araf and the penultimate 2- α -D-Araf are the attachment sites for the mycolic acids (McNeil *et al.*, 1991). Follow-up studies determined that mycolyl residues are located in clusters of four on the terminal hexaarabinofuranoside motifs, with only two-thirds of these being mycolated.

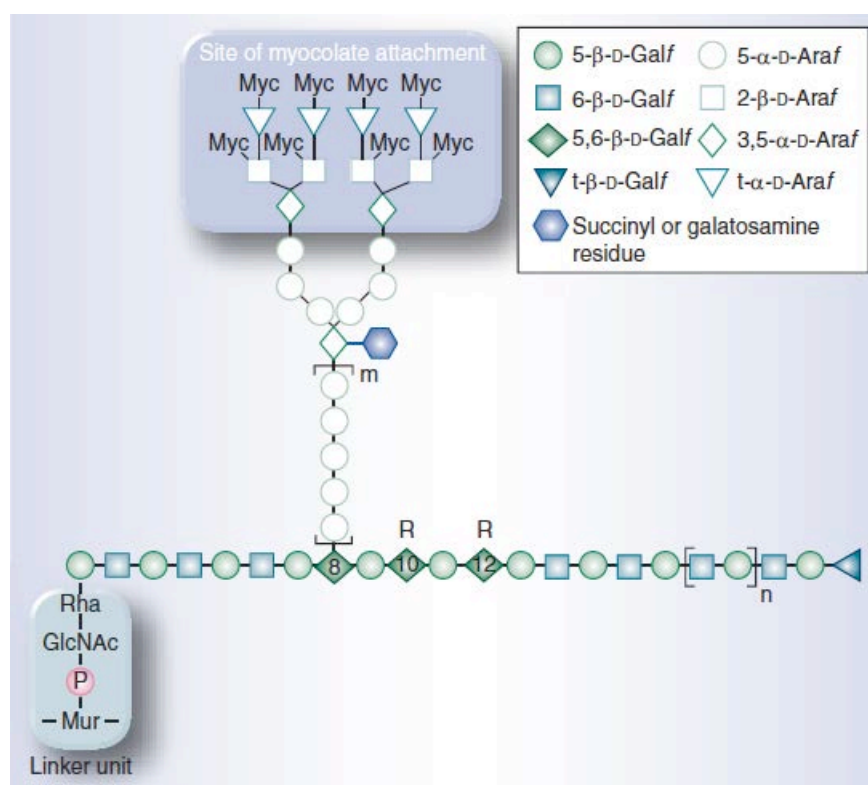


Figure 7 The general structure of arabinogalactan present in *M. tuberculosis*. The galactan domain containing approximately 30 Galf residues is attached to the peptidoglycan layer *via* a linker unit. This linear galactan serves as a base for the three arabinan chains that are connected to the galactan at positions 8, 10 and 12. The branched arabinan domain is covalently attached to the mycolates through a unique hexaarabinofuranosyl motif. Succinyl or galactosamine residue is attached to the inner 3,5- α -D-Araf residues, thus completing the structure of AG. Reprinted by permission from Future Medicine Ltd: Future Microbiology, Jankute *et al.*, (2012).

An endogenous arabinase, which can cleave the arabinan, has been partially purified from *M. smegmatis* (Dong *et al.*, 2006). The use of this enzyme together with matrix-assisted laser desorption/ionization-time of flight-mass spectrometry and NMR allowed the sequencing of very large fragments of arabinan chains released from the mycobacterial cell wall. Importantly, galactosamine (GalN) residues, previously detected as a minor covalently bound sugar residue of the cell envelope of slow growing mycobacteria, such as *M. tuberculosis* and *M. avium* (Draper *et al.*, 1997), were shown to be located on the C-2 position of some of the internal 3,5- α -D branched Araf residues (Lee *et al.*, 2006) and the stereochemistry of the GalN moiety was confirmed to be an α -anomer (Peng *et al.*, 2012) (Figure 7). In addition, succinyl groups were found located on the interior branched arabinosyl residues (Bhamidi *et al.*, 2008). Approximately one of the three arabinan chains linked to the linear galactan contain a GalN group and one of three are succinylated (Bhamidi *et al.*, 2008). In addition, succinyl residues were shown to be present only on the non-mycolated chains, thus suggesting a possible negative control mechanism for mycolylation. It is speculated that the GalN residue of AG may serve as a specific function during host infection (Bhamidi *et al.*, 2008).

1.8.2.2 Precursor formation

The biosynthesis of the linkage unit employs two high-energy substrates, UDP-GlcNAc and dTDP-Rha (**Figure 8**). UDP-GlcNAc, a sugar donor for both the AG linker unit and the biosynthesis of PG, is formed *via* a four-step reaction. Three enzymes, glutamine fructose-6-phosphate transferase GlmS, phosphoglucosamine mutase GlmM and glucosamine-1-phosphate acetyl transferase/*N*-acetylglucosamine-1-phosphate urididyl

transferase GlmU catalyse the conversion of fructose-6-phosphate to UDP-GlcNAc in *E. coli* (Mengin-Lecreulx and van Heijenoort, 1993, Mengin-Lecreulx and van Heijenoort, 1994, Mengin-Lecreulx and van Heijenoort, 1996, Klein and Ferre-D'Amare, 2006). Analysis of the genome sequence of *M. tuberculosis* determined that *Rv3436c*, *Rv3441c* and *Rv1018c* are homologous to the *E. coli* *glmS*, *glmM*, and *glmU* genes, respectively (Zhang *et al.*, 2008). GlmS is responsible for the conversion of fructose-6-phosphate to glucosamine-6-phosphate, which is then converted to glucosamine-1-phosphate by GlmM. Recent gene deletion studies demonstrated that *MSMEG_1556*, a *M. smegmatis* gene encoding the homologue of *glmM* from *E. coli*, is essential for survival (Li *et al.*, 2012). Furthermore, it was shown that *M. tuberculosis* *Rv3441c* possesses phosphoglucosamine mutase activity and was able to compensate for the loss of *MSMEG_1556* in the conditional mutant (Li *et al.*, 2012), thus demonstrating that they share the same function. Mycobacterial GlmU is a bifunctional enzyme involved in the last two sequential steps of UDP-GlcNAc synthesis (**Figure 8**). Disruption of *glmU* in *M. smegmatis* resulted in gross morphological changes and loss of viability (Zhang *et al.*, 2008). Biochemical characterisation as well as the structure of *M. tuberculosis* GlmU has recently been established (Zhang *et al.*, 2009, Zhou *et al.*, 2011, Zhou *et al.*, 2012).

The second nucleotide donor utilised in the biosynthesis of the linkage unit is dTDP-Rha (**Figure 8**). Presence of L-rhamnose, a sugar absent in humans, makes the biosynthetic machinery of the mycobacterial linkage unit an attractive drug target. As a result, the rhamnosyl biosynthetic pathway has come under close investigation and a number of inhibitors targeting this pathway have been described (Ma *et al.*, 2001, Babaoglu *et al.*, 2003, Kantardjieff *et al.*, 2004). Synthesis of dTDP-Rha occurs *via* a linear four-step

reaction. Recognition of the genes involved in this pathway was revealed by comparison to known polysaccharide biosynthetic enzymes found in other bacteria, notably *E. coli* (Ma *et al.*, 1997). RmlA (*Rv0334*) converts dTTP and α -D-glucose-1-P into dTDP-glucose. A strain of *E. coli* lacking four rhamnose biosynthetic genes was complemented with *rmlA* from *M. tuberculosis*. The α -D-Glc-P thymidylyltransferase activity was observed in the cellular extract samples, therefore confirming its proposed function (Ma *et al.*, 1997). The formed intermediate is utilised by dTDP-D-glucose-4,6-dehydratase RmlB (*Rv3464*), dTDP-4-oxo-6-deoxyglucose-3,5-epimerase RmlC (*Rv3465*), and dTDP-6-deoxy-L-lyxo-4-hexulose reductase RmlD (*Rv3266*) to finally form the nucleotide donor dTDP-Rha. Gene deletion studies in the presence of a rescue plasmid with a temperature sensitive origin of replication determined that all four *rmlA* (Qu *et al.*, 2007), *rmlB* (Li *et al.*, 2006), *rmlC* (Li *et al.*, 2006) and *rmlD* (Ma *et al.*, 2002) genes were essential to *M. smegmatis*. Hence, dTDP-Rha is an essential sugar donor for mycobacterial growth and enzymes involved in its synthesis are potential chemotherapeutic targets. Finally, enzyme assays employing RmlA-D from *M. tuberculosis* to screen inhibitors for developing novel anti-TB therapeutics have been established (Ma *et al.*, 2001, Sha *et al.*, 2012).

The high-energy nucleotide substrate UDP-Galp is formed *via* a three-step reaction (**Figure 8**). GalU (*Rv0993*) a glucose-1-phosphate uridylyltransferase catalyses the formation of UDP-Glcp from UTP and glucose-1-P (Lai *et al.*, 2008). Recently, *galU* from *M. tuberculosis* was successfully expressed, purified and biochemically characterised (Lai *et al.*, 2008). The second enzyme GalE is responsible for the epimerisation reaction, which forms UDP-Galp from UDP-Glcp in *E. coli* (Lemaire and

Muller-Hill, 1986). Studies in *M. smegmatis* examined the reverse reaction using radiolabelled UDP-Galp and observed UDP-glucose-4-epimerase activity. Sequentially, the *M. smegmatis* protein was purified and its N-terminal sequence was shown to be similar to that of the *M. tuberculosis* product GalE1 (*Rv3634c*) (Weston *et al.*, 1997). The transformation of UDP-Galp to the furanose form is catalysed by UDP-galactopyranose mutase Glf encoded by *MSMEG_6404* from *M. smegmatis* and *Rv3809c* from *M. tuberculosis* (Nassau *et al.*, 1996). Allelic exchange experiments of *MSMEG_6404* highlighted the essentiality of *glf* to mycobacteria (Pan *et al.*, 2001). It has also been structurally characterised (Sanders *et al.*, 2001, Beis *et al.*, 2005).

Arabinan biosynthesis utilises β -D-arabinofuranosyl-1-monophosphodecaprenol (DPA), the only known donor of Araf residues in mycobacteria and corynebacteria (Wolucka, 2008). Recently, its membrane linked synthesis has been investigated in detail (Alderwick *et al.*, 2005, Alderwick *et al.*, 2011b) (**Figure 8**). The initial reaction involves activation of ribose-5-phosphate by a phosphoribosyl-1-pyrophosphate synthetase PrsA (*Rv1017c*) to yield 5-phosphoribosyl-1-pyrophosphate (pRpp) (Alderwick *et al.*, 2011b). UbiA (*Rv3806c*) then transfers pRpp to a decaprenylmonophosphate producing decaprenylphosphoryl-5-phosphoribose (DPPR) (Alderwick *et al.*, 2011b). Disruption of *ubiA* (*NCgl2781*) in *C. glutamicum* resulted in a complete loss of cell wall arabinan, demonstrating that DPA is the only Araf sugar donor used in AG biosynthesis (Alderwick *et al.*, 2005). Remarkably, the mutant still generated a modified LAM version, which was arabinosylated even in the absence of DPA (Tatituri *et al.*, 2007a). An alternative source and mechanism by which these Araf residues are added to this glycolipid is yet to be resolved. DPPR is then dephosphorylated to decaprenyl-5-phosphoribose (DPR) by the

putative phospholipid phosphatase encoded by *Rv3807c*. Its homologue in *M. smegmatis* (*MSMEG_6402*) was shown to be a non-essential gene (Jiang *et al.*, 2011). The DprE1 (*Rv3790*) and DprE2 (*Rv3791*) heterodimer catalyses the epimerisation of DPR to DPA, which occurs *via* an oxidation-reduction mechanism. DPR is initially oxidised at the C2-OH group to form the keto-sugar intermediate, decaprenol-1-monophosphoryl-2-keto- β -erythro pentofuranose (DPX) that is subsequently reduced to DPA (Mikusova *et al.*, 2005). Deletion studies in *C. glutamicum* showed that *dprE1* (*NCgl0187*) is essential to bacterial growth whereas *dprE2* (*NCgl0186*) is not (Meniche *et al.*, 2008). In the absence of *dprE2*, a different enzyme encoded by *NCgl1429* was proposed to carry out the function of DprE2 since *NCgl1429* showed a similar function *in vivo* and appeared to be essential in the *NCgl0186*-inactivated mutant (Meniche *et al.*, 2008). Further investigation demonstrated that *dprE1* (*MSMEG_6382*) is also an essential gene in *M. smegmatis* (Crellin *et al.*, 2011). These results highlighted DprE1 as a novel drug target. Indeed, recent studies led to the discovery of two classes of potent compounds with specific activities against mycobacteria: dinitrobenzamide derivatives (DNB) and nitro-compounds related to DNBs – nitro-benzothiazinones (BTZ), both of which were revealed to target the decaprenylphosphoryl- β -D-ribose 2' epimerase encoded by *dprE1* (Christophe *et al.*, 2009, Makarov *et al.*, 2009). The structural complex of DprE1-BTZ has been determined revealing the mode of inhibitor binding (Christophe *et al.*, 2009, Batt *et al.*, 2012). In addition, a BTZ mechanism of action was proposed in *C. glutamicum*. Grover *et al.* (2014) determined that in presence of the drug, bacteria accumulate the intermediate DPR and, therefore, fail to recycle decaprenyl phosphate. As a result, AG and LAM biosynthesis is inhibited ultimately leading to the cell death. Notably, an increase of decaprenyl phosphate synthesis allows *C. glutamicum* to evade

the action of BTZ, demonstrating a new mechanism of drug resistance to the BTZ class of drugs (Grover *et al.*, 2014).

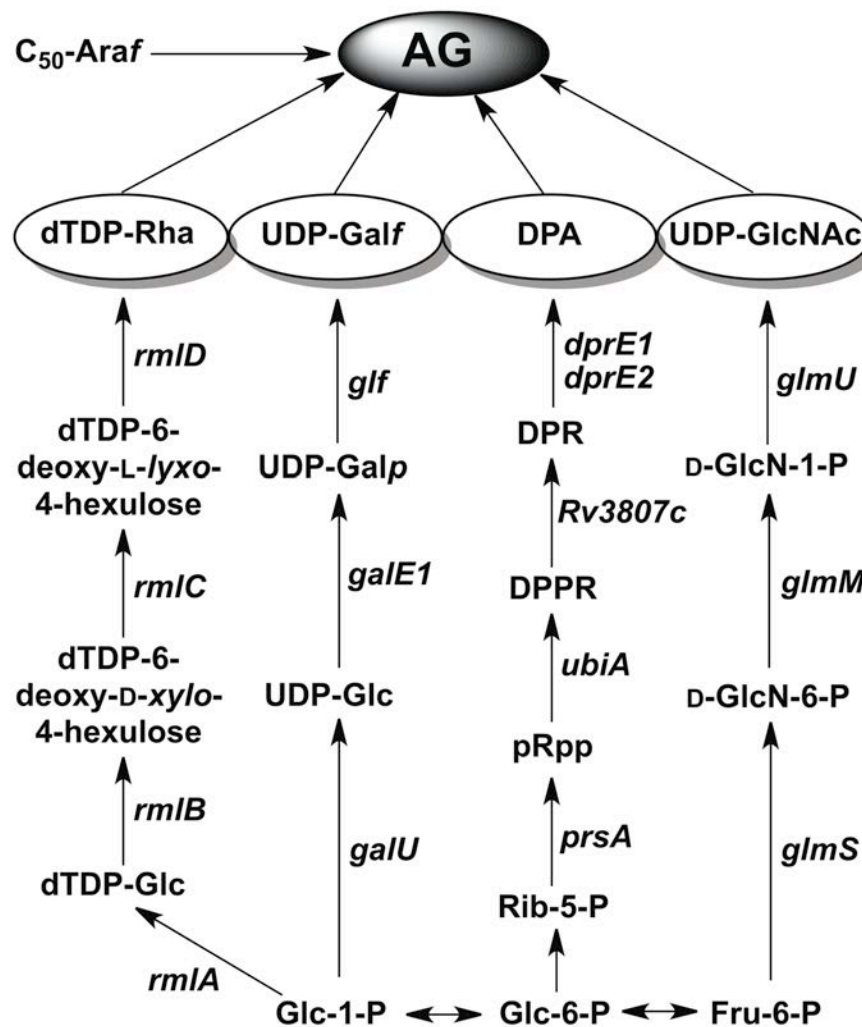


Figure 8 The biosynthesis of sugar donors required for mycobacterial arabinogalactan biosynthesis. Both UDP-GlcNAc and dTDP-Rha are utilised in the formation of the linker unit. UDP-Galf is the sugar donor of the galactofuranosyl residues used in the galactan chain formation. Decaprenylphosphoryl-D-arabinofuranose is the only known high energy nucleotide providing arabinofuranosyl residues to the arabinan domain of arabinogalactan. Reprinted by permission from ASM: *Microbiology Spectrum*, Jankute *et al.*, (2013).

1.8.2.3 Biosynthesis of arabinogalactan

The biosynthesis of AG begins with the formation of the linker unit synthesised on a decaprenyl phosphate lipid carrier. WecA (*Rv1302*) catalyses the first reaction by transferring GlcNAc-1-P from the sugar donor UDP-GlcNAc to the lipid carrier (Mikusova *et al.*, 1996, Jin *et al.*, 2010). Lipopolysaccharide analysis of a *wecA*-defective strain of *E. coli* complemented with either *M. tuberculosis* (*Rv1302*) or *M. smegmatis* (*MSMEG_4947*) homologue showed restoration of lipopolysaccharide biosynthesis, thus providing evidence it has the same function as the WecA protein from *E. coli* (Jin *et al.*, 2010). In addition, inactivation of *wecA* from *M. smegmatis* using a homologous recombination strategy resulted in drastic morphological changes and loss of viability (Jin *et al.*, 2010). Rhamnosyltransferase WbbL (*Rv3265c*) is responsible for the transfer of the rhamnose residue from dTDP-Rha substrate to the 3-position of the GlcNAc of C₅₀-P-P-GlcNAc, thus yielding the full linkage unit C₅₀-P-P-GlcNAc-Rha of AG. The key to the discovery of mycobacterial WbbL was the successful complementation of an *E. coli* mutant lacking WbbL activity with the *Rv3265c* gene from *M. tuberculosis* (Mills *et al.*, 2004). *M. tuberculosis wbbL* was expressed in *E. coli* and together with bioinformatics analysis used to establish its preliminary structure and characteristics (Wu *et al.*, 2011). Moreover, it was demonstrated that *wbbL* (*MSMEG_1826*) is crucial to the growth and viability of *M. smegmatis* (Mills *et al.*, 2004).

The previously synthesised linker unit serves as an acceptor for the addition of Galf residues from the sugar donor UDP-Galf. Galactofuranosyltransferase (GalfT) GltT1 (*Rv3782*) recognises the linker unit and transfers the initial two Galf residues to C₅₀-P-P-GlcNAc-Rha resulting in C₅₀-P-P-GlcNAc-Rha-Galf₂ (Mikusova *et al.*, 2006, Alderwick

et al., 2008, Belanova *et al.*, 2008) (Figure 9). Further galactan polymerisation is carried out by the second transferase GlfT2 (*Rv3808c*), identified through the use of a neoglycolipid acceptor assay together with UDP-Galf and isolated *E. coli* membranes expressing *glfT2*. It was demonstrated that GlfT2 acts both as a UDP-Galf: β -D-(1 \rightarrow 5) GalfT and a UDP-Galf: β -D-(1 \rightarrow 6) GalfT, and is responsible for the majority of galactan biosynthesis by sequentially polymerising the galactan polysaccharide in alternating β (1 \rightarrow 5) and β (1 \rightarrow 6) linkages (Kremer *et al.*, 2001, Rose *et al.*, 2006, Wheatley *et al.*, 2012). In addition, structural data together with site-directed mutagenesis and kinetic studies provided evidence for a mechanism that explains the unique ability of GlfT2 to generate β (1 \rightarrow 5) and β (1 \rightarrow 6) linkages using a single active site (Wheatley *et al.*, 2012).

An arabinofuranosyltransferase (ArafT) from the *emb* locus, AftA (*Rv3792*), is responsible for addition of the first key Araf residue to the 8th, 10th and 12th Galf residues, thus “priming” the galactan chain for further attachment of α (1 \rightarrow 5)-linked Araf units (Alderwick *et al.*, 2005, Alderwick *et al.*, 2006) (Figure 9). A homologue of *aftA* in *M. smegmatis*, *MSMEG_6386*, was shown to be essential for survival of mycobacteria. However, deletion of *aftA* in *C. glutamicum* resulted in a slow growing, but viable mutant. Cell wall analysis revealed the complete loss of arabinose leading to a truncated cell wall structure containing only a galactan chain (Alderwick *et al.*, 2006). Further polymerisation of arabinan is predicted to be catalysed by EmbA (*Rv3794*) and EmbB (*Rv3795*) enzymes with α -1,5 transferase activity. Also, Emb proteins were demonstrated to play a key role in forming the terminal hexaarabinofuranosyl motif (Escuyer *et al.*, 2001). Although AG formation of mycobacteria is essential for its viability, the homologues of either *Rv3794* or *Rv3795* in *M. smegmatis* can be disrupted, leading to a

viable mutant with an impaired AG structure (Escuyer *et al.*, 2001). Interestingly, in comparison with *M. tuberculosis*, *C. glutamicum* has been shown to have only one *emb* gene, which is not essential in this organism (Alderwick *et al.*, 2005). Nevertheless, the deletion of this gene caused severe reduction of arabinose, resulting in a truncated AG structure with only terminal *Araf* residues (Alderwick *et al.*, 2005). Recent deletion studies in *M. smegmatis* identified a branching enzyme AftC (*MSMEG_2785*, orthologue to *Rv2673*) that is responsible for the transfer of *Araf* residues from DPA to the arabinan domain to form $\alpha(1\rightarrow3)$ -linked *Araf* residues of the internal arabinan domain at the non-reducing end of AG and LAM (Birch *et al.*, 2008, Birch *et al.*, 2010). Yet another functional *Araf*T encoded by *aftD* (*Rv0236c*) has been shown to have α -1,3-branching activity on linear α -1,5-linked synthetic acceptors *in vitro*. Inactivation of *aftD* homologue in *M. smegmatis*, *MSMEG_0359*, was shown to be lethal to mycobacteria, while overexpression of *aftD* in *M. smegmatis* resulted in an overall increase of *Araf* residues (Skovierova *et al.*, 2009). Finally, AftB (*Rv3805c*) catalyses the transfer of *Araf* residues to the arabinan domain to form the terminal $\beta(1\rightarrow2)$ linked *Araf* residues (Figure 9). Disruption of *aftB* in *C. glutamicum* resulted in a viable mutant with complete absence of terminal $\beta(1\rightarrow2)$ -linked arabinofuranosyl residues and decreased abundance of cell wall bound mycolic acids, consistent with a partial loss of mycolylation sites (Seidel *et al.*, 2007).

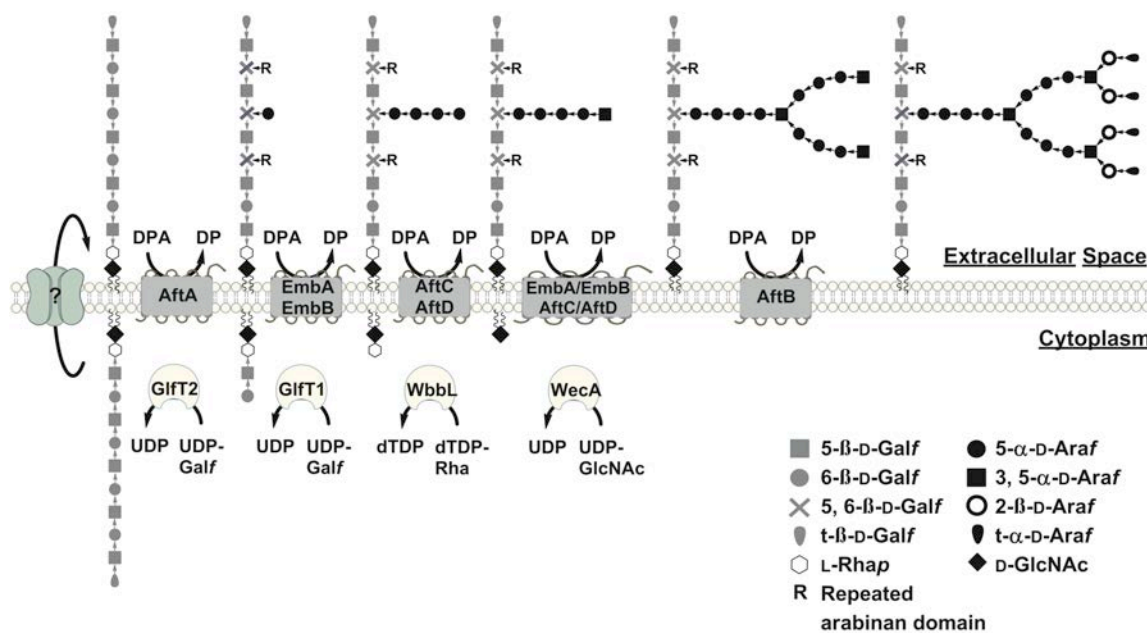


Figure 9 Schematic representation of mycobacterial arabinogalactan biosynthesis. WecA catalyses the transfer of GlcNAc to decaprenyl phosphate, which is then used as an acceptor for addition of rhamnosyl residue by WbbL, therefore forming the full linker unit. The first two galactofuranosyl (Galf) residues are added to the linkage unit *via* GifT1. The bifunctional GifT2 adds the remaining Galf residues forming a linear galactan chain. Before the polymerisation with arabinofuranosyl (Araf) residues, the galactan domain is thought to be translocated across the plasma membrane by the unknown flippase. AftA initiates the transfer of Araf residues from the sugar donor decaprenylphosphoryl-D-arabinofuranose (DPA) to the 8th, 10th and 12th β(1→6)-linked Galf residues of the galactan chain. EmbA and EmbB proteins act as α-1,5-arabinosyltransferases utilising the same nucleotide donor DPA. The 3,5-linked Araf branching is introduced by AftC and AftD enzymes. Finally, the terminal Araf residues are added to the arabinan domain by the “capping” enzyme AftB. Reprinted by permission from ASM: *Microbiology Spectrum*, Jankute *et al.* (2013).

1.8.2.4 Decoration of arabinogalactan and its attachment to peptidoglycan and mycolic acids

Two decorating structures, succinyl and GalN residues, have been identified in the interior AG arabinan domain of *M. tuberculosis*, thus concluding a model of the complete primary structure of mycobacterial AG. Enzymes involved in succinylation of arabinan

chains towards the non-reducing end are yet to be determined (Bhamidi *et al.*, 2008), however, the key components of the biosynthetic pathway of GalN have recently been elucidated (Skovierova *et al.*, 2010). Glycosyltransferase PpgS (*Rv3631*) catalyses the transfer of GalNAc from UDP-GalNAc to polyprenyl-P yielding a sugar donor polyprenyl-P-GalNAc. This high-energy substrate is then presumably deacetylated by an as yet unknown deacetylase before or after being translocated to the extracellular space where the membrane associated enzyme Rv3779 transfers the GalN_p (or GalNAc) residue to the C2 position of a portion of the internal 3,5-branched-D-Araf residues of AG. The synthesis of GalN was removed in both *ppgS* and *Rv3779* deletion mutants in *M. tuberculosis* (Skovierova *et al.*, 2010). It is worth noting that the GalN residue is only found in slow-growing mycobacteria. Hence expression of *ppgS* in the fast growing *M. smegmatis* species, otherwise devoid of the *ppgS* orthologue and any detectable polyprenyl-P-GalNAc synthase activity, allowed mycobacteria to synthesise polyprenyl-P-GalNAc *in vivo*. The physiological role and pathogenesis of both succinyl and GalN residues as well as the biosynthetic origin of succinylation remains to be elucidated.

Very little evidence has been obtained of how AG is ligated to PG to generate the complete cell wall core. *In vitro* assay in *M. smegmatis* utilizing cell-free extracts and radiolabelled substrates demonstrated the formation of simpler polyprenyl-P-P-GlcNAc-Rha-(Gal_f)_n intermediates followed by addition of AG and finally ligation to PG (Yagi *et al.*, 2003). However, there are no reports on which enzymes are involved in the attachment of AG to PG. In addition, it is not fully understood when the AG structure is mycolylated, before or after, the attachment to PG. *In vitro* enzymatic assays have identified members of the Antigen 85 complex, FbpA (*Rv3804c*), FbpB (*Rv1886c*) and

FbpC (*Rv0129c*), responsible for the transfer of mycolic acids onto trehalose leading to the formation of trehalose monomycolate (TMM) and trehalose dimycolate (TDM) (Belisle *et al.*, 1997). Inactivation of *fbpC* (previously termed *antigen 85C*) by transposon mutagenesis resulted in a mutant with reduced capacity to transfer mycolic acids to the mycobacterial cell wall (Jackson *et al.*, 1999). Similar mycolyltransferases from *C. glutamicum*, *cmytA* and *cmytB*, were deleted leading to a viable double mutant with significantly impaired ability to transfer corynomycolates to AG (Kacem *et al.*, 2004).

1.8.3 Lipoarabinomannan

PIMs, LM and LAM are glycopospholipids built on a phosphatidyl-myo-inositol (PI) backbone by a series of modifications including glycosylation and acylation. It is still unclear whether these PI based glycopospholipids and lipoglycans are embedded in the plasma membrane or found in the outer membrane of mycobacteria. However, using surface labelling experiments it has recently been demonstrated that the lipoglycans are exposed at the surface of mycobacteria, located in the outer leaflet of the outer membrane (Pitarque *et al.*, 2008). Such lipoglycans possess significant immunomodulatory effects in the host macrophages, including cytokine production, inhibition of phagosome maturation, and cross-protective immunity of mycobacteria (Mishra *et al.*, 2011a). Recent biochemical and genetic studies have indicated that modifications of the PI anchor follow the order: PI → PIM → LM → LAM (Mishra *et al.*, 2011a), where mannosylation of the PI anchor with four mannose residues produces PIM₄, which is further modified with mannose and arabinose residues to generate LM and LAM.

1.8.3.1 Structure of phosphatidyl-*myo*-inositol mannosides, lipomannan and lipoarabinomannan

PIMs are glycolipids comprised of fatty acids attached to a glycerol unit linked to *myo*-inositol *via* a phosphodiester moiety (Ballou *et al.*, 1963). The precursor of these glycolipids is the PI unit, which is based on *sn*-glycero-3-phosphate-(1-*D*-*myo*-inositol). PI, when modified with mannose residues at positions O-2 and O-6, forms the mannosyl phosphate inositol anchor (MPI) (Chatterjee *et al.*, 1992a, Severn *et al.*, 1998), which may differ in number, position and nature of the fatty acids attached. PIM₁ contains glycosylated PI with a α -*D*-mannopyranosyl (Manp) residue at position O-2, whereas PIM₂ contains glycosylated PI with Manp at positions O-2 and O-6. In addition, FAB-MS and GC-MS analysis of the perdeuteroacetyl and permethyl derivatives of PIMs from *M. tuberculosis* and *M. leprae*, has shown that position C-6 of PIM₂ mannose is esterified with a C₁₆ fatty acyl substituent (Khoo *et al.*, 1995b). There are four predicted sites of acylation in the MPI anchor: 1-OH and 2-OH of the glycerol unit, 3-OH of the inositol ring, and 6-OH of the Manp attached to position O-2 of the *myo*-inositol ring (Khoo *et al.*, 1995b). Based on the acylation sites, two types of PIMs were described in mycobacteria. The first type is Ac₁PIM_Y possessing an acyl group either at 3-OH of the inositol ring or at 6-OH of the Manp, whereas the second type is Ac₂PIM_Y, which is acylated at both of these positions. The acylated form of Ac₁PIM₂ serves as a substrate for higher order PIMs, such as Ac₁PIM₆ and Ac₂PIM₆. A combination of NMR spectroscopy and mass spectrometry determined that PIM₆ consists of Manp- α (1 \rightarrow 2)-Manp- α (1 \rightarrow 2)-Manp- α (1 \rightarrow 6)-Manp- α (1 \rightarrow 6)-Manp- α (1 \rightarrow .) (Figure 10) (Nigou *et al.*, 1999).

LAM is the extended form of LM and is comprised of a MPI anchor, polysaccharide backbone and capping moieties (Figure 9). The polysaccharide chain, common to both LAM and LM, contains 21-34 Man_p residues linked as a linear $\alpha(1\rightarrow6)$ chain with 5-10 units of single $\alpha(1\rightarrow2)$ Man_p side chains (Chatterjee *et al.*, 1992a, Dinadayala *et al.*, 2006) (Figure 9). In LAM, this linear chain also possesses an arabinan domain containing approximately 55-70 Araf residues linked as a linear $\alpha(1\rightarrow5)$ chain with α -1,3-branching (Dinadayala *et al.*, 2006, Birch *et al.*, 2010). Two types of branching patterns are observed in the arabinan domain of LAM. The first is a linear tetra-arabinoside comprising of β -D-Araf (1 \rightarrow 2)- α -D-Araf (1 \rightarrow 5)- α -D-Araf (1 \rightarrow 5)- α -D-Araf, whilst the second is a hexa-arabinoside [β -D-Araf (1 \rightarrow 2)- α -D-Araf]₂-3,5- α -D-Araf (1 \rightarrow 5)- α -D-Araf (Chatterjee *et al.*, 1991, Chatterjee *et al.*, 1993). The disaccharide unit, Araf- $\beta(1\rightarrow2)$ -Araf- $\alpha(1\rightarrow\cdot)$, is a common structural feature in both non-reducing ends of LAM arabinan. Finally, the capping moieties of arabinan contain manno-oligosaccharides (ManLAM) in slow growers, such as *M. tuberculosis* (Chatterjee *et al.*, 1993), phosphoinositide units (PILAM) in fast growers, such as *M. smegmatis* (Khoo *et al.*, 1995a), and uncapped LAM, known as AraLAM, in *M. chelonae* (Guerardel *et al.*, 2002). In addition, the mannan cap of ManLAM in *M. tuberculosis* contains mono-, di- and tri- $\alpha(1\rightarrow2)$ -D-Man_p saccharide units (Chatterjee *et al.*, 1992b, Chatterjee *et al.*, 1993).

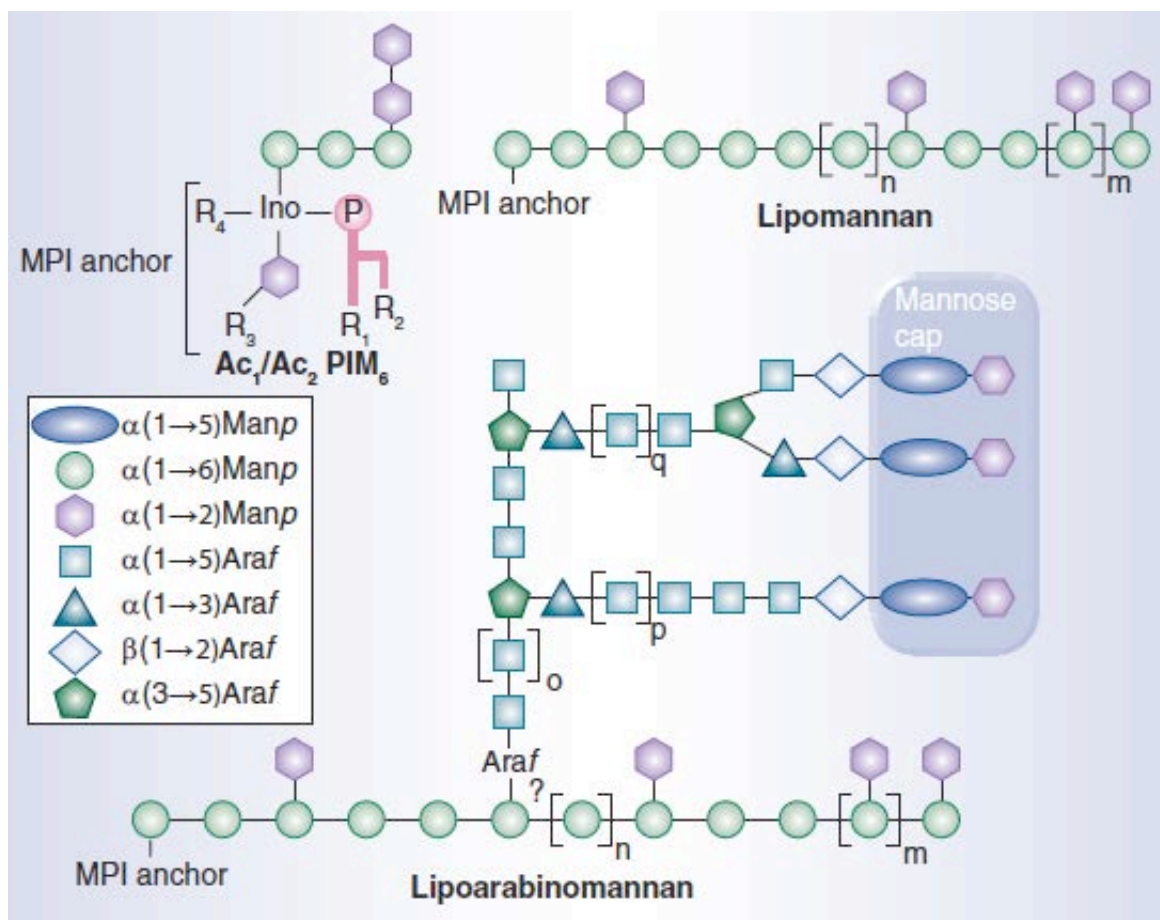


Figure 10 Structural features of acylated phosphatidyl-*myo*-inositol hexamannoside, lipomannan and lipoarabinomannan. Ac_1PIM_6/Ac_2PIM_6 are the end products of PIM biosynthesis generated by addition of four mannose residues to Ac_1PIM_2/Ac_2PIM_2 at position O-6 of inositol. The sugar backbone contains $Manp-\alpha(1\rightarrow2)-Manp-\alpha(1\rightarrow2)-Manp-\alpha(1\rightarrow6)-Manp-\alpha(1\rightarrow6)-Manp-\alpha(1\rightarrow\cdot)$ unit with acylation at either 3-OH of the inositol ring (R_3) or 6-OH of the $Manp$ attached to the O-2 position of the inositol ring (R_4). In case of Ac_2PIM_6 , both 3-OH and 6-OH are acylated. The 1-OH and 2-OH of the glycerol unit in the MPI anchor represent the acylation sites R_1 and R_2 , respectively. Lipomannan and lipoarabinomannan share the common MPI anchor, glycosylated at 2-OH and 6-OH of inositol. The common mannan core is composed of MPI anchor linked to 17-19 residues of $Manp$ via $\alpha(1\rightarrow6)$ linkage, branched frequently at position 2 with 5-10 single α - $Manp$ units. The distinctive feature of the lipoarabinomannan is the arabinan domain, which contains approximately 55-70 residues and is linked to the mannan core. To date, the point of attachment of arabinan to the mannan remains unknown. Finally, the non-reducing termini of arabinan are capped with mannose residues, thus forming the mannose cap. The alphabets n, m, o and p depict different extent of species specific glycosylations in LAM. Reprinted by permission from Future Medicine Ltd: *Future Microbiology*, Jankute *et al.* (2012).

1.8.3.2 Biosynthesis of phosphatidyl-*myo*-inositol mannosides, lipomannan and lipoarabinomannan

PIMs are synthesised employing glycosyltransferases PimA, PimB', PimC and PimE that sequentially catalyse the glycosylation of PI in the following order: PI \rightarrow PIM₂ \rightarrow PIM₄ \rightarrow PIM₆ (Besra and Brennan, 1997, Morita *et al.*, 2006, Mishra *et al.*, 2011a) (Figure 11). The biosynthesis is initiated by addition of a Manp residue to position O-2 of the inositol yielding PIM₁, which is then mannosylated at position O-6 of inositol ring, and further acylated to form Ac₁PIM₂ (Kordulakova *et al.*, 2002, Kordulakova *et al.*, 2003). Four more mannose residues are added at position O-6 of inositol in Ac₁PIM₂, thus generating Ac₁PIM₆ (Mishra *et al.*, 2011a). Ac₁PIM₂ acts as a scaffold in the PIM biosynthetic pathway on which other PIMs are assembled (Khoo *et al.*, 1995b, Besra and Brennan, 1997, Mishra *et al.*, 2011a).

PI synthesis begins with cyclisation of glucose-6-phosphate by inositol phosphate synthase (Ino1), encoded by *Rv0046c*, thus generating *myo*-inositol-1-phosphate. This intermediate is further dephosphorylated by inositol mono-phosphatase (IMP) yielding *myo*-inositol (Bachhawat and Mande, 1999, Movahedzadeh *et al.*, 2004). In *M. tuberculosis*, putative ImpA (*Rv1604*), SuhB (*Rv2701c*), CysQ (*Rv2131c*) and ImpC (*Rv3137*) proteins contain the characteristic IMP domain, are homologues of IMP enzyme and, therefore, are suggested to contribute to the biosynthesis of inositol (Nigou and Besra, 2002, Brown *et al.*, 2007, Movahedzadeh *et al.*, 2010). Surprisingly, only *impC* has been shown to be essential to the growth and viability of mycobacteria (Movahedzadeh *et al.*, 2010). PI biosynthesis is completed by PgsA (*Rv2612c*), which

catalyses esterification of the *myo*-inositol ring by transferring diacyl glycerol (DAG) from CDP-DAG (Jackson *et al.*, 2000).

PIM biosynthesis is initiated by PimA (*Rv2610c*), an α -mannopyranosyltransferase from GT-B superfamily (Guerin *et al.*, 2007), that transfers the first Manp residue from UDP-Manp to PI, thus generating PIM₁. PIM₁ can then be subsequently acylated by Rv2611, at the 6th position of the Manp residue to yield either Ac₁PIM₁ or Ac₂PIM₁ (Kordulakova *et al.*, 2002, Kordulakova *et al.*, 2003). PimB' (*Rv2188c*), an α -D-mannose- α -(1→6) phosphatidyl-*myo*-inositol-mannopyranosyltransferase, is responsible for addition of Manp to the 6-OH position of inositol to yield either Ac₁PIM₂ or Ac₂PIM₂. Interestingly, acylation of PIM₁ can precede the subsequent mannosylation step and is species dependent (Lea-Smith *et al.*, 2008, Mishra *et al.*, 2008b, Guerin *et al.*, 2009). Deletion of *pimB'* in *C. glutamicum* resulted in accumulation of Ac₁PIM₁, implying that acylation of PIM₁ occurred before the second mannosylation step (Kordulakova *et al.*, 2003, Lea-Smith *et al.*, 2008, Mishra *et al.*, 2008b). However, it was demonstrated that in *M. smegmatis*, Ac₁PIM₂ is generated after yielding PIM₂ (Guerin *et al.*, 2009). The *Rv0557c* was initially designated as PimB, an α -D-mannose- α -(1→6)-phosphatidyl-*myo*-inositol-mannopyranosyltransferase, and was estimated to perform the same function as PimB' (Schaeffer *et al.*, 1999). However, deletion of *pimB* in *M. tuberculosis* showed no effect on the biosynthesis of Ac₁PIM₂, indicating a case of redundancy or a different function (Torrelles *et al.*, 2009). Interestingly, *Rv0557c* homologue in *C. glutamicum* was demonstrated to contribute to the synthesis of Cg-LM type-B (a LM variant) and ManGlcAGroAc₂, a glucuronic acid diacyl-glycerol based glycolipid (Mishra *et al.*, 2009). Complementation studies utilising *Rv2188c* and *Rv0557c* orthologues as well as

C. glutamicum Δ pimB' Δ mgtA double mutant were performed in order to elucidate the role of Rv0557. The mutant complemented with Rv2188c demonstrated restoration of α -D-mannose- α -(1 \rightarrow 6)-phosphatidyl-*myo*-inositol-mannopyranosyltransferase activity, which led to the synthesis of Ac₁PIM₂ and the corresponding lipoglycan. However, complementation of *C. glutamicum* Δ pimB' Δ mgt with Rv0557c resulted in the production of ManGlcAGroAc₂ (Mishra *et al.*, 2009). As a result, *C. glutamicum* homologue of Rv0557c has been designated as MgtA with an α -mannosyl-glucopyranosylglucouronic acid transferase activity (Tatituri *et al.*, 2007b). Recently, the crystal structure of PimB' co-crystallized with GDP and GDP-Man has been elucidated (Batt *et al.*, 2010).

RvD2-ORF1 from *M. tuberculosis* CDC1551, designated as PimC with Ac₁PIM₂. α -D-mannose- α -(1 \rightarrow 6)-phosphatidyl-*myo*-inositol-mannopyranosyl transferase activity, catalyses the addition of a Manp residue to 6-OH of mannose at the non-reducing end of Ac₁/Ac₂PIM₂ to generate Ac₁/Ac₂PIM₃. Overexpression of PimC in a cell-free assay containing GDP-[¹⁴C]mannose, amphomycin (inhibitor of polyprenol phosphate-requiring transferases) and membranes from *M. smegmatis* led to the synthesis of Ac₁/Ac₂PIM₃. However, its inactivation in *M. bovis* BCG had no effect on PIM, LM and LAM biosynthesis, suggesting existence of redundant gene(s) or an alternative pathway that may compensate for PimC deficiency (Kremer *et al.*, 2002). Polyprenylphosphate based mannose donors are employed for further elongation and branching of LM and LAM (Guerin *et al.*, 2010). The α (1 \rightarrow 6) mannosylation of Ac₁/Ac₂PIM₃ to Ac₁/Ac₂PIM₄ is catalysed by either PimC or an unidentified mannopyranosyltransferase (Mishra *et al.*, 2011a). The pathway divides into two branches from Ac₁/Ac₂PIM₄: one leads to the formation of Ac₁/Ac₂PIM₆, whilst the other mediates the formation of LM and LAM. In

M. smegmatis, the LpqW channel regulates the amount of higher PIMs and lipoglycans formed, and directs Ac_1/Ac_2PIM_4 towards LM synthesis (Kovacevic *et al.*, 2006, Crellin *et al.*, 2008). In the first branch, $\alpha(1\rightarrow2)$ Manp is added to Ac_1/Ac_2PIM_4 to form Ac_1/Ac_2PIM_6 by two consecutive mannose additions catalysed by PimE (*Rv1159*), an $\alpha(1\rightarrow2)$ -mannopyranosyl-transferase, and an additional uncharacterised putative glycosyltransferase speculated to be encoded by either *Rv0051* or *Rv0541* (Liu and Mushegian, 2003, Berg *et al.*, 2007). In the second branch, Ac_1/Ac_2PIM_4 is hypermannosylated by glycosyltransferases of the GT-C superfamily to generate LM and LAM. Recent studies have established the orthologs of MptB (*Rv1459c*) and MptA (*Rv2174*) in *C. glutamicum* to be involved in the synthesis of the mannan backbone, where MptB catalyses the synthesis of the proximal end through addition of 12-15 Manp residues to the backbone (Mishra *et al.*, 2008a) and MptA that synthesises the distal end of the $\alpha(1\rightarrow6)$ mannan core of LM (Kaur *et al.*, 2007, Mishra *et al.*, 2007). However, failure of MptB from *M. tuberculosis* to complement the *C. glutamicum* Δ *mptB* strain and its redundancy in *M. smegmatis* suggests that an uncharacterised orthologue substitutes for MptB, or MptA itself is involved in the extension of the proximal end (Mishra *et al.*, 2008a). The $\alpha(1\rightarrow6)$ -mannan core synthesised by MptA and MptB is further branched by MptC (*Rv2181*) which adds $\alpha(1\rightarrow2)$ -Manp residues to the side chain of LAM (Mishra *et al.*, 2011b).

The transition from LM to LAM is exclusively catalysed by EmbC (*Rv3793*) and involves the utilisation of LM primed with Araf units (Shi *et al.*, 2006, Alderwick *et al.*, 2011a). An uncharacterised ArafT is suggested to perform the addition of first Araf residues to LM in a manner similar to that of AG that is catalysed by AftA (Alderwick *et*

al., 2006). EmbC is responsible for extension of the primed LM through addition of 12-16 $\alpha(1\rightarrow5)$ -Araf residues (Mishra, 2011). The addition of branch points is similar to AG, catalysed by AftC (Birch *et al.*, 2008, Mishra, 2011). Skovierova *et al.* (2009) has identified a second branching enzyme termed AftD (*Rv0236c*) with an $\alpha(1\rightarrow3)$ -ArafT activity (Skovierova *et al.*, 2009). The arabinan domain is estimated to be terminated by AftB (*Rv3805c*) (Seidel *et al.*, 2007). Although AftB primarily has $\beta(1\rightarrow2)$ -ArafT activity in AG biosynthesis, its probable role in LAM biosynthesis renders it bifunctional (Mishra *et al.*, 2011a).

Finally, the homologue of *Rv1635c* in *M. tuberculosis* CDC1551 was proposed to be involved in Man-LAM capping. *Rv1635c* mutants in *M. bovis* BCG and *M. marinum* are known to synthesise LAM lacking Man-caps (Appelmelk *et al.*, 2008). This enzyme was later revealed to be an undecaprenyl phosphomannose-dependent $\alpha(1\rightarrow5)$ mannosyltransferase, termed CapA. It recognises the fully synthesised arabinan domain and transfers the first Man_p residue onto the non-reducing arabinan termini of LAM (Dinadayala *et al.*, 2006, Appelmelk *et al.*, 2008, Mishra *et al.*, 2011a). ManLAM is completed by MptC (*Rv2181*) that elongates the cap with a second Man_p residue (Mishra *et al.*, 2011a).

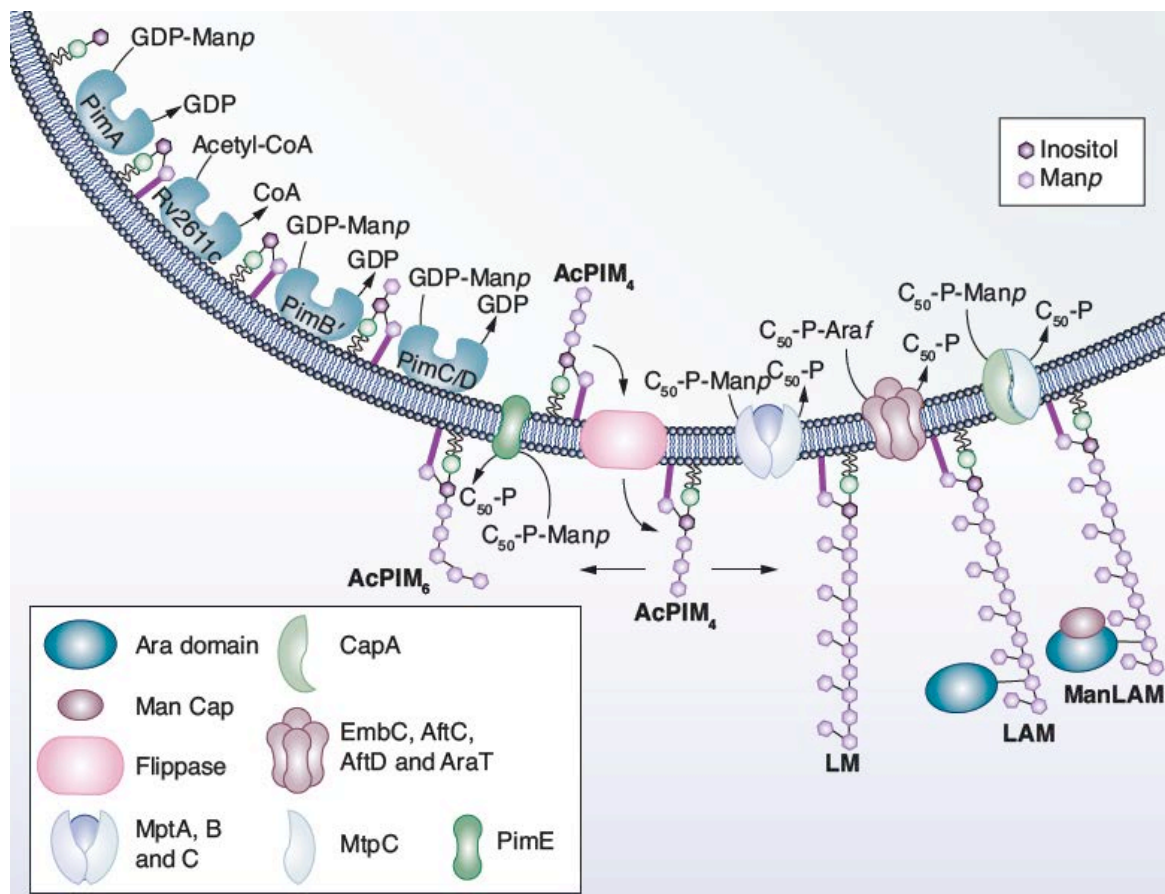


Figure 11 Biosynthesis of phosphatidyl-*myo*-inositol mannosides, lipomannan and lipoarabinomannan in *M. tuberculosis*. The first three mannosylation steps employ GDP-Man dependent PimA, PimB', PimC, and possibly PimC or an unidentified mannosyltransferase. Acylation of PIM₁ is mediated by Rv261c. AcPIM₄ is flipped over to the periplasmic side by an unidentified flippase, where it is mannosylated by PimE to generate the metabolic end product - AcPIM₆. AcPIM₄ is sequentially extended by MptB, MptA, and MptC to yield LM. An unidentified arabinofuranosyltransferase primes LM with Araf residues. EmbC, AftC and AftD subsequently extend and branch arabinan domain, while AftB terminates the chain extension, thus completing LAM. LAM is further capped by CapA in conjunction with Rv2181 to generate ManLAM. Reprinted by permission from Future Medicine Ltd: *Future Microbiology*, Jankute *et al.* (2012).

1.8.4 Ethambutol inhibits arabinogalactan and lipoarabinomannan biosynthesis

Ethambutol [EMB; (S, S)-2, 2'-(ethylenediimino)di-1-butanol] is a bacteriostatic agent extensively used as a front-line drug for the treatment of TB since 1966. It is an integral

component of the multi-drug regimen used worldwide for TB therapy in combination with INH, PZA and RIF to prevent the emergence of drug-resistant TB strains. Although, this anti-mycobacterial agent has been administered for several decades, the mechanism of action and molecular genetic basis of resistance of this drug is yet to be fully understood. However, over the last forty years significant progress has been made in efforts to elucidate the intracellular mode of action of EMB.

Earlier studies into the investigation of the intracellular target showed that treatment of *M. smegmatis* with EMB leads to rapid declumping of bacterial cells, indicating EMB involvement in alteration of the cell wall (Kilburn *et al.*, 1977). The examination of the effects of EMB on mycobacteria has led to several hypotheses being proposed in relation to the mechanism of action over the years including inhibition of nucleic acid metabolism (Forbes *et al.*, 1965), phospholipid metabolism (Kilburn *et al.*, 1981), spermidine biosynthesis (Poso *et al.*, 1983) and glucose metabolism (Silve *et al.*, 1993). EMB was shown to act on the accumulation and secretion of trehalose mycolates and free mycolic acids in *M. smegmatis* leading to the rapid cessation of mycolic acid transfer to the cell wall (Kilburn and Takayama, 1981), establishing its effect on cell wall synthesis. However, the primary mode of action of EMB was proposed to specifically target the inhibition of AG biosynthesis (Takayama and Kilburn, 1989) and the synthesis of the arabinan core of LAM (Deng *et al.*, 1995), which leads to the accumulation of DPA, a donor in the pathway of arabinan biosynthesis (Wolucka *et al.*, 1994).

The *embAB* gene cluster in *Mycobacterium avium*, was found to confer resistance to EMB when subcloned and overexpressed into *M. smegmatis*, and was first presented as the precise cellular target of EMB (Belanger *et al.*, 1996). The *embAB* genes encode

AraT involved in the polymerisation of arabinose into AG (Belanger *et al.*, 1996). Subsequent studies identified a three gene operon, *embCAB* in *M. tuberculosis* and *M. smegmatis* as a putative target of EMB, which encode three AraT enzymes that display approximately 65% similarity to each other (Telenti *et al.*, 1997). The discovery of the *emb* gene cluster provided the opportunity to examine the molecular basis of resistance of mycobacteria. Telenti *et al.* (1997) demonstrated that high-level resistance to EMB might result from either overproduction of Emb protein(s) or a structural mutation in the EmbB protein. Further studies have shown genetic evidence for a key role of the EmbB protein in cell wall biosynthesis highlighting the fact it is the most EMB-sensitive protein in the gene cluster (Alcaide *et al.*, 1997).

Point mutations were identified in codons 289 and 292 of the *embB* gene of *M. smegmatis* that conferred resistance to EMB (Lety *et al.*, 1997). As the *embB* region is highly conserved amongst *M. tuberculosis*, *M. avium*, *M. leprae* and *M. smegmatis*, similar mutations were found in codons 303 and 306 in the *embB* gene of EMB-resistant clinical isolates of *M. tuberculosis* (Lety *et al.*, 1997, Telenti *et al.*, 1997). The most frequently detected mutation occurred in codon *embB*306 resulting in a substitution of the wild-type methionine with isoleucine, leucine or valine, and was found in 50% of all EMB-resistant clinical isolates, which has been proposed to be positioned in a cytoplasmic loop that forms an EMB resistance-determining region (ERDR) (Telenti *et al.*, 1997). This mutation was proposed as a key mechanism of resistance and a possible candidate marker for rapid detection of MDR-TB and XDR-TB strain (Starks *et al.*, 2009). However, it is controversial as to whether it mediates EMB resistance due to the presence of this mutation in EMB-susceptible *M. tuberculosis* isolates (Lee *et al.*, 2004, Hazbon *et al.*,

2005, Shen *et al.*, 2007). Furthermore, it was shown that mutations in *embB306* may predispose *M. tuberculosis* isolates to broad drug resistance rather than EMB resistance alone (Hazbon *et al.*, 2005). Recent studies have reinforced the association of *embB306* mutations with EMB resistance (Starks *et al.*, 2009) as shown by experimental evidence that the transfer of *embB306* mutations to EMB susceptible *M. tuberculosis* strain increases the EMB minimum inhibitory concentration (MIC) *in vitro* (Safi *et al.*, 2008). Although, the role of *embB* in relation to mediating resistance has been under much scrutiny, studies have indicated the existence of multiple molecular pathways to the EMB resistance phenotype (Ramaswamy *et al.*, 2000). For example, there are a number of EMB resistant strains that do not possess ERDR mutations, thus other genes may be involved in EMB resistance.

Inactivation of *embA* and *embB* in *M. smegmatis* resulted in a greater reduction of the arabinose content of arabinogalactan compared to levels observed in an *embC* deletion mutant (Escuyer *et al.*, 2001). The hexaarabinofuranosyl motif was altered in both *embA* and *embB* deletion mutants in *M. smegmatis* when compared to the wild-type demonstrating both EmbA and EmbB proteins are involved in the formation of AG (Escuyer *et al.*, 2001). Inactivation of *embC* in *M. smegmatis* resulted in no detectable LAM content demonstrating EmbC is required for synthesis of LAM (Zhang *et al.*, 2003). The deletion of Emb proteins and generation of viable mutants in *M. smegmatis* suggests that EMB targets two or more Emb proteins to cause growth inhibition. Since the *embC* deletion mutant could not be restored to its phenotype by complementation, it could not be confirmed that EmbC is critical for AG production in *M. smegmatis* (Escuyer *et al.*, 2001). EmbA and EmbC has recently been shown to be independently

essential in *M. tuberculosis* implying that EMB can target either of these proteins to arrest bacterial growth (Amin *et al.*, 2008, Goude *et al.*, 2008). Thus, the consequences of EMB action may differ between fast growing and slow growing mycobacteria.

1.8.5 Mycolic acids

Mycolic acids are vital components of the waxy cell wall of *Corynebacterineae* including *M. tuberculosis*, *M. smegmatis*, and *C. glutamicum* and are located either covalently attached to AG, or as free glycolipids: trehalose monomycolates (TMM), trehalose dimycolates (TDM), and glucose monomycolates (GMM) (Bhatt *et al.*, 2007, de Souza *et al.*, 2008, Luo *et al.*, 2012, Verschoor *et al.*, 2012). These α -branched, β -hydroxylated fatty acids account for reduced cell wall permeability and virulence in pathogenic mycobacteria (Asselineau and Lederer, 1950). Notably, the length of the fatty acids is species dependent. For instance, mycobacterial mycolic acids are 70-90 carbons in lengths (Goodfellow and Minnikin, 1977), whereas in corynebacteria they are 22-38 carbons (Collins *et al.*, 1982a, Collins *et al.*, 1982b). In *M. tuberculosis*, mycolates are generally composed of C₂₄-C₂₆ saturated straight chain fatty acids to provide the α -alkyl branch and the meromycolate chain of up to 56 carbons in length. Three different distinct classes of *M. tuberculosis* mycolic acids have been identified: α -, methoxy-, and keto-. α -Mycolates are the most abundant variety, which is non-oxygenated at both proximal and distal positions of the meromycolate chain and comprising of two cyclopropane rings in the *cis* configuration (Qureshi *et al.*, 1978, Yuan *et al.*, 1995, Minnikin *et al.*, 2002). Methoxy- and keto-mycolates are both oxygenated lipid species containing one cyclopropane ring either in the *cis* and *trans* configuration, respectively (George *et al.*, 1995). Mycolic acids are also present in the cell envelope as noncovalently associated

glycolipids. TMM contains a trehalose disaccharide esterified to one mycolic acid residue, whereas TDM is esterified to two mycolic acid residues. GMM is composed of a glucose unit and one mycolic acid residue.

The biosynthesis of mycolic acids in mycobacteria is mediated through two fatty acid synthases: fatty acid synthase I (FAS-I) generates the C₂₄-C₂₆ saturated straight chain fatty acids, whereas fatty acid synthase II (FAS-II) provides the meromycolate backbone - C₅₆ fatty acids (de Souza *et al.*, 2008, Luo *et al.*, 2012). FAS-I is a large eukaryote-like multifunctional enzyme encoded by a single *fas* (*Rv2524c*) gene. FAS-II further elongates the FAS-I end products, thus forming meromycolates (Bhatt *et al.*, 2007). FAS-II multi-enzyme synthase is comprised of a number of proteins including *trans*-2-enoyl-ACP reductase *InhA* (*Rv1484*), β -ketoacyl-ACP synthase composed of *KasA* (*Rv2245*) and *KasB* (*Rv2246*), β -ketoacyl-ACP reductase *MabA* (*Rv1483*), β -hydroxyacyl-ACP dehydratase comprised of *HadA* (*Rv0635*), *HadB* (*Rv0636*) and *HadC* (*Rv0637*). Subsequently, meromycolate chains are then modified introducing functional groups by various methyltransferases (Glickman *et al.*, 2000, Glickman, 2003, Takayama *et al.*, 2005). The α -branch and the meromycolate backbone are then linked in a condensation reaction catalyzed by *Pks13* (*Rv3800c*) (Portevin *et al.*, 2004), and reduced by *Rv2509* to finally form mature mycolic acids (Bhatt *et al.*, 2008). Mycolyltransferases (the Antigen 85 complex) catalyse the transfer of these newly synthesized mycolic acids to AG or other acceptors (Belisle *et al.*, 1997, Jackson *et al.*, 1999). A number of enzymes involved in mycolic acid biosynthesis have been identified, however further characterization of these key enzymes is needed to fully understand their roles in the physiology of mycobacteria and the identification of new drug targets.

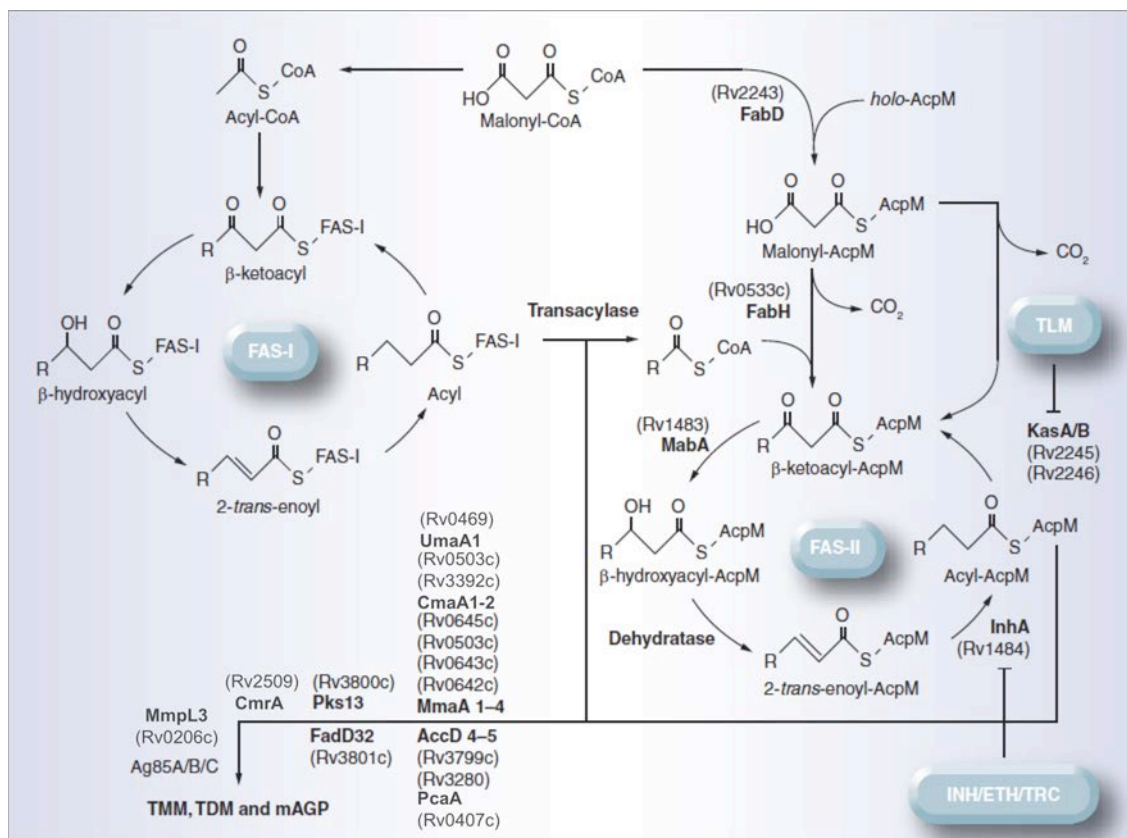


Figure 12 *Mycobacterium tuberculosis* mycolic acid biosynthesis. FAS-I, and then FAS-II synthesise the long aliphatic acyl-chains that are further modified to produce the mature meromycolate chains of mycolic acids. Key steps in the biosynthesis of mycolic acids as well as the targets for the action of activated isoniazid (INH), ethionamide (ETH), triclosan (TRC), and thiolactomycin (TLM) are shown as indicated. TMM: trehalose monomycolate; TDM: trehalose dimycolate; mAGP: mycolyl-arabinogalactan-peptidoglycan complex. Reprinted by permission from Future Medicine Ltd: Future Microbiology *et al.* (2012).

1.9 Surrogate systems in *M. tuberculosis* research

A distinctive quality of mycobacterial pathogens, such as *M. tuberculosis* and *M. leprae*, is their slow growth rate and fastidious culturing process. Therefore, various surrogate systems have been employed to facilitate *M. tuberculosis* research. A significant surrogate organism used in *M. tuberculosis* research is *Mycobacterium smegmatis*. In 1884, Lustgarten discovered and isolated *M. smegmatis* wild-type strain, which is fairly

common in soil and aquatic environments (Lustgarten, 1884). Commercially, *M. smegmatis* is utilised for production of xylitol (dietary sweetener) and L-arabinose, whereas in scientific research it is most widely used in investigating mycobacterial pathogens. Classified as saprophytic species, *M. smegmatis* is one of the most rapid growing mycobacteria with a generation time of approximately 4 hours. Importantly, *M. smegmatis* shares a highly similar cell wall structure to other mycobacteria, including *M. tuberculosis*. Ease of genetic manipulation has allowed characterisation of a large number of *M. smegmatis* genes, some of which are homologues to *M. tuberculosis* (Bhatt and Jacobs, 2009, Rana *et al.*, 2012, Varela *et al.*, 2012). However, a limitation of using *M. smegmatis* as a model organism is the study of inhibitors of virulent *M. tuberculosis*. *M. smegmatis* is a non-pathogenic species lacking approximately 30 % of conserved homologues in *M. tuberculosis* (Altaf *et al.*, 2010). Therefore, although *M. smegmatis* offers technical benefits, it is inappropriate to study pathogenesis and is less efficient when investigating potential drug targets.

M. bovis, a member of MTBC, is the causative agent of bovine TB. It has been also found to cause disease in humans and other mammals that are not considered to be its primary host, therefore making *M. bovis* a zoonotic organism. *M. bovis* Bacillus Calmette Guerin (BCG) is a strain that was attenuated from *M. bovis* by serial passages in the laboratory and is the vaccine used for TB (King and Park, 1929, Kaplan, 1947). Although *M. bovis* BCG is a slow grower, it can be used under less stringent conditions. More importantly, *M. bovis* BCG shares approximately 99 % of conserved genes identical to *M. tuberculosis* and is amenable to genetic and molecular manipulations. However, the disadvantage of

using *M. bovis* BCG is that it lacks virulence factors that were lost in the process of attenuation.

Another widely used model organism to study *M. tuberculosis* and other pathogenic mycobacteria is *Mycobacterium marinum*. It is an opportunistic human pathogen that causes TB like infections in fish and amphibians. *M. marinum* has a shorter generation time (8-12 hours) than compared to *M. tuberculosis*, responds to genetic manipulation and is a Category 2 organism. In addition, the genome comparisons of both *M. marinum* and *M. tuberculosis* demonstrate a close genetic relationship sharing over 3000 conserved orthologs (Stinear *et al.*, 2008). Significantly, *M. marinum* has been demonstrated to enter and replicate within the macrophages in a similar manner to *M. tuberculosis* (Stamm and Brown, 2004). Therefore, *M. marinum* is widely used to study mechanisms of virulence as well as host-pathogen interactions utilising its primary host. In addition, fish infected with *M. marinum* develop granulomas, which is one of the main features of the host immune response to *M. tuberculosis* infection (Chan *et al.*, 2002, Davis *et al.*, 2002). Finally, infection studies are usually utilizing zebrafish those embryos are transparent, thus facilitating the observation process.

Corynebacterium glutamicum has been isolated in the early 20th century in Japan where it was established in the industrial production of L-glutamate (Abe, 1967). *C. glutamicum* along with *Corynebacterium diphtheriae*, the causative agent of diphtheria, and *Corynebacterium pseudotuberculosis*, the causative agent of caseous lymphadenitis, belong to the same *Corynebacteriaceae* family as *M. tuberculosis* and, thus, share a similar cell wall architecture. However, the cell envelope of *C. glutamicum* is less complex than compared to mycobacteria and, therefore, more tolerant towards deletions

of homologous genes that are essential in, for example, *M. tuberculosis* (Alderwick *et al.*, 2006, Seidel *et al.*, 2007, Birch *et al.*, 2008). In addition, *C. glutamicum* is a non-pathogenic bacterium with a generation time of 30 minutes and a number of genetic and molecular techniques that are available to manipulate it, thus making it an attractive model organism to study mycobacterial species.

1.10 Project aims

Over the past decade, researchers have been successful in identifying and characterising a number of enzymes involved in AG and LAM biosynthesis in *C. glutamicum*, *M. smegmatis* and *M. tuberculosis*. However, many components of these complex pathways, including multiprotein complex formation, still remain elusive. Furthermore, aspects of cell wall biosynthesis, such as regulation and turnover, remain poorly studied.

The overall aim of this project was to gain an insight into the biosynthetic pathways leading to formation of AG and LAM with specific focus on characterisation of glycosyltransferases and their protein-protein interactions. The main project was split into three minor objectives:

A. Construction and characterisation of knockout mutants of *aftA*, *aftB* and *emb* from *Corynebacterineae*

This aim involved the construction of null mutants of either *M. smegmatis* mc²155 or *C. glutamicum* ATCC 13032 non-essential genes in order to determine their function in cell wall biosynthesis. In the case of genes being essential to *M. smegmatis* mc²155, conditional mutants were generated. Conditional expression-specialised transduction essentiality test (CESTET) is an efficient and relatively quick method of determining

essentiality of individual genes in mycobacteria (Bhatt and Jacobs, 2009). Biochemical phenotypes and intermediates were analysed.

B. Biochemical and structural characterisation of AftA, AftB and Emb from *Corynebacterineae*.

Proteins encoded by candidate genes were overexpressed either in *E. coli* or *C. glutamicum* cells and purified for further analysis. *In vitro* activity assays using radioactive labelled substrates for individual glycosyltransferases were attempted. Furthermore, C-terminal hydrophilic domains of membrane proteins encoded by some of the candidate genes were overexpressed and purified for crystallisation studies. This approach was recently proved to be successful when investigating membrane glycosyltransferases with large characteristic hydrophilic loops (Alderwick *et al.*, 2011a).

C. Investigation of protein interaction network of AG biosynthetic enzymes.

In order to assess protein-protein interactions of proteins encoded by candidate genes various methods were approached. The list of techniques included the *in vivo* bacterial two-hybrid system (BACTH), a novel lipodisq technology for membrane proteins, and pull-down assays. These methods provided an insight into physical interactions between proteins encoded by candidate genes and their interacting partners.

Chapter 2

Elucidation of a protein-protein interaction network involved in *Corynebacterium glutamicum* cell wall biosynthesis as determined by bacterial two-hybrid analysis

2 Elucidation of a protein-protein interaction network involved in *Corynebacterium glutamicum* cell wall biosynthesis as determined by bacterial two-hybrid analysis

2.1 Introduction

Over the past decade sophisticated genomic and molecular tools have advanced our understanding of biochemical and physiological processes of mycobacteria. The availability of the full genome sequences of various mycobacterial species, including *M. tuberculosis* (Cole *et al.*, 1998), *Mycobacterium marinum* (Stinear *et al.*, 2008), *Mycobacterium leprae* (Eiglmeier *et al.*, 2001) and *Mycobacterium* BCG (Garnier *et al.*, 2003), have greatly facilitated the identification of large numbers of drug targets and antigens. Moreover, numerous genome sequences and rapidly evolving field of comparative genomics have allowed the research to focus on identifying and characterising gene products that play an important role in growth, viability, and virulence of *M. tuberculosis*. It has become increasingly obvious that networks of interacting proteins mediate pathways of these cellular processes. Therefore, methods to investigate protein-protein interactions have emerged as a valuable tool in the field of *M. tuberculosis* research. Currently, mycobacterial protein associations have been employed to study numerous cellular processes, such as signal transduction pathways (Lee *et al.*, 2012), inhibitors of mycobacterial interactions (Mai *et al.*, 2011), formation of enzymatic complexes (Baulard *et al.*, 2003, Veyron-Churlet *et al.*, 2004, Veyron-Churlet *et al.*, 2005), and drug resistance pathways (Raman and Chandra, 2008, Cui *et al.*, 2009, Wang *et al.*, 2010).

The essential mycobacterial cell wall structure –mAGP– represents a very attractive drug target. A key structural component of the mAGP complex is AG that plays an important role in covalently anchoring the mycolic acid layer to the inner PG. Although, years of

research have obtained structural, physical and chemical evidence of this unique cell structure (General Introduction 1.8.2), there are still substantial gaps in our knowledge of AG biosynthesis. For instance, little is known about multi-protein complexes involved in AG assembly, perhaps due to a number of cell wall biosynthetic proteins being transmembrane or membrane bound. Recently, Zheng *et al.* (2011) shed light on potential protein-protein associations in membrane fractions of *M. bovis* BCG utilising blue native polyacrylamide gel electrophoresis (BN-PAGE) combined with liquid chromatography tandem mass spectrometry (LC-MS). Nine protein clusters have been identified representing multi-protein complexes. Notably, AraFTs encoded by *emb* genes *embA*, *embB*, and *embC* were identified in the study as a possible Emb protein complex (Zheng *et al.*, 2011). EmbA and EmbB are responsible for elongation of the arabinan domain of AG (Escuyer *et al.*, 2001, Alderwick *et al.*, 2005), while EmbC is involved in the synthesis of LAM (Goude *et al.*, 2008). Studies demonstrated that the EMB, a first line anti-TB drug, specifically inhibits AG biosynthesis with the precise molecular target being the *embCAB* locus in *M. tuberculosis* (Telenti *et al.*, 1997). Disruption of these proteins leads to inhibition of mAGP synthesis and perhaps increased permeability of mycobacterial cell envelope (Telenti *et al.*, 1997). Therefore, Emb proteins represent ideal targets for anti-TB drug development. Another recently published study has identified an *in vivo* interaction between GlfT1, responsible for initiating the polymerisation of the galactan domain of AG, and the small multi-drug resistance-like transporter Rv3789 (Larrouy-Maumus *et al.*, 2012). It has been proposed that Rv3789 is involved in AG and LAM biosynthesis, where it translocates the DPA from the cytosol to the periplasmic space, enabling its utilisation by AraFTs, such as EmbA, EmbB and EmbC (Larrouy-Maumus *et al.*, 2012).

The closely related species *M. smegmatis* and *C. glutamicum* have proven useful as laboratory model organisms in the study of orthologues *M. tuberculosis* genes and proteins involved in mAGP synthesis (Alderwick *et al.*, 2006, Seidel *et al.*, 2007, Birch *et al.*, 2008). Based on the use of pull-down assays and a bacterial two-hybrid system (BACTH) this study examines the construction of a protein-protein interaction network for the enzymes involved in the synthesis of AG, a major cell wall component in *C. glutamicum*. The pull-down binding assay is a relatively easy and straightforward technique to identify physical interactions between a tagged protein and its interacting partners *in vitro*. A tagged 'bait' protein is immobilised by the affinity of the ligand unique for the tag, followed by the incubation with the source of protein (e.g cell lysate or membrane preparation) that contains 'prey' proteins (Figure 1). Eluted protein complexes are then analysed by gel electrophoresis and mass spectrometry. BACTH, on the other hand, is based on the functional complementation between two fragments of the adenylate cyclase *in vivo* to restore cAMP signaling cascade in *Escherichia coli* (Figure 2) (Karimova *et al.*, 1998). Importantly, BACTH is able to detect physical interactions between both cytoplasmic, as well as membrane proteins (Karimova *et al.*, 2005, Maxson and Darwin, 2006, Hara *et al.*, 2008, Karimova *et al.*, 2009, Georgiadou *et al.*, 2012). This is extremely useful, since many *C. glutamicum* cell wall glycosyltransferases are membrane proteins.

In this study, we have identified twenty-four putative homotypic and heterotypic protein interactions *in vivo*. Our results demonstrate an association between glycosyltransferases, GlfT1 and AftB, and interaction between the sub-units of decaprenylphosphoribose epimerase, DprE1 and DprE2. These analyses have also show that AftB interacts with AftA, which catalyses the addition of the first three arabinose units onto the galactan

chain. Both AftA and AftB associate with other Ara/Ts, including Emb and AftC that elongate and branch the arabinan domain, respectively. Moreover, a number of proteins involved in AG biosynthesis were shown to form dimers or multimers. These findings provide a useful recourse for understanding the biosynthesis and function of the mycobacterial cell wall, as well as providing new therapeutic targets.

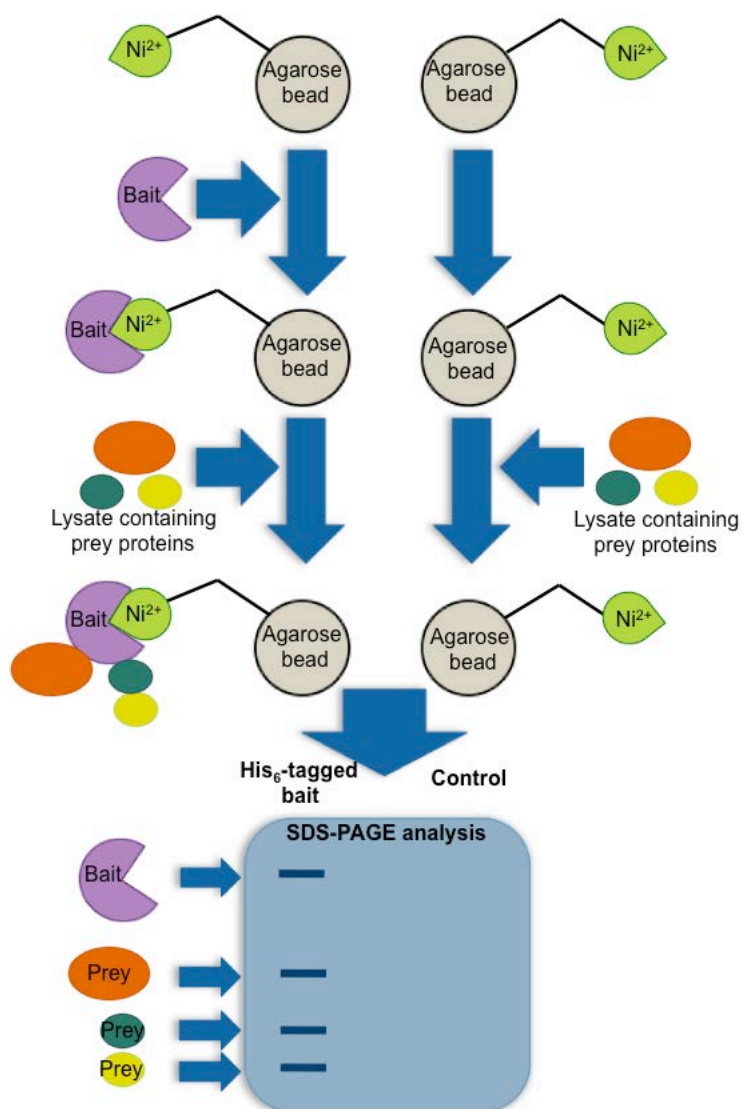


Figure 1 Generalised scheme of a pull-down assay for detection of putative protein-protein interactions. The purified 'bait' protein carrying a His₆-tag is captured on agarose beads charged with Ni²⁺. The unbound protein is removed by a washing step and the lysate containing putative 'prey' proteins is incubated with immobilised 'bait'. Several washing steps are performed to wash away the unbound or weakly bound proteins and possible protein-protein complexes are eluted with buffer containing increasing amounts of imidazole. Samples are analysed by 4-15 % SDS-PAGE and Western blot using monoclonal anti-His₆ antibody.

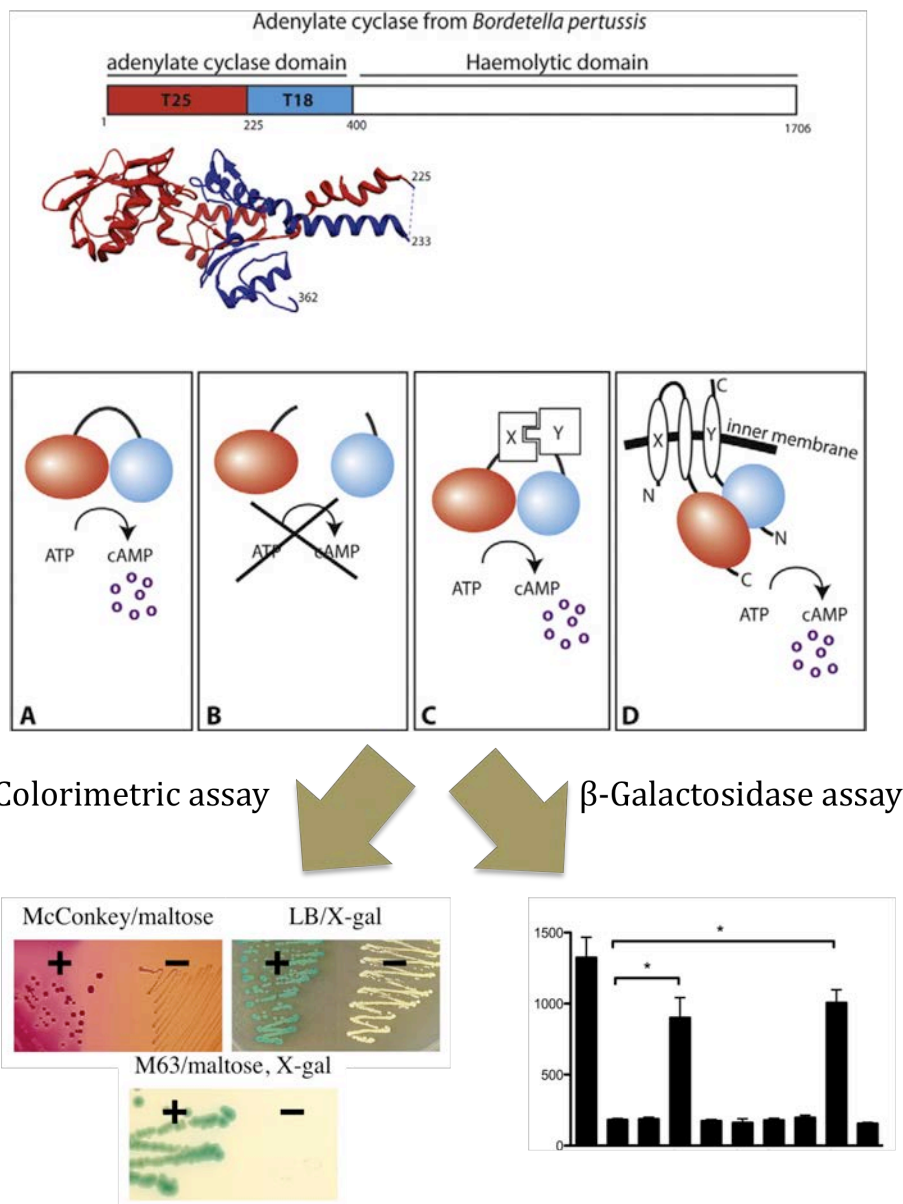


Figure 2 Organisation of *Bordetella pertussis* adenylate cyclase and principle of detecting protein-protein interactions using BACTH. The structure of adenylate cyclase domain is shown with T25 (1-224 residues) and T18 (225-399 residues) fragments colored in red and blue, respectively. **A** Expression of chimeric adenylate cyclase alone in *E. coli cya⁻* (BTH101) strain leads to residual cAMP synthesis. **B** Expression of individual T25 and T18 fragments results in lack of cAMP production. **C** If the hybrid proteins carrying T25 and T18 fragments interact, the adenylate cyclase domains are brought together and, thus, cAMP synthesis is restored. **D** BACTH method works for membrane or membrane associated proteins, providing that T25 and T18 domains are facing the cytoplasm. The cAMP binds catabolite activator protein (CAP) and the complex regulates the expression of various genes, including lactose and maltose operons, which result in *cya⁺* phenotype. Indicator plates (MacConkey/maltose, LB/X-gal, and M63/maltose/X-gal) are used to identify positive interactions. Further, quantitative β -galactosidase assay is used to characterise the extent of protein-protein interaction. Adapted by permission from Elsevier: *Methods*, Battesti and Bouveret (2012).

2.2 Materials and methods

2.2.1 Construction of vectors overexpressing GlfT2 and AftA

The pET28b-*MSMEG_6403* expression vector was kindly donated by Dr. Talat Jabeen. Briefly, the coding region for GlfT2, annotated as *MSMEG_6403*, was amplified by PCR from *M. smegmatis* mc²155 genome using the primer pair: 5'-GA TCG ATC CATATG AGT GAC ATC CCT TCC GGC GCA CTC GAA-3' and 5'-GA TCG ATC CTCGAG TCA TCG TCC GAC TTT CTC CGG TGT CTC-3'. The PCR product was restricted with NdeI and XhoI enzymes and ligated into pET28b digested with identical enzymes, thus yielding pET28b-*MSMEG_6403*. The pET28b vector carries an N-terminal His₆-tag sequence. The full length *aftA* (*NCgl0185*) was amplified from *C. glutamicum* ATCC 13032 genomic DNA by PCR using forward 5'-CA TGC ATG CATATG ATG ATT AAC ACC TCT GAA G-3' and reverse 5'-CA TGC ATG GCGGCCGC CTC ATT GTG CGT TAC CAC CA-3' primers, where underlined nucleotides are NdeI and NotI restrictions sites, respectively. The purified DNA fragment was cloned into NdeI and NotI cutting sites of pMSX to express recombinant AftA tagged with His₆ at its C-terminus. DNA sequencing and construct verification was carried out at the Eurofins DNA Ltd.

2.2.2 Pull-down detection of putative GlfT2 and AftA binding proteins

In vitro binding assays were carried out by immobilising 'bait' proteins, pure His₆-tagged GlfT2 or AftA, to Ni²⁺-NTA charged resin (Qiagen). The unbound protein was washed off with 20 ml purification buffer 50 mM KH₂PO₄ pH 7.9, 300 mM NaCl, 20 % glycerol and 50 mM KH₂PO₄ pH 7.9, 100 mM NaCl, 10% glycerol for GlfT2 and AftA, respectively. Clarified cell lysate of *M. smegmatis* and *C. glutamicum* containing

endogenous 'prey' proteins were incubated with immobilised GltT2 and AftA proteins, respectively, and left to interact for 1 hour with gentle shaking at 4 °C. Two washing steps were performed with 20 ml purification buffer containing 50 mM imidazole to remove any unbound or weakly bound proteins. The possible protein complexes of GltT2 and AftA were eluted with 100-500 mM gradient of imidazole. Fractions were analysed by 4-15 % SDS-PAGE and Western blot using monoclonal anti-His₆ antibody.

2.2.3 Mass spectrometry and analysis of putative GltT2 and AftA binding proteins

Proteins were visualised on the SDS-PAGE gels by Coomassie blue staining as described in the General Materials and Methods 6.6.1. The intense bands or regions of interest were excised from the gels, reduced with 10 mM dithiothreitol (DTT), and subsequently alkylated with 75 mM iodoacetamide. Trypsin digestion of in-gel proteins was performed and extracted peptides were analysed by LC-MS/MS (Orbitrap Velos). MASCOT (Matrix Science Ltd) protein search software together with NCBI database was employed to identify peptide sequences.

2.2.4 Plasmid construction for BACTH analysis

All recombinant DNA methods were performed using standard protocols. Briefly, the genes involved in *C. glutamicum* AG biosynthesis were amplified from genomic DNA of *C. glutamicum* ATCC 13032. The plasmids have been constructed by inserting gene sequences of interest in pKT25 (T25 fusion at N-terminus of the gene), pKNT25 (T25 fusion at C-terminus of the gene), pUT18 (T18 fusion at C-terminus of the gene) and pUT18c (T18 fusion at N-terminus of the gene) (Karimova *et al.*, 1998), using oligonucleotides provided in Table 1. Constructs were confirmed using DNA sequencing by Eurofins MWG Ltd. employing the standard M13 reverse primers (29mer) and gene

specific internal primers. The bacterial BACTH system kit was obtained from Euromedex and contained empty vectors together with positive control plasmids pKT25-*zip* and pUT18c-*zip*. Expressed T25-Zip and T18-Zip fusion proteins contain the leucine zipper region of yeast transcriptional activator GCN4. Dimerisation of leucine zipper motif results in reconstitution of adenylate cyclase domain and, sequentially, restoration of cAMP production.

Table 1 Oligonucleotides used in this study.

Plasmid ^a	Primer sequence (5' – 3') ^b
pKT25- <i>wecA</i>	5'-CATGCATGTCTAGAGATGGGAGTCGGTTTCGCG-3' 5'-CATGCATGGAATTCCTTAATCAAGTTTGCGGCT-3'
pKNT25- <i>wecA</i>	5'-CATGCATGAAGCTTGATGGGAGTCGGTTTCGCG-3' 5'-CATGCATGGAATTCGAATCAAGTTTGCGGCTC-3'
pUT18- <i>wecA</i>	5'-CATGCATGAAGCTTGATGGGAGTCGGTTTCGCG-3' 5'-CATGCATGGAATTCGAATCAAGTTTGCGGCTC-3'
pUT18c- <i>wecA</i>	5'-CATGCATGTCTAGAGATGGGAGTCGGTTTCGCG-3' 5'-CATGCATGGAATTCCTTAATCAAGTTTGCGGCT-3'
pKT25- <i>wbbL</i>	5'-CATGCATGCTGCAGGGGTGATCACAGTGACCTA-3' 5'-CATGCATGGAATTCCTAAGAGGCTTTCGTTCTC-3'
pKNT25- <i>wbbL</i>	5'-CATGCATGCTGCAGGGGTGATCACAGTGACCTAT-3' 5'-CATGCATGGAATTCGAAGAGGCTTTCGTTCTCA-3'
pUT18- <i>wbbL</i>	5'-CATGCATGCTGCAGGGGTGATCACAGTGACCTAT-3' 5'-CATGCATGGAATTCGAAGAGGCTTTCGTTCTCA-3'
pUT18c- <i>wbbL</i>	5'-CATGCATGCTGCAGGGGTGATCACAGTGACCTAT-3' 5'-CATGCATGGAATTCCTAAGAGGCTTTCGTTCTC-3'
pKT25- <i>glfT1</i>	5'-CATGCATGCTGCAGGGATGGCACAAACCACTAC-3' 5'-CATGCATGGGTACCCTAGGGCCTATTGAATTTCT-3'
pKNT25- <i>glfT1</i>	5'-CATGCATGCTGCAGGATGGCACAAACCACTACC-3' 5'-CATGCATGGGTACCCGGGGCCTATTGAATTTCT-3'
pUT18- <i>glfT1</i>	5'-CATGCATGCTGCAGGATGGCACAAACCACTACC-3' 5'-CATGCATGGGTACCCGGGGCCTATTGAATTTCT-3'
pUT18c- <i>glfT1</i>	5'-CATGCATGCTGCAGGATGGCACAAACCACTACC-3' 5'-CATGCATGGGTACCCTAGGGCCTATTGAATTTCT-3'
pKT25- <i>glfT2</i>	5'-CATGCATGCTGCAGGGATGAAGGGTGAAGATAC-3' 5'-CATGCATGGGTACCTTATTGCTCATCGAAGACC-3'
pKNT25- <i>glfT2</i>	5'-CATGCATGCTGCAGGATGAAGGGTGAAGATACG-3' 5'-CATGCATGGGTACCCCTTGCTCATCGAAGACCT-3'
pUT18- <i>glfT2</i>	5'-CATGCATGCTGCAGGATGAAGGGTGAAGATACG-3' 5'-CATGCATGGGTACCCCTTGCTCATCGAAGACCT-3'
pUT18c- <i>glfT2</i>	5'-CATGCATGCTGCAGGATGAAGGGTGAAGATACG-3' 5'-CATGCATGCTGCAGGGATGAAGGGTGAAGATAC-3'

pKT25- <i>aftA</i>	5'-CATGCATGCTGCAGGGATGATTAACACCTCTGA-3' 5'-CATGCATGGGTACCTTACTCATTGTGCGTTACC-3'
pKNT25- <i>aftA</i>	5'-CATGCATGCTGCAGGATGATTAACACCTCTGAA-3' 5'-CATGCATGGGTACCCCTCATTGTGCGTTACCA-3'
pUT18- <i>aftA</i>	5'-CATGCATGCTGCAGGATGATTAACACCTCTGAA-3' 5'-CATGCATGGGTACCCCTCATTGTGCGTTACCA-3'
pUT18c- <i>aftA</i>	5'-CATGCATGCTGCAGGATGATTAACACCTCTGAA-3' 5'-CATGCATGCTGCAGGGATGATTAACACCTCTGA-3'
pKT25- <i>aftB</i>	5'-CATGCATGTCTAGAGATGACGTTTAGCCCCCAG-3' 5'-CATGCATGGAATTCTTACTGAGAGCTATATAAA-3'
pKNT25- <i>aftB</i>	5'-CATGCATGTCTAGAGATGACGTTTAGCCCCCAG-3' 5'-CATGCATGGAATTTCGACTGAGAGCTATATAAAG-3'
pUT18- <i>aftB</i>	5'-CATGCATGTCTAGAGATGACGTTTAGCCCCCAG-3' 5'-CATGCATGGAATTTCGACTGAGAGCTATATAAAG-3'
pUT18c- <i>aftB</i>	5'-CATGCATGTCTAGAGATGACGTTTAGCCCCCAG-3' 5'-CATGCATGGAATTCTTACTGAGAGCTATATAAA-3'
pKT25- <i>aftC</i>	5'-CATGCATGCTGCAGGGATGTTGTTGATGGCGCA-3' 5'-CATGCATGGGTACCTCATGCTGTCCTCTCAAGA-3'
pKNT25- <i>aftC</i>	5'-CATGCATGCTGCAGGATGTTGTTGATGGCGCAT-3' 5'-CATGCATGGGTACCCGTGCTGTCCTCTCAAGAT-3'
pUT18- <i>aftC</i>	5'-CATGCATGCTGCAGGATGTTGTTGATGGCGCAT-3' 5'-CATGCATGGGTACCCGTGCTGTCCTCTCAAGAT-3'
pUT18c- <i>aftC</i>	5'-CATGCATGCTGCAGGATGTTGTTGATGGCGCAT-3' 5'-CATGCATGGGTACCTCATGCTGTCCTCTCAAGA-3'
pKT25- <i>aftD</i>	5'-CATGCATGCTGCAGGGGTGCTGGGTTTTGTGGT-3' 5'-CATGCATGCCCGGGTTAGCGCTTTGGAGGCCTT-3'
pKNT25- <i>aftD</i>	5'-CATGCATGCTGCAGGGGTGCTGGGTTTTGTGGTG-3' 5'-CATGCATGCCCGGGGGCGCTTTGGAGGCCTTAA-3'
pUT18- <i>aftD</i>	5'-CATGCATGCTGCAGGGGTGCTGGGTTTTGTGGTG-3' 5'-CATGCATGCCCGGGGGCGCTTTGGAGGCCTTAA-3'
pUT18c- <i>aftD</i>	5'-CATGCATGCTGCAGGGGTGCTGGGTTTTGTGGTG-3' 5'-CATGCATGCTGCAGGGGTGCTGGGTTTTGTGGT-3'
pKT25- <i>ubiA</i>	5'-CATGCATGTCTAGAGGTGAGCGAACACGCCGCT-3' 5'-CATGCATGGAATTCTCAAACATCGGCATGATG-3'
pKNT25- <i>ubiA</i>	5'-CATGCATGAAGCTTGGTGAGCGAACACGCCGCT-3' 5'-CATGCATGGAATTTCGAAAACATCGGCATGATGT-3'
pUT18- <i>ubiA</i>	5'-CATGCATGAAGCTTGGTGAGCGAACACGCCGCT-3' 5'-CATGCATGGAATTTCGAAAACATCGGCATGATGT-3'
pUT18c- <i>ubiA</i>	5'-CATGCATGTCTAGAGGTGAGCGAACACGCCGCT-3' 5'-CATGCATGGAATTCTCAAACATCGGCATGATG-3'
pKT25- <i>dprE1</i>	5'-CATGCATGTCTAGAGATGAACAGTTCTCACGGC-3' 5'-CATGCATGGGTACCTTAAGAAAGCTCAAGTCG-3'
pKNT25- <i>dprE1</i>	5'-CATGCATGGCATGCCATGAACAGTTCTCACGGC-3' 5'-CATGCATGGGTACCCGAGAAAGCTCAAGTCGGC-3'
pUT18- <i>dprE1</i>	5'-CATGCATGGCATGCCATGAACAGTTCTCACGGC-3' 5'-CATGCATGGGTACCCGAGAAAGCTCAAGTCGGC-3'
pUT18c- <i>dprE1</i>	5'-CATGCATGTCTAGAGATGAACAGTTCTCACGGC-3' 5'-CATGCATGGGTACCTTAAGAAAGCTCAAGTCG-3'
pKT25- <i>dprE2</i>	5'-CATGCATGTCTAGAGATGCTTAACGCAGTGGGC-3'

	5'-CATGCATGGAATTCTTAGAACGGCAGCTTGCGG-3'
pKNT25- <i>dprE2</i>	5'-CATGCATGTCTAGAGATGCTTAACGCAGTGGGC-3'
	5'-CATGCATGGAATTCGAGAACGGCAGCTTGCGGA-3'
pUT18- <i>dprE2</i>	5'-CATGCATGTCTAGAGATGCTTAACGCAGTGGGC-3'
	5'-CATGCATGGAATTCGAGAACGGCAGCTTGCGGA-3'
pUT18c- <i>dprE2</i>	5'-CATGCATGTCTAGAGATGCTTAACGCAGTGGGC-3'
	5'-CATGCATGGAATTCTTAGAACGGCAGCTTGCGG-3'
pKT25- <i>emb</i>	5'-CATGCATGTCTAGAGATGCGCCAAGTCGGTGGT-3'
	5'-CATGCATGGAATTCTTATTCATCTACCTTCATA-3'
pKNT25- <i>emb</i>	5'-CATGCATGTCTAGAGATGCGCCAAGTCGGTGGT-3'
	5'-CATGCATGGAATTCGATTCATCTACCTTCATAT-3'
pUT18- <i>emb</i>	5'-CATGCATGTCTAGAGATGCGCCAAGTCGGTGGT-3'
	5'-CATGCATGGAATTCGATTCATCTACCTTCATAT-3'
pUT18c- <i>emb</i>	5'-CATGCATGTCTAGAGATGCGCCAAGTCGGTGGT-3'
	5'-CATGCATGGAATTCTTATTCATCTACCTTCATA-3'

^aPlasmids contain the sequence generated by PCR amplification with the pair of primer on the right column

^bRestriction enzyme sites are underlined

2.2.5 Bacterial two-hybrid system

E. coli BTH101 cells were co-transformed with two plasmids (Table 2) expressing recombinant proteins bearing N- or C- terminal T25 and T18 fusions as described in the General Materials and Methods 6.5.10. Cells were spread on LB plates containing streptomycin (100 µg/ml), ampicillin (100 µg/ml), kanamycin (50 µg/ml) and incubated at 30 °C for 48 hours. Several colonies were picked and used to inoculate 5 ml of LB supplemented with appropriate antibiotics and 0.5 mM IPTG to induce protein expression. Cultures were grown overnight at 30 °C with shaking. Samples were then washed three times in minimal M63 media and spotted (2 µl) onto LB, MacConkey or M63 minimal media agar plates supplemented with appropriate antibiotics and nutrients.

Table 2 Bacterial strains and plasmids used in this study.

Plasmid or strain	Description or genotype
<i>E. coli</i> XL-1 Blue	Cloning strain
<i>E. coli</i> BTH101	F ⁻ , <i>cya</i> -99, <i>araD</i> 139, <i>galE</i> 15, <i>galK</i> 16, <i>rpsL</i> 1 (Str ^r), <i>hsdR</i> 2, <i>mcrA</i> 1, <i>mcrB</i> 1 strain
pKT25	Cloning and expression vector, pSU40 derivative with T25 domain of CyaA, multiclonal sequence site (MCS) at the 3' end of T25, Kan ^r
pKNT25	Cloning and expression vector, pSU40 derivative with T25 domain of CyaA, MCS at the 3' start of T25, Kan ^r
pUT18	Cloning and expression vector, pUC19 derivative with T18 domain of CyaA, MCS at the 3' start of T18, Amp ^r
pUT18c	Cloning and expression vector, pUC19 derivative with T18 domain of Cya, MCS at the 3' end of T18, Amp ^r
pKT25-zip	Control plasmid, T25 domain of Cya fused in frame with leucine zipper of GCN4, Kan ^r
pUT18c-zip	Control plasmid, T18 domain of Cya fused in frame with leucine zipper of GCN4, Amp ^r
pKT25- <i>wecA</i>	pKT25 plasmid with <i>cyaAT25-wecA</i> fusion, Kan ^r
pKNT25- <i>wecA</i>	pKNT25 plasmid with <i>wecA-cyaAT25</i> fusion, Kan ^r
pUT18- <i>wecA</i>	pUT18 plasmid with <i>wecA-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>wecA</i>	pUT18c plasmid with <i>cyaAT18-wecA</i> fusion, Amp ^r
pKT25- <i>wbbL</i>	pKT25 plasmid with <i>cyaAT25-wbbL</i> fusion, Kan ^r
pKNT25- <i>wbbL</i>	pKNT25 plasmid with <i>wbbL-cyaAT25</i> fusion, Kan ^r
pUT18- <i>wbbL</i>	pUT18 plasmid with <i>wbbL-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>wbbL</i>	pUT18c plasmid with <i>cyaAT18-wbbL</i> fusion, Amp ^r
pKT25- <i>glfT1</i>	pKT25 plasmid with <i>cyaAT25-glfT1</i> fusion, Kan ^r
pKNT25- <i>glfT1</i>	pKNT25 plasmid with <i>glfT1-cyaAT25</i> fusion, Kan ^r
pUT18- <i>glfT1</i>	pUT18 plasmid with <i>glfT1-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>glfT1</i>	pUT18c plasmid with <i>cyaAT18-glfT1</i> fusion, Amp ^r
pKT25- <i>glfT2</i>	pKT25 plasmid with <i>cyaAT25-glfT2</i> fusion, Kan ^r
pKNT25- <i>glfT2</i>	pKNT25 plasmid with <i>glfT2-cyaAT25</i> fusion, Kan ^r
pUT18- <i>glfT2</i>	pUT18 plasmid with <i>glfT2-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>glfT2</i>	pUT18c plasmid with <i>cyaAT18-glfT2</i> fusion, Amp ^r
pKT25- <i>aftA</i>	pKT25 plasmid with <i>cyaAT25-aftA</i> fusion, Kan ^r
pKNT25- <i>aftA</i>	pKNT25 plasmid with <i>aftA-cyaAT25</i> fusion, Kan ^r
pUT18- <i>aftA</i>	pUT18 plasmid with <i>aftA-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>aftA</i>	pUT18c plasmid with <i>cyaAT18-aftA</i> fusion, Amp ^r
pKT25- <i>aftB</i>	pKT25 plasmid with <i>cyaAT25-aftB</i> fusion, Kan ^r
pKNT25- <i>aftB</i>	pKNT25 plasmid with <i>aftB-cyaAT25</i> fusion, Kan ^r
pUT18- <i>aftB</i>	pUT18 plasmid with <i>aftB-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>aftB</i>	pUT18c plasmid with <i>cyaAT18-aftB</i> fusion, Amp ^r
pKT25- <i>aftC</i>	pKT25 plasmid with <i>cyaAT25-aftC</i> fusion, Kan ^r
pKNT25- <i>aftC</i>	pKNT25 plasmid with <i>aftC-cyaAT25</i> fusion, Kan ^r
pUT18- <i>aftC</i>	pUT18 plasmid with <i>aftC-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>aftC</i>	pUT18c plasmid with <i>cyaAT18-aftC</i> fusion, Amp ^r
pKT25- <i>aftD</i>	pKT25 plasmid with <i>cyaAT25-aftD</i> fusion, Kan ^r

pKNT25- <i>aftD</i>	pKNT25 plasmid with <i>aftD-cyaAT25</i> fusion, Kan ^r
pUT18- <i>aftD</i>	pUT18 plasmid with <i>aftD-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>aftD</i>	pUT18c plasmid with <i>cyaAT18-aftD</i> fusion, Amp ^r
pKT25- <i>ubiA</i>	pKT25 plasmid with <i>cyaAT25-ubiA</i> fusion, Kan ^r
pKNT25- <i>ubiA</i>	pKNT25 plasmid with <i>ubiA-cyaAT25</i> fusion, Kan ^r
pUT18- <i>ubiA</i>	pUT18 plasmid with <i>ubiA-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>ubiA</i>	pUT18c plasmid with <i>cyaAT18-ubiA</i> fusion, Amp ^r
pKT25- <i>dprE1</i>	pKT25 plasmid with <i>cyaAT25-dprE1</i> fusion, Kan ^r
pKNT25- <i>dprE1</i>	pKNT25 plasmid with <i>dprE1-cyaAT25</i> fusion, Kan ^r
pUT18- <i>dprE1</i>	pUT18 plasmid with <i>dprE1-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>dprE1</i>	pUT18c plasmid with <i>cyaAT18-dprE1</i> fusion, Amp ^r
pKT25- <i>dprE2</i>	pKT25 plasmid with <i>cyaAT25-dprE2</i> fusion, Kan ^r
pKNT25- <i>dprE2</i>	pKNT25 plasmid with <i>dprE2-cyaAT25</i> fusion, Kan ^r
pUT18- <i>dprE2</i>	pUT18 plasmid with <i>dprE2-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>dprE2</i>	pUT18c plasmid with <i>cyaAT18-dprE2</i> fusion, Amp ^r
pKT25- <i>emb</i>	pKT25 plasmid with <i>cyaAT25-emb</i> fusion, Kan ^r
pKNT25- <i>emb</i>	pKNT25 plasmid with <i>emb-cyaAT25</i> fusion, Kan ^r
pUT18- <i>emb</i>	pUT18 plasmid with <i>emb-cyaAT18</i> fusion, Amp ^r
pUT18c- <i>emb</i>	pUT18c plasmid with <i>cyaAT18-emb</i> fusion, Amp ^r

2.2.6 β -Galactosidase assay

Several colonies were picked from co-transformation plates and used to inoculate 5 ml LB supplemented with appropriate antibiotics and 0.5 mM IPTG to induce protein expression. Cultures were grown overnight at 30 °C with shaking. The liquid cultures were diluted 1:5 with M63 minimal media and the OD₆₀₀ of each culture recorded. Bacterial cells were then permeabilised with 30 μ l toluene and 30 μ l 0.1 % SDS solution per 2.5 ml of diluted cell culture. The samples were vortexed for 30 seconds, lightly plugged with cotton wool and incubated at 37 °C with shaking for 45 minutes. Permeabilised cells (0.5 ml) were added to 0.5 ml of PM2 assay buffer (70 mM Na₂HPO₄·12H₂O, 30 mM NaH₂PO₄·H₂O, 1 mM MgSO₄, 0.2 mM MnSO₄ pH 7.0, 100 mM β -mercaptoethanol) and incubated at 28 °C for 5 minutes. A control tube containing 1 ml of PM2 assay buffer was also prepared to serve as a blank. Enzymatic reaction was

started with 0.25 ml 0.4 % o-nitrophenol- β -galactosidase (ONPG) and samples were incubated at 28 °C until the sufficient yellow color has developed. The reaction was stopped with 0.5 ml 1M Na₂CO₃. The OD₄₂₀ and OD₅₅₀ were then recorded for each sample. The β -galactosidase activity was calculated using the formula:

$$1 \text{ Miller unit} = 1000 * \frac{(\text{OD}_{420} - (1.75 * \text{OD}_{550}))}{(t * v * \text{OD}_{600})}$$

The values presented are the mean of 3 independent activity assays.

2.2.7 Statistical analysis

The results are expressed as the means \pm S.D. and were analyzed using a Student's t-test to determine significant differences ($p < 0.01$) between samples.

2.3 Results

2.3.1 Network analysis of AG biosynthetic proteins

We initially aimed to identify whether any of the proteins involved in AG biosynthesis have been predicted or demonstrated to interact as part of a multi-protein functional network. Focusing on the list of proteins associated with AG biosynthesis we used the STRING database of interactions (Szklarczyk *et al.*, 2011) to reveal a putative protein association network with GlfT2 chosen as the network node (Figure 3). The interaction patterns of proteins had a high confidence score (>0.7) and served as a basis for selection of *C. glutamicum* proteins that were further analysed using either pull-down assays or BACTH. The generated network contained ABC family transporters (RfbD and RfbE), GT-A type GalT's GlfT1, and proteins involved in rhamnose sugar donor formation, all

centered on GlfT2. Transmembrane Ara/T Emb showed strong evidence for interaction with AftA and AftB, as well as the uncharacterised protein NCgl2596. The network also contained a putative phospholipid phosphatase NCgl2782 and proteins involved in DPA synthesis: DprE1, DprE2 and UbiA.

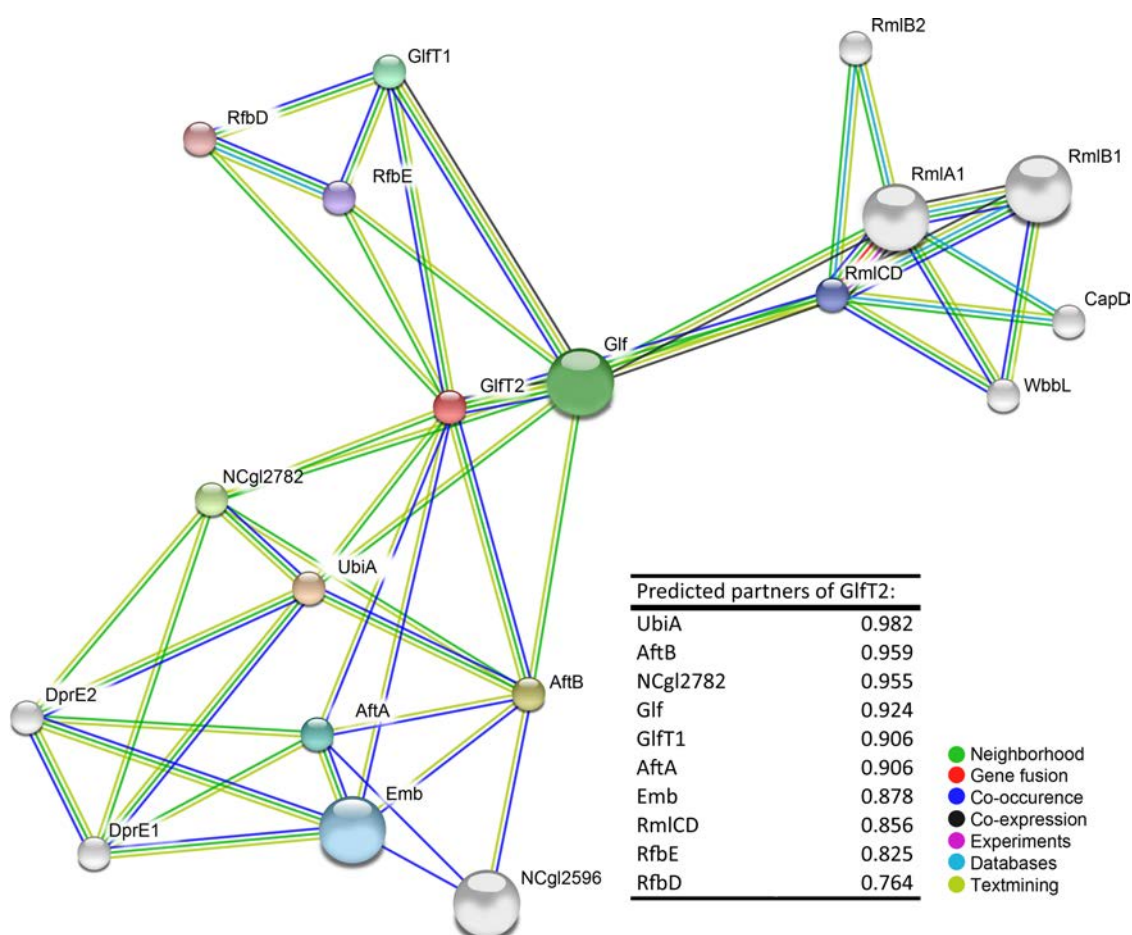


Figure 3 Network of *C. glutamicum* proteins found to be important for cell wall assembly as determined by STRING 9.1 analysis. Lines connecting the nodes indicate various interaction data supporting the network, coloured by evidence type. Combining the probabilities from these different evidence channels, and then correcting for the probability of randomly observing an interaction STRING 9.1 database generates the confidence score. Gene number encoding respective proteins were used in the input and output, which were subsequently. Protein number encoding GlfT2 was used in the input resulting in the output data to contain the same nomenclature. For clarity, all protein numbers were further replaced by protein name according to the literature.

2.3.2 Pull-down assay and detection of putative GlfT2 binding proteins from *M.*

smegmatis

A pull-down assay employing His₆-tagged GlfT2 was performed in order to identify putative protein-protein interactions. Unfortunately, attempts to obtain highly purified *C. glutamicum* GlfT2 in soluble form were unsuccessful. Therefore, we employed standard pET expression system in *E. coli* BL21 (DE3) cells to over-express recombinant full-length GlfT2 of *M. smegmatis* (Figure 4A). *E. coli* cells harbouring pET28b were grown to an early exponential phase and expression of *MSMEG_6403* was induced with 1 mM IPTG at 16 °C. His₆-tagged GlfT2 was purified by affinity chromatography and SDS-PAGE analysis revealed an appreciable amount of protein at 200 mM, 300 mM, and 500 mM imidazole fractions (Figure 4B). *M. smegmatis* GlfT2 migrated at the expected molecular weight – 72 kDa – and was estimated to be at least >95% pure (Figure 4C).

In the pull-down assay, purified GlfT2 was bound to the nickel affinity chromatography column followed by incubation with *M. smegmatis* clarified cell lysate. The column was subsequently washed with 50 mM imidazole to remove any non-bound or weakly bound components of the cell lysate. Proteins possibly interacting with GlfT2 were eluted along with GlfT2 using 250 mM imidazole and analysed by SDS-PAGE and Western blot (Figure 4D). To identify protein compositions in the SDS-PAGE gel, eight bands with high staining intensity were excised from the gel, digested with trypsin, and acquired tryptic peptides were analysed by LC-MS/MS. A parallel assay was performed lacking the GlfT2 protein and, hence, served as a negative control (Figure 4D). SDS-PAGE analysis demonstrated the absence of protein bands, thus reaffirming that eluted proteins in the pull-down assay are indeed associated with GlfT2.

In total, 33 distinct proteins were identified using Mascot algorithm against NCBI database (Table 3). Most of the detected proteins are involved in genetic information processing (27%). Additionally, six proteins are required for virulence, detoxification, and adaptation (18%), while the rest are involved in lipid metabolism (18%). Importantly, the essential GroEL2 molecular chaperone was identified in five out of nine analysed bands strongly suggesting an association with GlfT2. In mycobacteria, GroEL2 is proposed to provide a housekeeping chaperone function, where it promotes refolding and proper assembly of unfolded polypeptides and proteins (Stapleton *et al.*, 2012). Thus, suggesting that over-expressed GlfT2 requires GroEL2 for appropriate folding. Moreover, Alderwick *et al.* (2008) have demonstrated that GlfT2 from *M. tuberculosis* requires chaperones for overexpression in *E. coli*. In addition, polyketide synthase Pks13 and condensing enzyme KasA, both involved in mycolic acid biosynthesis, were also detected. Interactions between GlfT2 and enzymes involved in the assembly of mycolic acids are not surprising, since formation of large multi-protein complexes may enhance efficiency and fidelity of mAGP formation. Nevertheless, none of the identified proteins were related to AG synthesis and, hence, were not studied further.

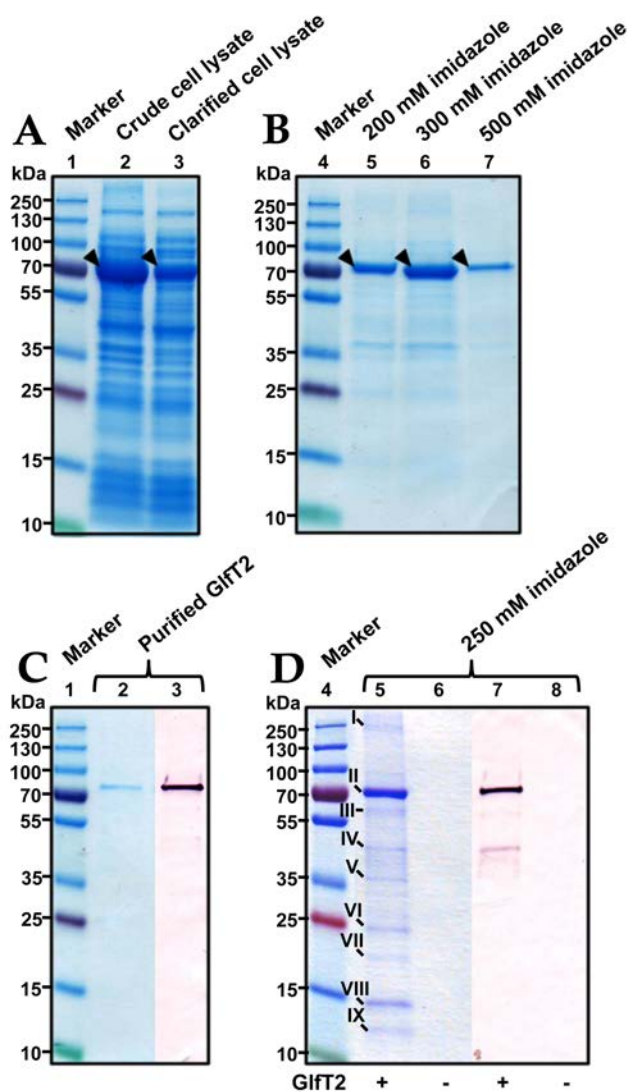


Figure 4 SDS-PAGE and Western blot analysis of *M. smegmatis* GlfT2 over-expression, purification, and pull-down assay. **A** *E. coli* BL21 (DE3) cells were grown at 16 °C overnight and 1 mM IPTG was used to induce expression of *MSMEG_6403*. Lanes 2 and 3 correspond to crude and clarified cell lysates, respectively. **B** Recombinant GlfT2 with a polyhistidine tag was purified by nickel affinity chromatography column. Lanes 5, 6 and 7 correspond to elution fractions at 200 mM, 300 mM and 500 mM of imidazole, respectively. The arrows indicate the over-expressed GlfT2 protein. **C** Lanes 2 (SDS-PAGE) and 3 (Western blot) correspond to purified GlfT2 fraction. **D** Pull-down assay where purified recombinant GlfT2 was used as a ‘bait’ and cell lysate of *M. smegmatis* served as a pool of ‘prey’ proteins. Lanes 5 (SDS-PAGE) and 7 (Western blot) represent the 250 mM imidazole elution fraction containing recombinant GlfT2 and co-eluted proteins. Lanes 6 (SDS-PAGE) and 8 (Western blot) represent the 250 mM imidazole fraction, where GlfT2 protein was excluded from the affinity column and, therefore, served as a negative control. Samples were analysed by 4-15 % SDS-PAGE gel and stained with Coomassie Blue; the His-tagged GlfT2 was detected by Western blot using a monoclonal anti-His₆ tag antibody.

Table 3 List of proteins that potentially form multi-protein complexes with GlfT2 from *M. smegmatis* identified in this study

Band ^a	Protein	Description	Score ^b	Coverage ^c	Peptides ^d	MW (kDa) ^e
I	GlfT2	galactofuranosyltransferase	252.31	61.92	46	71.53
	Pks13	polyketide synthase	28.23	17.90	27	194.3
	GroEL2	60kDa chaperonin 2	22.35	26.67	10	56.12
II	GlfT2	galactofuranosyltransferase	954.97	80.19	56	71.53
	DnaK	molecular chaperone protein	37.53	36.82	15	66.61
	GroEL2	60kDa chaperonin 2	9.22	33.15	12	56.12
	MSMEG4474	acyl-CoA oxidase	6.60	8.28	5	70.45
III	GlfT2	galactofuranosyltransferase	221.14	65.94	43	71.53
	GroEL2	60kDa chaperonin 2	108.28	51.85	28	56.12
	RpsA	30S ribosomal protein S1	32.44	41.34	17	53.28
	SppA	signal peptide peptidase, 67K type	30.19	24.24	13	62.68
	GroEL1	60kDa chaperonin 1	24.44	26.43	12	56.45
	FadD15	AMP-binding enzyme	14.60	13.69	7	64.45
	AccD4	propionyl-CoA carboxylase beta chain	10.64	14.12	5	56.13
	SdhA	succinate dehydrogenase flavoprotein	6.06	8.73	5	64.35
IV	GlfT2	galactofuranosyltransferase	203.09	61.76	39	71.53
	FabG4	3-ketoacyl-(acyl-carrier-protein) reductase	101.31	65.78	24	46.47
	FadA2	acetyl-CoA acetyltransferase	60.33	46.06	20	45.21
	KasA	3-oxoacyl-(acyl carrier protein) synthase	48.24	48.56	13	43.76
	RpoA	DNA-directed RNA polymerase subunit	16.73	23.43	7	37.90
	MSMEG4632	saccharopine dehydrogenase	12.09	33.41	10	43.61
	Tuf	translation elongation factor	11.76	24.75	8	43.71
	DnaJ	chaperone protein	10.69	23.36	5	40.31
	AtpFH	F0F1 ATP synthase subunit delta	8.98	22.02	8	47.42
	ClpX	ATP-dependent protease ATP-binding unit	8.98	26.06	8	46.65
	CysS	cysteinyl-tRNA synthetase	7.63	9.22	10	45.37
	MSMEG2070	acyl-CoA dehydrogenase family protein	6.87	13.77	4	47.26
	MSMEG6284	cyclopropane-fatty-acyl- synthase	6.53	22.65	5	48.61
GroEL2	60kDa chaperonin 2	6.10	21.85	7	56.12	
V	GlfT2	galactofuranosyltransferase	82.97	50.46	30	71.53
	MSMEG5117	proline dehydrogenase	14.11	31.23	9	35.40
	GroEL2	60kDa chaperonin 2	8.32	7.41	4	56.12
	NudC	NADH pyrophosphatase	7.55	16.40	5	34.12
VI	GlfT2	galactofuranosyltransferase	31.88	40.87	21	71.53
VII	GlfT2	galactofuranosyltransferase	26.03	40.87	20	71.53
VIII	GlfT2	galactofuranosyltransferase	53.91	34.21	24	71.53
	RpsG	30S ribosomal protein S7	23.27	68.59	11	17.62
	RplM	50S ribosomal protein L13	18.27	58.50	10	16.11
	RplL	50S ribosomal protein L7/L12	17.52	46.92	6	13.45
	RpsK	30S ribosomal protein S11	16.47	40.58	4	14.63
	MSMEG4692	conserved hypothetical protein	10.58	57.96	5	15.89
	HspX	14 kDa antigen	9.07	47.59	7	15.92
	RplP	50S ribosomal protein L16	8.83	33.33	4	15.73
	RplT	50S ribosomal protein L20	7.46	27.13	3	14.46
	MSMEG1077	conserved hypothetical protein	5.26	59.06	7	16.28
IX	GlfT2	galactofuranosyltransferase	20.38	22.60	12	71.53

a) number of excised SDS-PAGE gel band for protein identification; **b)** the Mascot score of protein identification; **c)** sequence coverage in % of the identified protein; **d)** number of identified unique peptides obtained for a particular protein; **e)** molecular weight in kDa of the respective protein. Yellow background indicates proteins of interest that were identified in the assay.

2.3.3 Pull-down assay and detection of putative AftA binding proteins from *C.*

glutamicum

Recombinant full-length AftA from *C. glutamicum* was overexpressed in *C. glutamicum* ATCC13303. Cells previously transformed with pMSX-NCgl0185 were grown to an early exponential phase and expression of AftA with C-terminus His₆ tag was induced using 1 mM IPTG at 16 °C. Notably, a high molecular weight (150 kDa and above) bands were observed in Western blots of samples overexpressing *C. glutamicum* AftA, which may correspond to aggregated material (Figure 5A). However, it is tempting to speculate that this may be the evidence of multi-protein complex formation, where *C. glutamicum* AftA is forming a dimer or tetramer. AftA was purified by immobilised metal ion exchange chromatography and eluted fractions were analysed by SDS-PAGE. Protein bands of weak intensity were obtained at the expected size of 74.1 kDa at 100-200 mM imidazole fractions (Figure 5B). The purity of *C. glutamicum* AftA preparation was estimated to be greater than 95% (Figure 5C).

A pull-down assay was performed employing purified AftA in order to identify its interacting partners. Immobilised AftA was incubated with *C. glutamicum* clarified lysate, washed with 50 mM imidazole, and ‘prey’ proteins bound to AftA were eluted with 500 mM imidazole (Figure 5D). Samples were subjected to SDS-PAGE analysis, major bands of possible protein complexes excised from the gel (I-VIII, Figure 5D), digested with trypsin, and analysed by LC-MS/MS. The negative pull-down control assay, excluding AftA as the ‘bait’, was performed analogously (Figure 5D). SDS-PAGE analysis demonstrated bands of weak intensity in the elution fraction, possibly represented by proteins that bound to the nickel affinity chromatography column either in a non-specific manner or contain a number of histidine residues. However, due to faint

nature of these bands when compared to the pull-down assay fraction, these were considered insignificant.

Subsequent analysis using Mascot algorithm against NCBI database identified a list of proteins that potentially interact with AftA. The summary containing significant hits (score > 25, coverage > 20) and proteins associated with cell envelope biosynthesis is presented in Table 4. Proteins involved in the assembly of PG, notably MurA, MurC and MurG, were identified. Moreover, GalE epimerase and RfbE ABC transporter that play a role in UDP-Gal f donor formation and transportation of cell wall lipopolysaccharides, respectively, were also observed. Importantly, five proteins involved in AG biogenesis were identified: GlfT2, GalT, AftD, AraG, DprE1 and DprE2, as well as decaprenylphosphate phosphoribosyltransferase UbiA. These latter proteins were selected for further studies using BACTH.

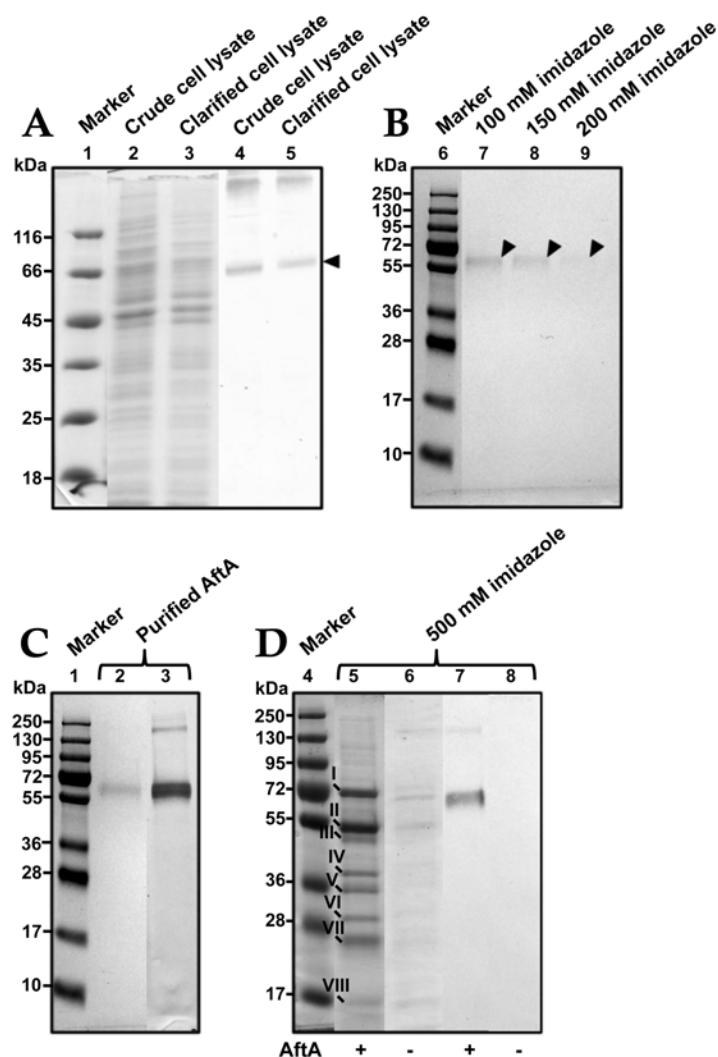


Figure 5 SDS-PAGE and Western blot analysis of *C. glutamicum* AftA over-expression, purification, and a pull-down assay. A *C. glutamicum* cells were grown at 16 °C overnight and 1 mM IPTG was used to induce expression of *NCgl0185*. Lanes 2 (SDS-PAGE) and 4 (Western blot) represent the crude cell lysate of *C. glutamicum::pMSX-NCgl0185*, while lanes 3 (SDS-PAGE) and 5 (Western blot) correspond to clarified cell lysate. B Recombinant AftA with a His₆-tag was purified by affinity chromatography. Lanes 7, 8 and 9 correspond to elution fractions at 100 mM, 150 mM and 200 mM of imidazole, respectively. Note that due to the low amount of protein it is difficult to detect AftA in elution fractions. The arrows indicate the over-expressed AftA protein. C Pull-down assay employing recombinant AftA, where purified AftA was a ‘bait’ and membrane fraction preparation of *C. glutamicum* served as a pool of ‘prey’ proteins. Lanes 2 (SDS-PAGE) and 3 (Western blot) represent purified AftA fraction used in the assay. D Lanes 5 (SDS-PAGE) and 7 (Western blot) represent the 500 mM imidazole fraction containing recombinant AftA and co-eluted proteins. Lanes 6 (SDS-PAGE) and 8 (Western blot) represent the 500 mM imidazole fraction, where AftA protein was excluded from the affinity column and, therefore, served as a negative control. Samples were analysed by 12 % SDS-PAGE gel and stained with Coomassie Blue; the His-tagged AftA was detected by a Western blot using a monoclonal anti-His₆ tag antibody.

Table 4 List of proteins that potentially form multi-protein complexes with AftA from *C. glutamicum* identified in this study.

Band ^a	Protein	Description	Score ^b	Coverage ^c	Peptides ^d	MW (kDa) ^e
I	GlgB	glycogen branching enzyme	288.6	33.65	28	82.54
	NCgl0634	monomeric isocitrate dehydrogenase	104.1	41.73	30	80.03
	NCgl0104	acyl-CoA synthetase	94.75	37.57	26	75.41
	NCgl2859	cation transport ATPase	94.01	32.34	22	77.37
	DnaK	molecular chaperone	91.37	54.05	36	66.25
	NCgl1471	methylmalonyl-CoA mutase	67.93	31.75	25	80.06
	NCgl2765	phosphoenolpyruvate carboxykinase	63.25	33.44	21	66.83
	NCgl1590	guanosine polyphosphate synthetase	56.48	38.16	38	84.39
	NCgl1053	membrane GTPase	55.74	33.44	26	68.66
	GlfT2	galactosyltransferase	52.03	38.57	26	72.93
	AftA	arabinosyltransferase	36.86	19.52	11	72.52
Fas-IB	fatty-acid synthase	30.55	12.39	38	314.9	
II	GorEL	60 kDa molecular chaperone	590.8	49.63	39	56.71
	ArgS	arginyl-tRNA synthetase	125.8	52.36	24	59.69
	NCgl1165	FOF1 ATP synthase subunit beta	74.99	41.61	18	52.51
	NCgl2294	ABC-type transport systems	74.92	28	12	57.67
	AftA	arabinosyltransferase	69.93	21.63	12	72.52
	AtpA	FOF1 ATP synthase subunit alpha	69.01	45.16	27	58.71
	NCgl1233	hypothetical protein	67.22	29.59	23	67.15
	SerA	phosphoglycerate dehydrogenase	63.81	50.38	21	55.27
	GlgB	glycogen branching enzyme	56.81	28.59	18	82.54
	NCgl0578	inositol dehydrogenase	55.19	49.8	27	53.33
	PrpD	propionate catabolism protein	54.66	32.74	19	55.36
GlfT2	galactosyltransferase	19.22	27.13	18	72.93	
III	GroEL	60 kDa molecular chaperone	267.8	48.7	33	56.71
	NCgl0461	putative regulatory protein	132.1	49.6	26	55.71
	AmyE	ABC-type transporter	127.7	36.79	17	49.42
	ActA	Acetyl-CoA hydrolase	99.85	33.53	20	55.91
	MgtE	Mg/Co/Ni transporter	97.17	47.91	22	47.29
	NCgl0622	flotillin-like protein	84.30	56.74	33	49.72
	GlgB	glycogen branching enzyme	65.64	29	20	82.54
	DprE1	decaprenylphosphoryl-beta-D-ribose 2'-oxidase	62.59	32.17	15	53.03
IV	AftA	arabinosyltransferase	52.51	20.12	11	72.52
	MurC	UDP-N-acetylmuramate-L-alanine ligase	36.31	20.58	9	51.08
	GlfT2	galactosyltransferase	14.89	29.73	20	72.93
	NCgl2383	hypothetical protein	233.0	54.06	20	42.33
	NCgl2377	ABC-type transporter	101.9	51.86	23	40.29
	GroEL	60 kDa molecular chaperone	92.49	39.96	22	56.71
	IspG	1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase	90.09	57.41	20	40.04
	NCgl1526	glyceraldehyde-3-phosphate dehydrogenase	73.59	39.22	13	36.02
	PfkA	phosphofructokinase	70.40	46.94	16	37.06
	AftA	hypothetical membrane protein	27.30	15.89	9	72.52
	MurA	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	24.80	39.47	17	44.17

	GlgB	glycogen branching enzyme	24.27	20.66	13	82.54
	LpqI	beta-N-acetylglucosaminidase-like protein	4.76	13.92	5	41.53
	AftD	arabinosyltransferase	2.06	3.97	3	108.1
	PfkA	6-phosphofructokinase	195.8	63.85	27	37.06
	NCgl2383	hypothetical protein	103.1	55.33	19	42.33
	GroEL	60 kDa molecular chaperone	90.57	47.03	28	56.71
	RpsB	30S ribosomal protein S2	62.71	55.51	18	30.10
	AdhA	alcohol dehydrogenase	62.50	27.83	10	36.79
	NCgl0347	glycosyltransferase	55.96	53.91	18	38.81
V	GlgB	glycogen branching enzyme	39.93	25.03	16	82.54
	GalE	UDP-glucose 4-epimerase	36.45	29.48	9	35.24
	MurG	UDP-N-acetylglucosamine:LPS N-acetylglucosamine transferase	29.62	34.68	15	38.44
	AftA	arabinosyltransferase	12.13	9.23	5	72.52
	Glft2	galactosyltransferase	2.31	17.23	12	72.93
	RfbE	ABC-type transporter	1.86	21.67	6	28.66
	PimT	SAM-dependent methyltransferase	218.0	65.47	21	31.14
	GroEL	60 kDa molecular chaperone	109.0	45.35	28	56.71
	NCgl2487	histone acetyltransferase HPA2-like protein	80.00	46.58	17	32.11
	NCgl2156	Zn-ribbon protein	62.37	50.21	16	26.12
	NCgl2383	hypothetical protein	53.34	43.91	13	42.33
	PfkA	6-phosphofructokinase	52.08	51.31	16	37.06
VI	DprE2	decaprenylphosphoryl-D-2-keto erythro pentose reductase	41.83	45.06	12	26.71
	GluB	glutamate ABC-type transporter	34.06	14.92	4	31.65
	NCgl1179	ABC-type transporter	32.39	37.2	10	32.31
	NCgl0353	glycosyltransferase	22.99	46.35	18	31.36
	AftA	arabinosyltransferase	14.55	12.1	7	72.52
	Glft2	galactosyltransferase	6.91	30.03	9	35.28
	DprE1	decaprenylphosphoryl-beta-D-ribose 2'-oxidase	6.37	11.48	4	53.03
	RfbE	ABC-type transporter	2.31	22.05	7	28.66
VII	RplC	50S ribosomal protein L3	85.60	61.93	20	23.15
	NCgl1242	SAM-dependent methyltransferase	84.25	28.86	8	21.92
	GroEL	60 kDa molecular chaperone	83.15	45.35	26	56.71
	NCgl0811	inositol monophosphatase family protein	81.90	43.65	11	27.26
	RplD	50S ribosomal protein L4	77.76	60.55	17	23.59
	RpsC	30S ribosomal protein S3	73.26	76.21	28	28.09
	NCgl2383	hypothetical protein	63.69	45.69	14	42.33
	NCgl1065	Rossmann fold nucleotide-binding protein	56.46	49.22	18	28.67
	NCgl0286	cAMP-binding domain-containing protein	54.14	62.56	17	24.97
	AmyE	ABC-type transporter	37.30	21.9	10	49.42
	AftA	arabinosyltransferase	32.98	17.85	10	72.52
	NCgl0319	glycosyltransferase	24.48	47.46	13	26.31
	RfbE	ABC-type transporter	22.89	40.3	13	28.66
	GlgB	glycogen branching enzyme	20.44	20.38	15	82.54
	UbiA	decaprenyl-phosphate phosphoribosyltransferase	2.45	14.76	6	36.03

	DprE1	decaprenylphosphoryl-beta-D-ribose 2'-oxidase	2.33	20.7	8	53.03
VIII	NCgl2607	inorganic pyrophosphatase	92.53	60.13	14	17.89
	RplF	50S ribosomal protein L6	83.58	66.85	15	19.32
	GroEL	60 kDa molecular chaperone	82.80	45.91	28	56.71
	RpsG	30S ribosomal protein S7	76.86	85.16	17	17.47
	NCgl2383	hypothetical protein	73.63	49.24	15	42.33
	SmpB	SsrA-binding protein	44.75	45.73	15	18.99
	AtpF	FOF1 ATP synthase subunit B	44.59	40.43	13	21.08
	UreE	urease accessory protein	43.02	49.04	9	17.63
	RplJ	50S ribosomal protein L10	42.99	61.4	15	17.95
	RplC	50S ribosomal protein L3	34.59	63.3	17	23.15
	RplB	50S ribosomal protein L2	33.35	58.93	16	31.11
	AftA	arabinoxyltransferase	21.83	18	11	72.52

a) number of excised SDS-PAGE gel band for protein identification; b) the Mascot score of protein identification; c) sequence coverage in % of the identified protein; d) number of identified unique peptides obtained for a particular protein; e) molecular weight in kDa of the respective protein. Yellow background indicates proteins that are involved in the biosynthesis of cell wall that were identified in the assay.

2.3.4 Bacterial two-hybrid analysis of AG proteins

To characterise the physical interactions between components of the *C. glutamicum* cell wall biosynthetic machinery, the following full-length proteins listed in Table 5, were tested systematically for pair-wise interactions using BACTH. Each protein was fused to the fragment of the catalytic domain of chimeric adenylate cyclase (T25 or T18) of *Bordetella pertussis* at either the C- or N-terminus. Interaction between two hybrid proteins leads to reconstitution of the fragments of adenylate cyclase resulting in restoration of cAMP production in a *E. coli cya* mutant (Figure 2) (Karimova *et al.*, 1998). The resulting cAMP forms a complex with the catabolite activator protein and binds to various promoters, thus regulating transcription of several genes, including the lactose and maltose operons. The activation of these operons can be detected on selective agar plates or using a β -galactosidase assay. Importantly, this bacterial two-hybrid system was shown to be suitable to detect interactions between both cytoplasmic and

transmembrane or membrane associated proteins (Baulard *et al.*, 2003, Clarke *et al.*, 2009).

Table 5 Predicted topology and function of *C. glutamicum* proteins described in this study.

Protein	Predicted topology	Function
WecA	transmembrane	UDP-GlcNAc-1-phosphatetransferase
WbbL	soluble	α -3-L-rhamnosyltransferase
GlfT1	soluble	UDP-galactofuranosyltransferase
GlfT2	soluble	UDP-galactofuranosyltransferase
AftA	transmembrane	arabinofuranosyltransferase
AftB	transmembrane	arabinofuranosyltransferase
AftC	transmembrane	arabinofuranosyltransferase
AftD	transmembrane	arabinofuranosyltransferase
Emb	transmembrane	arabinofuranosyltransferase
DprE1	soluble	decaprenylphosphoryl- α -D-ribose 2'-oxidase
DprE2	soluble	decaprenylphosphoryl-D-2-keto erythro pentose reductase
UbiA	transmembrane	decaprenyl-phosphate 5- phosphoribosyltransferase

Despite several attempts, we did not succeed in generating pKT25 and pUT18 derivatives expressing UbiA and AftD proteins, respectively. This is probably due to the toxicity of hybrid proteins to bacterial cells when expressed at high levels, which is especially true of membrane proteins. Moreover, the UbiA-T18^N and UbiA-T18^C hybrid proteins, when co-expressed with several other hybrid proteins, reduced bacterial growth suggesting that UbiA production in large quantities is toxic to *E. coli* cells.

To examine putative interactions between the hybrid proteins, *E. coli* BTH101 cells were co-transformed with pairs of recombinant plasmids Table 2. In total, 577 pairs were screened for protein-protein interactions *in vivo*. All co-transformants, together with the positive and negative controls, containing either pKT25-*zip*/pUT18c-*zip* or empty pKT25/pUT18, were then spotted onto selective agar plates and the colouration of the

colonies were observed after 48 hours of growth at 30 °C. In the absence of association between T25 and T18 fragments colonies appear white, whereas they are blue or red when functional complementation occurs. The results of this protein-protein interaction study are summarised in Table 6 whilst representative plates from the screening are shown in Appendix (Figure 1-13). The efficiency of functional complementation between T25 and T18 domains was quantified by measuring β -galactosidase activity of each resultant strain. Ultimately, 50 pairs of hybrid proteins resulted in a positive signal representing 24 putative homotypic and heterotypic protein-protein interactions.

2.3.5 Self-association of *C. glutamicum* cell wall biosynthesis proteins

Among all the tested proteins, dimerisation or multimerisation of WecA, GlfT1, GlfT2, AftA, AftB, AftC, DprE1, and DprE2 have been demonstrated employing BACTH. Co-expression of transmembrane WecA-T25^C and WecA-T18^C hybrid proteins restored a *cya*⁺ phenotype and synthesis of cAMP in the *E. coli* cells, resulting in blue and red colonies on LB/M63-Xgal and MacConkey media, respectively (Appendix, Figure 1). The β -galactosidase assay revealed a significant increase in β -galactosidase activity (487 ± 47 Miller units) when compared to the negative control (86 ± 11 Miller units), containing empty pKT25, pKNT25, pUT18, and pUT18c plasmids (Appendix, Figure 1). Importantly, the transmembrane fusions have to be correctly inserted into the plasma membrane with the T18 and T25 domains facing the cytoplasm in order an interaction to be detected, therefore, suggesting that the C- terminus of GlcNAc-1-phosphate transferase WecA is cytoplasmic. This is in agreement with the predicted topology of WecA (Amer and Valvano, 2001). Physical self-dimerisation or multimerisation *in vivo* was also demonstrated for GlfT1 and GlfT2. Consistently, GlfT1-T25^N and GlfT1-T18^C, GlfT1-T25^N and GlfT1-T18^N, GlfT1-T25^C and GlfT1-T18^N, GlfT2-T25^N and GlfT2-

T18^C, and GlfT2-T25^C and GlfT2-T18^N hybrids restored *lac*⁺ and *mal*⁺ phenotypes and resulted in significant β -galactosidase activity ranging from 266 ± 69 to 679 ± 118 Miller units (Appendix, Figure 1). Recently, the structure of the polymerising GlfT2 orthologue in *M. tuberculosis* has been solved revealing its assembly as a homotetramer (Wheatley *et al.*, 2012), thus supporting the results obtained in this BACTH study.

Transmembrane AftA, AftB and AftC, proteins also tested positive for self-association. Co-expression of AftA-T25^N and AftA-T18^N, AftB-T25^N and AftB-T18^N, AftC-T25^C and AftC-T18^C combinations yielded β -galactosidase activity of 714 ± 92 , 1185 ± 265 , and 398 ± 23 Miller units, respectively (Appendix, Figure 2). The C-terminal region of AftA and AftB is predicted to face the periplasm (Alderwick *et al.*, 2006, Seidel *et al.*, 2007), therefore the lack of interaction between fusion pairs carrying C-terminal T25 or T18 fragment was expected and supported by this study. In addition, BACTH experiments propose that the N- termini of AftA and AftB are cytoplasmic. In contrast to AftA and AftB, AftC is characterised by the absence of a periplasmic C- terminal extension (Birch *et al.*, 2008). Hence it is unsurprising that multimerisation of AftC is observed with the fusion proteins tagged at the C- terminus. Interestingly, no evidence for homodimerisation could be obtained for Emb and AftD. Finally, DprE1 and DprE2, both involved in DPA synthesis, appeared positive for self-interaction. DprE1-T25^N and DprE1-T18^C, DprE1-T25^C and DprE1-T18^N fusions, as well as all four pairs of hybrid proteins co-expressing DprE2 led to a strong *lacZ* induction (ranged between 291 ± 33 and 1156 ± 54 Miller units), significantly exceeding the negative control (Appendix, Figure 3).

Table 6 Protein-protein interactions between *C. glutamicum* AG biosynthetic proteins determined by BACTH.

	WecA	WbbL	Glft1	Glft2	AftA	AftB	AftC	AftD	UbiA	DprE1	DprE2	Emb
WecA	✓											
WbbL	-	-										
Glft1	-	-	✓									
Glft2	-	-	-	✓								
AftA	-	-	-	-	✓							
AftB	✓	-	✓	-	✓	✓						
AftC	✓	-	-	-	✓	✓	✓					
AftD	-	-	-	-	-	-	-	-				
UbiA	✓	✓	-	-	✓	✓	✓	-	-			
DprE1	-	-	-	-	-	-	-	-	-	✓		
DprE2	-	-	✓	-	✓	✓	✓	-	-	✓	✓	
Emb	✓	-	-	-	✓	✓	-	-	-	-	-	-

The positive interaction is indicated as (✓), whereas the lack of interaction is marked as (-)

2.3.6 *In vivo* interaction network among AG proteins

Next, we examined the interactions between different proteins involved in AG biosynthesis. Our results indicate that in addition to homodimerisation, WecA is also able to interact with multiple partners of AG biosynthesis Table 6. WecA-T18^C, when co-expressed with AftB-T25^N, AftC-T25^C, Emb-T25^N, and UbiA-T25^C yielded significant β -galactosidase activities 1082 ± 268 , 1047 ± 186 , 1346 ± 217 , and 1018 ± 137 Miller units, respectively (Appendix, Figure 4-5). BACTH also revealed an interaction with the rhamnosyltransferase WbbL, when UbiA hybrids were used as the ‘bait’. Co-transformation of either WbbL-T18^C or WbbL-T18^N together with UbiA-T25^C led to a restoration of cAMP cascade with β -galactosidase activities of 1094 ± 93 and 1195 ± 78 Miller units, respectively (Appendix, Figure 5). Our studies have demonstrated the physical interaction between Glft1 and AftB (Appendix, Figure 6), as well as the DprE2 involved in DPA formation (Appendix, Figure 7). Recent studies reported the physical interaction between Glft1 and Rv3789, a small multidrug resistance-like transporter (Larrouy-Maumus *et al.*, 2012). Rv3789 was proposed to target and stabilise membrane associated Glft1 (Larrouy-Maumus *et al.*, 2012). Further experiments demonstrated

evidence for a physical interaction between UbiA and AftA-T25^N (387 ± 22 Miller units) (Appendix, Figure 8), AftB-T25^N (1015 ± 185 Miller units) (Appendix, Figure 10) and AftC-T25^C (755 ± 118 Miller units) (Appendix, Figure 12), responsible for the biosynthesis of the arabinan domain of AG. Most of these ArafTs could also establish multiple interactions with each other. AftA, which primes the galactan chain of AG, associated with Emb, AftC and AftB (Appendix, Figure 7-9). In addition, AftB also interacted with Emb and AftC hybrid proteins (Appendix, Figure 10-11). Finally, *C. glutamicum* DprE1 was also found to strongly interact with DprE2 as a heterodimer. DprE1-DprE2 association has been identified with seven different plasmid combinations resulting in a significant β-galactosidase activity ranging between 637 ± 52 and 1027 ± 86 Miller units (Appendix, Figure 13). Previous studies reported that orthologues of DprE1 and DprE2 in *M. tuberculosis* were able to catalyse the epimerisation reaction *in vitro*, however, neither protein alone was sufficient to support this activity (Larrouy-Maumus *et al.*, 2012). Thus, strongly suggesting that DprE1 and DprE2 work in concert to catalyse the conversion of DPR to DPA. However, when the same *M. tuberculosis* orthologues were experimentally tested for interaction using BACTH, co-transformants yielded negative results (Larrouy-Maumus *et al.*, 2012).

2.3.7 Verification of recombinant protein expression

Importantly, lack of *lacZ* induction might be a result of plasmid instability, insoluble or dissipating fusions, and not the lack of direct physical interaction. In order to determine whether the hybrid proteins are expressed appropriately for protein-protein interactions to be detected, Western blot analysis using antibodies against T25 and T18 fragments was performed. Single expression studies of hybrid proteins fused to T18 domain proved the production and accumulation of expected size AftA-T18^N (92 kDa), AftB-T18^N (94 kDa),

Glft1-T18^N (53 kDa), Glft1-T18^C (53 kDa), Glft2-T18^N (92 kDa), Glft2-T18^C (92 kDa), DprE1-T18^N (71 kDa), DprE1-T18^C (71 kDa), DprE2-T18^N (45 kDa), DprE2-T18^C (45 kDa), UbiA-T18^N (54 kDa), WbbL-T18^N (50 kDa), and WbbL-T18^C (50 kDa) proteins (Figure 6). Emb-T18^N (141 kDa), AftB-T18^C (94 kDa), and WecA-T18^C (58 kDa) fusion proteins were all expressed and detected by a Western blot, however, their determined molecular weight was lower than expected, possibly due to degradation or the hydrophilic nature of these proteins. We have also observed lack of expression of AftA-T18^C, AftD-T18^N, Emb-T18^C, UbiA-T18^C, AftC-T18^C, AftC-T18^N, and WecA-T18^N.

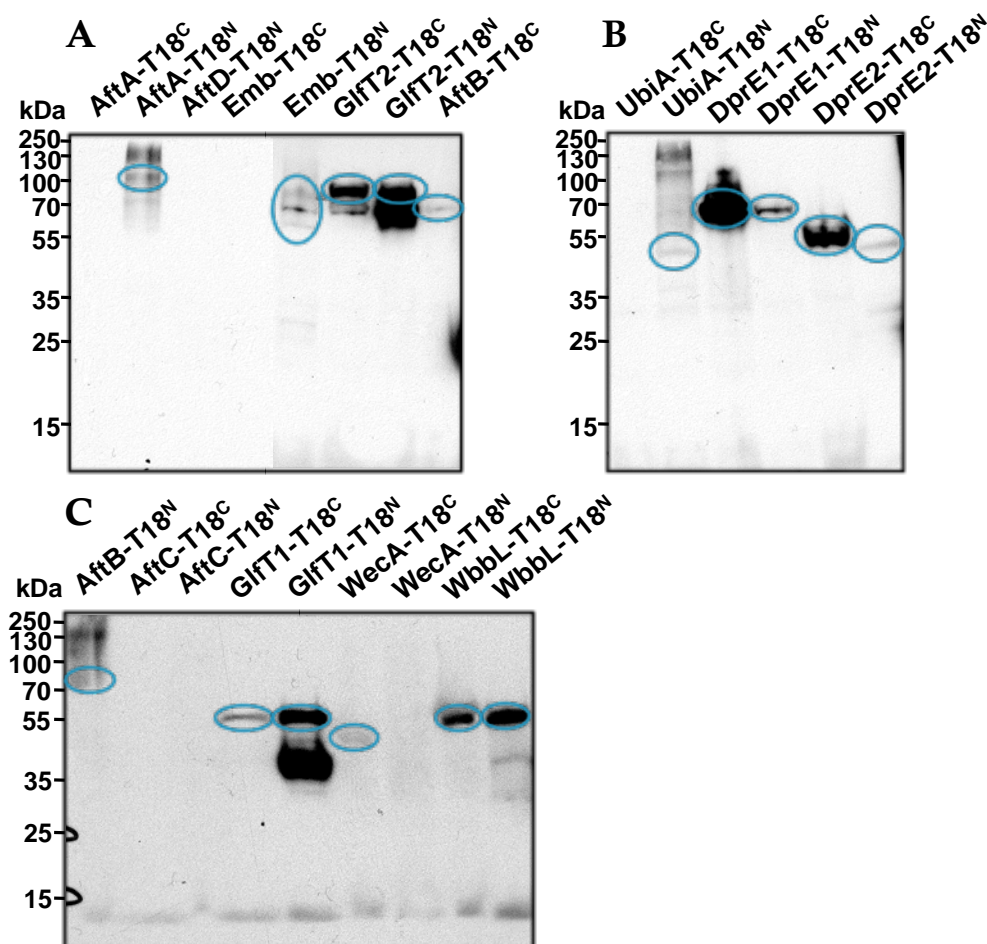


Figure 6 Single expression of hybrid proteins fused with T18 fragment as determined by Western blot analysis. A, B, and C Samples were analyzed by 12 % SDS-PAGE gel; fusions with T18 fragments were detected by Western blot analysis using monoclonal CyaA 3D1 antibody.

Single expression studies of hybrid proteins fused to T25 fragment led to very low levels of protein expression that were undetectable on a Western blot (Figure 7A). This is perhaps a result of different replication origins, p15A of pKT25 and pKNT25 plasmids (low copy number), compared to ColE1 in pUT18 and pUT18c plasmids (high copy number), thus resulting in production of poor amounts of T25 tagged proteins. In addition, we have investigated the co-expression of GlfT2 recombinant proteins (Figure 7B). Protein pairs (GlfT2-T25^C and GlfT2-T18^N; GlfT2-T25^N and GlfT2-T18^C), which were identified as interacting, showed expression and accumulation of both T25 and T18 fused proteins. However, in the samples where negative interactions were observed, only T18 tagged protein production was detected (Figure 7B). This could be explained by the positive feedback loop triggered by cAMP/CAP complex (Figure 7C). Hybrid protein expression is induced using IPTG, however, cAMP/CAP complex facilitates lactose operon induction. Therefore, positive protein-protein interaction significantly increases cAMP levels, which results in a contribution to the promoter activation and sequentially higher levels of recombinant protein production.

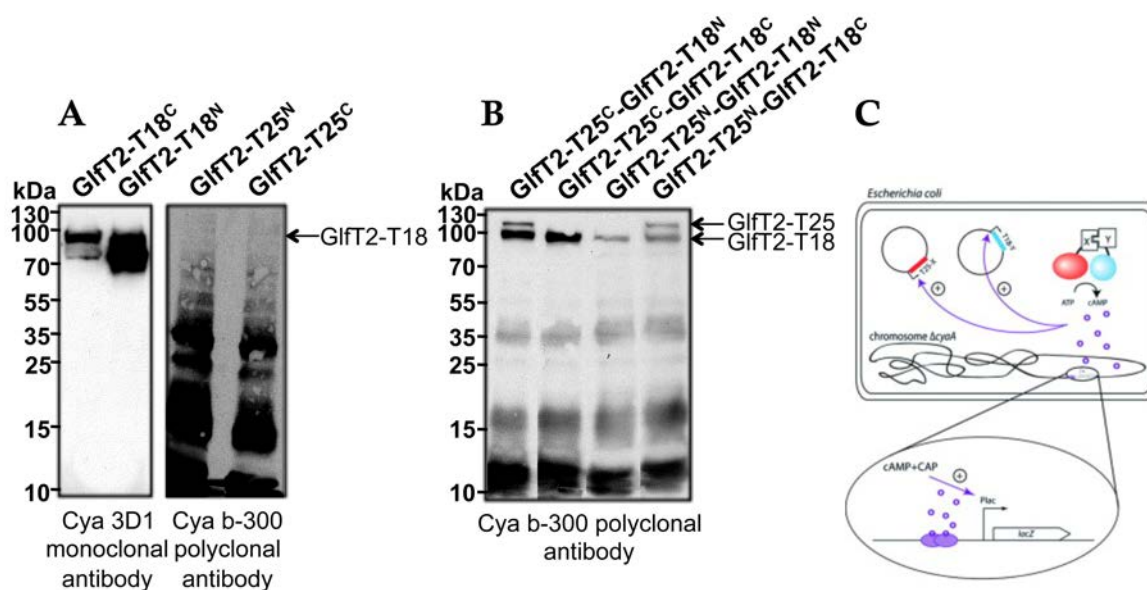


Figure 7 Expression of GlfT2 fusions as determined by Western blot analysis. A Single-expression studies of GlfT2 hybrid proteins. **B** Co-expression studies of GlfT2 hybrid proteins. **C** An interaction between proteins restores cAMP production, which results in a positive feedback loop – expression of lactose and maltose operons (Adapted from Battesti and Bouveret, 2012 with permission). Samples were analyzed by 12 % SDS-PAGE gel; the GlfT2 hybrids fused with T25 and T18 fragments were detected by Western blot analysis using polyclonal CyaA b-300 and monoclonal CyaA 3D1 antibodies, respectively. Adapted by permission from Elsevier: *Methods*, Battesti and Bouveret (2012).

2.4 Discussion

The majority of bacterial polysaccharides important for maintaining cell structure and viability are built on a carrier lipid in the cytosolic side of the plasma membrane. Although it is not fully clear how and when these polymers are translocated to the periplasm, one could speculate that anchoring these macromolecules to the membrane, positions them closely to the transporters and glycosyltransferases, therefore promoting productive export across the plasma membrane. Formation of multi-protein complexes, that contain glycosyltransferases, enzymes forming its sugar nucleotides and transporters,

is expected to be beneficial for the bacterial cell, since the tight arrangement of the biosynthetic reactions would retain productivity and accuracy of the polymerisation process.

We have demonstrated that proteins responsible for the formation of the AG linker unit, WecA and WbbL, form a complex with decaprenylphosphoryl-5-phosphoribose synthase UbiA at the cytoplasmic membrane (Figure 8). WecA and UbiA directly employ decaprenyl-phosphate for the linker unit and DPPR formation, respectively (Mikusova *et al.*, 1996, Mikusova *et al.*, 2005, Huang *et al.*, 2008). Proximal interactions between WecA, WbbL and UbiA could perhaps facilitate synchronised utilisation of decaprenyl-phosphate for coordinated AG biosynthesis. In addition, UbiA show evidence for physical interaction between AftA, AftB and AftC proteins, which employ DPA as a substrate (Figure 8). It is possible that this multi-protein complex formation assists a mechanism similar to substrate channeling, where intermediary metabolic products of one enzyme are passed directly to another enzyme. Other DPA forming proteins, DprE1 and DprE2, showed evidence for a physical interaction. Interestingly, while both DprE1 and DprE2 are required for the epimerisation reaction, there is evidence that *C. glutamicum* NCgl1429 may play a similar function to DprE2 (Meniche *et al.*, 2008). Investigation into potential DprE1-NCgl1429 complexes could provide insight into this gene redundancy. Notably, GT-A glycosyltransferases GlfT1 and GlfT2 showed evidence for homodimerisation using BACTH. GlfT1 transfers the first two *Galf* residues to the linker unit, whereas GlfT2 is responsible for addition of approximately 30 *Galf* residues in a linear chain (Kremer *et al.*, 2001, Mikusova *et al.*, 2006, Alderwick *et al.*, 2008, Belanova *et al.*, 2008). The recent crystal structure of *M. tuberculosis* GlfT2 and in complex with a UDP, established its homotetrameric architecture (Wheatley *et al.*, 2012).

Finally, AftA, AftB, AftC and Emb proteins involved in the assembly of arabinan domain in AG, indeed form a multi-protein complex at the cytoplasmic membrane (Figure 8). One could speculate that such a sophisticated complex would maintain the efficiency and fidelity of AG polymerisation.

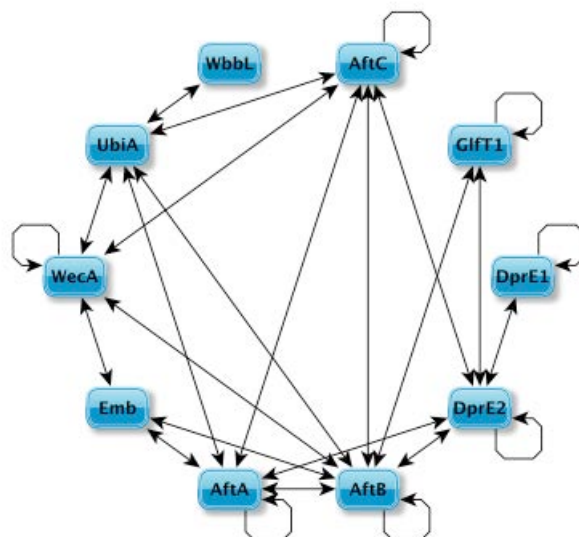


Figure 8 An interaction network of *C. glutamicum* proteins involved in AG biosynthesis. Graph was generated using yEd graph editor software. The circular arrows indicate self-association.

BACTH is a powerful technique for the investigation of protein-protein associations, however, several important notes should be highlighted regarding the significance of the interaction data obtained from BACTH. Firstly, the lack of cAMP production might be a direct result of hybrid proteins being expressed inappropriately or not expressed at all. We could not analyse the co-expression of all the protein pairs, which were identified as non-interacting due to large quantity of samples. Therefore, the hybrid proteins that test negative for interactions may still interact *in vivo*. Moreover, since the output of the interaction – cAMP – requires to be generated in the cytoplasm, these negative results may also result from the incorrect topological orientation of functional T25 and T18 domains into plasma membrane. In addition, using BACTH, the fusion proteins are

overexpressed when compared to the expression levels of native cells. Under these conditions, BACTH could have revealed a number of weak interactions between AG biosynthetic proteins. Although such associations would not take place at low protein concentrations, they can still occur when AG is being synthesised, where the local concentrations of proteins should be significantly higher. Finally, it is possible that some of the identified interactions are a consequence of a non-specific interactions initiated by endogenous *E. coli* host proteins that act as a tethering agent. These indirect associations, caused by a third protein, cannot be simply rejected.

In conclusion, our findings here suggest that enzymes involved in *C. glutamicum* cell wall assembly and precursor formation form complicated multi-protein complexes. We have identified 24 interactions *in vivo* between 12 proteins responsible for AG biosynthesis using BACTH. The challenge for the future will be to discover precisely how each of these multi-protein complexes form and function to synthesise and translocate AG.

Chapter 3

Molecular and biochemical
characterisation of AftA and
AftB arabinosyltransferases
from *Corynebacterineae*

3 Molecular and biochemical characterisation of AftA and AftB arabinosyltransferases from *Corynebacterineae*

3.1 Introduction

The thick, carbohydrate and lipid rich cell wall with distinct lipoglycans enables *M. tuberculosis* to survive under hostile conditions, such as shortage of nutrients and antimicrobial exposure. The key features of this highly complex cell wall are the mAGP complex and phosphatidyl-*myo*-inositol derived lipoglycans with potent immunomodulatory properties, notably LM and LAM. These structures are crucial for the growth, viability, and virulence of *M. tuberculosis* and therefore, are often the targets of effective chemotherapeutic agents against TB.

Mycobacterial LAM is generated by elaborating LM with 55 to 70 Araf units forming an arabinan domain similar to that found in AG (Besra and Brennan, 1997). It is speculated that a yet uncharacterised ArafT primes LM for LAM synthesis. Further transition occurs by addition of 12 to 16 $\alpha(1\rightarrow5)$ -Araf residues to the primed LM, a reaction exclusively catalysed by EmbC possessing $\alpha(1\rightarrow5)$ activity (Shi *et al.*, 2006, Goude *et al.*, 2008, Alderwick *et al.*, 2011). The linear arabinose polymer of LAM is branched similarly to AG by addition of $\alpha(1\rightarrow3)$ Araf residues employing AftC, resulting in 3,5 Araf branch points (Birch *et al.*, 2008, Birch *et al.*, 2010). Recently, AftD was designated as a second enzyme with $\alpha(1\rightarrow3)$ ArafT activity (Skovierova *et al.*, 2009). *In vitro* assays using neoglycolipid acceptors and cell-free extracts from *M. smegmatis* showed that the enzyme was able to add $\alpha(1\rightarrow3)$ Araf residues to the linear α -1,5-linked acceptor, resulting in branching of the linear arabinan. Therefore, its function is considered similar to that of AftC. The branched arabinose motif of LAM is further modified to either a

tetra-arabinoside or a hexa-arabinoside motif. The terminal $\beta(1\rightarrow2)$ Araf residues present in both tetra- and hexa-motifs are perhaps added by AftB (*Rv3805c*), which may have a dual functionality and play a role in both AG and LAM biosynthesis.

A priming enzyme AftA has been identified in *Corynebacteriaceae* as that attaches arabinose residues to the galactan chain of AG at 8th, 10th and 12th positions (Alderwick *et al.*, 2006, Shi *et al.*, 2008). The chromosomal region of *aftA* includes *dprE1* (*Rv3790*) and *dprE2* (*Rv3791*), which collectively encode a decaprenylphosphoryl-5-phosphoribose (DPPR) epimerase, and are responsible in generating the sugar donor DPA. Although, the function of AftA in AG biosynthesis has been defined in both *M. smegmatis* and *C. glutamicum*, its involvement in LAM biosynthesis remains unknown. The *aftB* locus of *M. tuberculosis* consists of UDP-Galp mutase Glf and the known GlfT1. Just upstream *aftB* is located *ubiA*, which encodes the decaprenylphosphoryl-5-phosphoribose (DPPR) synthase. The *fbpA* and *fbpD* genes are located downstream of *aftB*, and encode mycolyltransferases that decorate the terminal arabinan residues with mycolates (Belisle *et al.*, 1997, Kremer *et al.*, 2002). The essentiality of *aftB* is implied by its conservation within *Corynebacteriaceae*, including all *Mycobacterium* and *Corynebacterium* species (Seidel *et al.*, 2007a).

In mycobacteria, biosynthesis of essential cell wall structures, including PG, AG, and LM/LAM, are mediated by different glycosyltransferases. These enzymes catalyse the transfer of a sugar residue to various targets forming glycosidic bonds. Three main glycosyltransferase superfamilies have been established based upon their protein fold; (i) a nucleoside-diphosphosugar transferase (GT-A) family that is characterised by a conserved DxD motif; (ii) GT-B family also known as GPGTF family, and (iii) a GT-C

family of integral membrane proteins with a modified DxD signature (Liu and Mushegian, 2003). Examples of GT-C family enzymes are found in all eukaryotes, but are missing from archaea, and are uncommon in prokaryotes with an exception of mycobacteria and related species of the order *Actinomycetales*. Whilst proteins belonging to both GT-A and GT-B are globular in nature, the GT-C family enzymes are large hydrophobic glycosyltransferases that are presented in the cell as transmembrane or membrane associated proteins with 8 to 13 predicted multiple transmembrane domains. The conserved element in this family is a modified DxD sequence (e.g., DxEx, ExD, DDx, DEx, or EEx) typically located in the first or second extracytoplasmic loop and a long C-terminal region directed toward the periplasm (Liu and Mushegian, 2003). The conserved modified DxD sequence is predicted to bind the lipid linked sugar donor substrate *via* a divalent cation most commonly Mn^{2+} or Mg^{2+} (Tarbouriech *et al.*, 2001), however, the exact function of this motif has not been fully elucidated. Several studies employing site directed mutagenesis have demonstrated that substitution of specific conserved aspartic acid residues to alanine within the DxD sequence in PimE and EmbC of *M. smegmatis*, have led to a loss of enzymatic activity (Berg *et al.*, 2005, Morita *et al.*, 2006). Similarly, topology and mutational analysis of the single Emb from *C. glutamicum* highlighted the importance of the conserved aspartic residues within the modified DXD motif to the function of Emb (Seidel *et al.*, 2007b). In addition, 20-40 amino acids downstream of the DxD motif, conserved proline residue(s) and a partially conserved acidic amino acid have been identified and, in conjunction to DxD sequence, is referred to as the GT-C motif (Berg *et al.*, 2007). Finally, the study by Zhang *et al.* (2003) suggested that Emb proteins utilise their N-terminal region to recognise the acceptor, while the long periplasmic C-terminal region is responsible for arabinan formation. Recently, Alderwick *et al.* (2011)

unravelling the structure of the hydrophilic C-terminal domain of EmbC (EmbC^{CT}) GT-C family glycosyltransferase, which is essential to *M. tuberculosis* growth and viability. EmbC is responsible for the elongation of the arabinan domain of LAM and is inhibited by the front-line anti-TB drug EMB (Goude *et al.*, 2008, Goude *et al.*, 2009). The structure of EmbC^{CT} contains two sub-domains both associated with two separate carbohydrate binding sites (Alderwick *et al.*, 2011). The sub-domain I showed weak structural similarity with the C-terminal domain of oligosaccharyltransferase STT3 from *Pyrococcus furiosus*, whereas sub-domain II demonstrated distinct similarity to lectin-like carbohydrate-binding modules (Alderwick *et al.*, 2011). Site-directed mutagenesis of the conserved tryptophan residues located in sub-domain I and II led to the disruption of EmbC enzymatic activity, thus confirming EmbC^{CT} function as a carbohydrate binding module (Zhang *et al.*, 2003, Alderwick *et al.*, 2011). Mycobacterial glycosyltransferases, such as AraFTs, represent ideal targets for drug development and are targeted by several anti-TB drugs, however their structure and mode of action remain poorly understood.

Structural and functional analysis of membrane proteins lags behind that of soluble proteins. This stems from the inherent challenge of extracting pure membrane proteins from their native membrane in an active state, a process that is not required when purifying soluble proteins. The low success rate of conventional detergent or surfactant based solubilisation, which greatly limits the number and type of membrane proteins available for study, has been attributed to many factors, the most significant being the inability of detergents to recreate the complex architecture of native membranes within which proteins normally reside. In recent years new approaches have addressed this limitation (Ritchie *et al.*, 2009, Popot, 2010, Jamshad *et al.*, 2011). The most prominent being the use of membrane associated peptides to produce nano-dimensional, disc-shaped

segments of membrane bilayers into which membrane proteins can be inserted (Ritchie *et al.*, 2009, Velez-Ruiz and Sunahara, 2011). However, this method does not allow membrane proteins to be solubilised directly from a source membrane. Instead, the protein must first be extracted with detergent and then inserted into a membrane disc. The inexpensive and biocompatible poly(styrene-*co*-maleic acid) polymer, which added to a liposome result in instantaneous formation of 11 nm diameter bilayer discs, encapsulate active bacterial membrane proteins without the use of detergent (Figure 1) (Knowles *et al.*, 2009, Jamshad *et al.*, 2011). This polymer can be used to extract a range of membrane proteins from different host membranes, and the resultant ‘poly(Styrene-*co*-Maleic Acid)-Lipid Particles’ (SMALPs) can be subjected to purification and subsequent study.

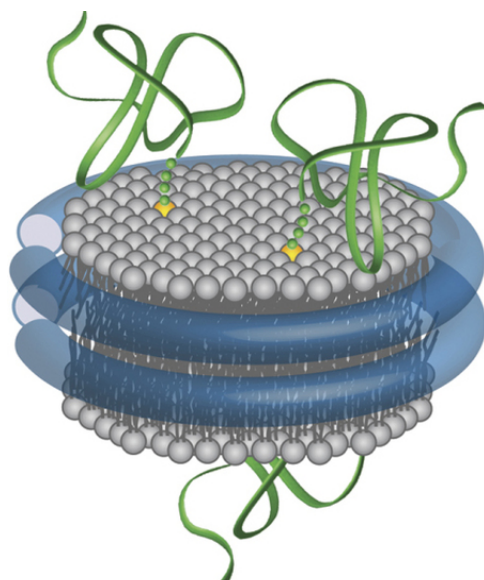


Figure 1 Transmembrane protein encapsulated in SMALP disc. SMALPs preserve the functional and structural integrity of α -helical and β -barrel transmembrane proteins. They form approximately 11 nm particles that are monodispersed, biocompatible, thermostable and water-soluble allowing diverse membrane proteins to be simply and rapidly prepared for an *in vitro* analysis. Adapted by permission from Macmillan Publishers Ltd: *Nature Protocols*, Shi *et al.* (2013).

This study focuses on defining the involvement of AftA and AftB in LM/LAM biosynthesis through the use of a novel conditional expression-specialised transduction essentiality test (CESTET), and subsequent analysis of the cellular composition of mutants. In addition, we employ a number of biochemical techniques in order to characterise full-length AftA and AftB membrane proteins as well as their C-termini.

3.2 Materials and methods

3.2.1 Construction of merodiploid strains

The 1875 bp and 1938 bp coding regions for *aftA* (*MSMEG_6386*) and *aftB* (*MSMEG_6400*), respectively, were amplified by PCR from *M. smegmatis* mc²155 genomic DNA using the primer pairs (restriction sites underlined): 5'-GAT CGA TCG GAT CCG TGG CGG CCA GGG TTC TCG-3' (*aftA* forward), GAT CGA TCA TCG ATT CAG TGG CCA TCG GTC TCC-3' (*aftA* reverse), 5'-GAT GAT CGG ATC CGT GCG CAT CAG CCT GTG GC-3' (*aftB* forward) and 5'-GAT CGA TCA TCG ATC TAC GGT CCC GTT GCC GGC-3' (*aftB* reverse). The single copy integrating constructs pMV306-*aftA* and pMV306-*aftB* were generated by ligating a 2.6 kb XbaI-BamHI inducible acetamide promoter and either *aftA* or *aftB* digested with BamHI-ClaI into XbaI-ClaI digested pMV306 yielding pMV306 constructs containing the *aftA* or *aftB* gene cloned downstream of the acetamide promoter. The merodiploid strains were generated by electroporating wild type *M. smegmatis* mc²155 with pMV306-*aftA* and pMV306-*aftB* and selecting for kanamycin resistant colonies.

3.2.2 Construction of *aftA* and *aftB* deletion mutants in *M. smegmatis*

Approximately 1 kb of upstream and downstream flanking sequences of *MSMEG6386* (*aftA*) and *MSMEG6400* (*aftB*) were PCR amplified from *M. smegmatis* mc²155 genomic DNA using the primer pairs: MS6386LL (5'-TTT TTT TTC CAT AAA TTG GGG CTT GGA GCT GCA GTA G-3') and MS6386LR (5'-TTT TTT TTC CAT TTC TTG GAG CGC ACC CAG GCG CGC C-3'), and MS6386RL (5'-TTT TTT TTC CAT AGA TTG GGC CGC ACT CGA CGA GCT G-3') and MS6386RR (5'-TTT TTT TTC CAT CTT TTG GGT CGG CGC AGT CCA GCA C-3') for *aftA* from *M. smegmatis*; MS6400LL (5'-TTT TTT TTC CAT AAA TTG GGA GTT ACA CCA GCA GCT ACC-3') and MS6400LR (5'-TTT TTT TTC CAT TTC TTG GAC CAG CAC ACC ATC ATC C-3'), and MS6400RL (5'-TTT TTT TTC CAT AGA TTG GAT GGG CAT GCT GGG CAT GAA CG-3') and MSMEG6400RR (5'-TTT TTT TTC CAT CTT TTG GCA CCG AGA TGC CCG AGT TGT AG-3') for *aftB* from *M. smegmatis*. Following restriction digestion of the primer incorporated *Van91I* sites, the PCR fragments were cloned into *Van91I*-digested p0004S to yield the knockout plasmids p Δ MSMEG6386 and p Δ MSMEG6400. The obtained vectors are sequenced and subsequently linearised by *PacI* and cloned into the *PacI*-digested phAE159 phasmid. The ph Δ MSMEG6386 and ph Δ MSMEG6400 phasmids were then packaged into the temperature-sensitive mycobacteriophage particles as described previously (Bardarov *et al.*, 2002). Generation of high titre phage particles and specialised transduction were performed as described earlier (Bardarov *et al.*, 2002). Deletion of the gene of interest was confirmed using Southern blot analysis. DNA sequencing and construct verification were carried out at Eurofins DNA Ltd.

3.2.3 Plasmid construction for overexpression of full-length AftA and AftB from *M. tuberculosis* and *C. glutamicum*

The coding regions for *Rv3792*, *Rv3805c*, *NCgl0185* and *NCgl2780* were amplified by PCR from *M. tuberculosis* H37Rv and *C. glutamicum* ATCC 13032 genomic DNA, respectively, using the primer pairs provided in the Table 1. The DNA fragments were restricted with the appropriate enzymes and ligated into pMSX or pEKEEx5 digested with identical enzymes, thus yielding pMSX-*Rv3792*, pMSX-*NCgl0185*, pEKEEx5-*Rv3792*, pEKEEx5-*NCgl0185*, pMSX-*Rv3805c*, pMSX-*NCgl2780*, pEKEEx5-*Rv3805c* and pEKEEx5-*NCgl2780*. Both pMSX and pEKEEx5 vectors carry a C-terminal and an N-terminal His₆-tag sequence, respectively. DNA sequencing and construct verification were carried out at Eurofins DNA Ltd.

Table 1 Primers used in this study.

Gene	Primer	Vector
<i>Rv3792</i> (<i>aftA</i>)	5'-CATGCATGCATATGATGCCGAGCAGACGCAAA-3' 5'-CATGCATGGAGCTCGCCGCGCTCTCCTGCGGCT-3'	pMSX
<i>NCgl0185</i> (<i>aftA</i>)	5'-CATGCATGCATATGATGATTAACACCTCTGAAG-3' 5'-CATGCATGGCGGCCGCCTCATTGTGCGTTACCAC-3'	pMSX
<i>Rv3792</i> (<i>aftA</i>)	5'-CATGCATGAGATCTATGCCGAGCAGACGCAAA-3' 5'-CATGCATGGAGCTCTCACGCGCTCTCCTGCGGC-3'	pEKEEx5
<i>NCgl0185</i> (<i>aftA</i>)	5'-CATGCATGGAGCTCATGATTAACACCTCTGAAG-3' 5'-CATGCATGGTTCGACTTACTCATTGTGCGTTACC-3'	pEKEEx5
<i>Rv3805c</i> (<i>aftB</i>)	5'-CATGCATGCATATGATGGTCCGGGTCAGCTTG-3' 5'-CATGCATGGAGCTCCTCCCGCGGTGGCGGGCC-3'	pMSX
<i>NCgl2780</i> (<i>aftB</i>)	5'-CATGCATGCATATGATGACGTTTAGCCCCCAG-3' 5'-CATGCATGGAGCTCCTGAGAGCTATATAAAGG-3'	pMSX
<i>Rv3805c</i> (<i>aftB</i>)	5'-CATGCATGGGATCCATGGTCCGGGTCAGCTTGT-3' 5'-CATGCATGGAGCTCTCACTCCCGCGGTGGCGG-3'	pEKEEx5
<i>NCgl2780</i> (<i>aftB</i>)	5'-CATGCATGGAGCTCATGACGTTTAGCCCCCAGC-3' 5'-CATGCATGGTTCGACTTACTGAGAGCTATATAAA-3'	pEKEEx5

3.2.4 Plasmid construction for overexpression of C-terminal domain of AftA and AftB from *M. tuberculosis*

The pET16b-*aftA*^{CT}, pET16b-*aftB*^{CT} constructs were kindly provided by Dr. L. Alderwick, University of Birmingham. Specifically, the C-terminus region of *aftA* (*Rv3792*), and *aftB* (*Rv3805c*) were amplified from *M. tuberculosis* H37Rv genomic DNA using oligonucleotides pairs (restriction sites underlined): 5'-GAT CGA TCC ATA TGC CCG ACG TGT TGC GGC CGG-3' (*aftA*^{CT} forward), 5'-GAT CGA TCG GAT CCT CAC GCG CTC TCC TGC GGC-3' (*aftA*^{CT} reverse), 5'-GAT CGA TCC ATA TGA ACT CGC CGG GCA TGG GTG-3' (*aftB*^{CT} forward), 5'-GAT CGA TC G GAT CCT CAC TCC CGC GGT GGC GGG-3' (*aftB*^{CT} reverse). Purified DNA fragments were cloned into appropriate cutting sites of pET16b vectors to express recombinant proteins tagged with His₆ at the N-terminus. DNA sequencing and construct verification was carried out at Eurofins DNA Ltd.

3.2.5 Genomic DNA extraction from *M. smegmatis*

M. smegmatis cells were resuspended in 450 µl of lysis buffer (25 mM Tris-HCl pH 8.0, 10 mM EDTA, 50 mM glucose, 1 mg/ml lysozyme) and the cell slurry incubated at 37 °C for 12 hours. A further 500 µl of 10 % SDS (Sigma Aldrich) and 250 µl of Protease K (10 mg/ml, Sigma Aldrich) were added and samples incubated at 55 °C for 4 hours. Ionic bonds between genomic DNA and proteins were disrupted by addition of 1 ml of 5 M NaCl at 65 °C and proteins were removed using chloroform: isoamyl alcohol (24:1, Sigma Aldrich) extraction. The DNA was precipitated using 2.8 ml of isopropanol, and the DNA pellet washed twice using 70 % ethanol and air-dried at room temperature. The

genomic DNA pellets were dissolved in Tris-EDTA buffer (10 mM Tris-HCl, 1 mM EDTA) and stored at -20 °C until further use.

3.2.6 Southern blot analysis

Genomic DNA extracted from wild type *M. smegmatis*, merodiploid strains (mc²155::pMV306-*aftA* and mc²155::pMV306-*aftB*) and deletion mutants (mc²155Δ*aftA*::pMV306-*aftA* and mc²155Δ*aftB*::pMV306-*aftB*) were digested with either EcoRI (for *aftA*) or SacI (for *aftB*). Digested samples were subjected to gel electrophoresis followed by gel depurination in 0.25 M HCl for 10 minutes. The agarose gel was then denatured in 1.5 M NaCl, 0.5 M NaOH for 15 minutes and subsequently neutralised for 30 minutes in 0.5 M Tris-HCl pH 7.2, 1 M NaCl. The separated digested genomic DNA samples were transferred from the agarose gel onto a nylon membrane (Roche) by capillary transfer using 3 M NaCl, 0.3 M sodium citrate pH 7.0. DNA was then fixed to the nylon membrane by UV cross-linking. The membrane was further used for hybridisation, labelling and detection as described in DIG High Primer DNA Labelling and Detection Starter Kit (Roche).

3.2.7 Conditional depletion of AftA and AftB

M. smegmatis strains were grown in either TSB or minimal medium (Chang *et al.*, 2009) supplemented with 0.05 % Tween-80 (TBST) and 0.2 % acetamide to an OD₆₀₀ of 0.5. Cells were washed twice with media to remove traces of acetamide and resuspended in the original volume of appropriate media. Culture was used as a 20 % inoculum in either TBST or minimal media and grown for 12 hours to deplete intracellular AftA or AftB.

The depleted sub-culture served as inoculum (5 %) for cultures with or without 0.2 % acetamide.

3.2.8 Protein production

DNA constructs were appropriately electroporated into *C. glutamicum* ATCC 13032 cells or transformed into *E. coli* BL21 (DE3), C43 (DE3), Tuner (DE3), and Rosetta (DE3) pLysS cells, and selected for resistance to appropriate antibiotics on BHI or LB agar plates, respectively. Recombinant cells harbouring constructs of interest were used to inoculate a pre-culture of 5 ml media supplemented with appropriate antibiotic. This pre-culture was then used to inoculate 1 l media containing appropriate antibiotic and incubated at 30 °C or 37 °C with shaking until an OD₆₀₀ of 0.5-0.8. IPTG was then used at a final concentration of 0.1 mM or 1 mM to induce protein expression and cells were further cultivated for additional 4 hours at 30 °C/37 °C or 12 hours at 16 °C. Cells were harvested, washed with phosphate buffered saline and pellets frozen at -20 °C until further use.

3.2.9 Purification of *M. tuberculosis* AftA^{CT}

For protein purification, pellet was thawed and resuspended in 50 mM KH₂PO₄ pH 7.9, 300 mM NaCl and 20 mM imidazole, containing one complete protease inhibitor cocktail tablet (Roche). The cell suspension was disrupted by sonication (MSE Soniprep 150, 12 micron amplitude, 30 seconds ON, 90 seconds OFF for a total of 10 cycles) and the cell slurry centrifuged for 30 minutes, 27,000 g at 4 °C. Supernatant was collected and passed over a 1 ml HiTrap Ni²⁺-NTA agarose column (GE Healthcare), which was previously equilibrated with 50 mM KH₂PO₄ pH 7.9, 300 mM NaCl and 20 mM imidazole. Elution

occurred *via* a stepwise gradient of 50-500 mM imidazole in 50 mM KH₂PO₄ pH 7.9, 300 mM NaCl and 10 ml fractions collected. Pure protein fractions were determined by 12% SDS-PAGE analysis. AftA^{CT} was dialysed against either 4 L of 50 mM Tris-HCl pH 7.5, 300 mM NaCl or 100 mM sodium acetate pH 4.2, 200 mM NaCl or 20 mM NaAc pH 6.0, 200 mM NaCl.

3.2.10 Differential scanning fluorimetry with AftA^{CT}

Protein unfolding was observed using the SYPRO Orange fluorescence dye (Invitrogen), which binds non-specifically to hydrophobic surfaces of unfolded proteins. Differential scanning fluorimetry (DSF) was set up in 96-well plates using a reaction volume of 40 µl. Firstly, the following buffers: sodium citrate pH 4.0, sodium acetate pH 4.2, sodium acetate pH 4.8, sodium citrate pH 5.0, sodium acetate pH 5.4, MES pH 5.7, sodium citrate pH 6.0, Bis-Tris pH 6.0, MES pH 6.1, imidazole pH 6.4, MES pH 6.5, Bis-Tris pH 6.5, Bis-Tris pH 7.0, HEPES pH 7.0, imidazole pH 7.1, Tris pH 7.2, HEPES pH 7.5, imidazole pH 7.6, HEPES pH 8.0, Tris pH 8.0, Tris pH 8.6, CHES pH 8.6, CHES pH 9.1, and CHES pH 9.6 containing either 190 mM or 340 mM NaCl were aliquoted to the 96-well plate. The final concentration of all buffers was 100 mM. Secondly, since SYPRO Orange dye is supplied in 100 % (v/v) DMSO and high concentrations of solvent have a tendency to damage proteins, the dye was diluted (1:1,000) into the buffers prior to addition of the protein. Finally, the AftA^{CT} protein dialysed in 100 mM sodium acetate pH 4.2 containing 200 mM NaCl was aliquoted into the 96-well plate to a final concentration of 100 ng/ml. Fluorescence measurements were collected employing real-time PCR machine (Agilent MX3005P with MxPro QPCR Software) using a temperature scan from 25 °C to 80 °C at 1 °C/min. Data was analysed using GraphPad Prism

software. Fluorescence intensities were plotted as a function of temperature. The effects of increasing temperature were expressed as ΔT_m .

3.2.11 AftA^{CT} crystallisation trials

Purified AftA^{CT} was concentrated to either 15 mg/ml or 30 mg/ml. The Molecular Dimensions optimised sparse matrix screens Structure screen I & IITM, Pact *premier*TM, JCSG-*plus*TM, MorpheusTM, MIDASTM as well as Hampton Research Crystal screen I & IITM were used to identify conditions that were most suitable for crystal growth. The sitting drop vapour diffusion format was used in a 96-well plate. The Mosquito nano-drop crystallisation robot was employed to aliquot the concentrated pure protein and different buffers.

3.2.12 Styrene maleic acid solubilisation

Membranes were prepared from *C. glutamicum*::pMSX-*aftA* as described in General Materials and Methods 6.5.13, re-suspended to a final concentration of 60 mg/ml (wet membrane weight) and homogenised in 50 mM Tris-HCl pH 8.0 supplemented with 500 mM NaCl and 5 mM imidazole. An equal volume of 5 % (w/v) styrene maleic acid resuspended in the same buffer was added to the membrane sample to yield a final concentration of 2.5 %. The sample was then incubated for 12 hours at 30 °C with gentle agitation. Removal of insoluble material was performed by centrifugation at 100,000 g for 1 hour at 4 °C. Once formed, SMALP encapsulated proteins were found to be stable and could be purified using a Ni²⁺ affinity purification method.

3.2.13 Purification of *C. glutamicum* AftA within SMALP

SMALP encapsulated *C. glutamicum* membranes containing recombinant AftA were incubated with 2 ml of Ni²⁺-NTA linked agarose (Qiagen) resin for 12 hours at 4 °C with gentle agitation. The sample was then transferred to 10 ml of disposable gravity flow column (Thermo Fisher), which was previously equilibrated with 5 column volumes of 50 mM Tris-HCl pH 8.0, 500 mM NaCl and 5 mM imidazole. Elution of *C. glutamicum* AftA was achieved using a step-wise imidazole gradient of 50-500 mM imidazole in 50 mM Tris-HCl pH 8.0, 500 mM NaCl and 1 ml fractions were collected and analysed using standard 12 % SDS-PAGE and Western blot analysis. Immunological detection of immobilised proteins was performed using a primary anti-hexa-His monoclonal antibody (Takara) and an anti-mouse IgG HRP-conjugated secondary antibody (Sigma Aldrich) according to manufacturer's instructions. Fractions containing *C. glutamicum* AftA were pooled, dialysed in 50 mM Tris-HCl pH 8.0 and concentrated. The protein concentration was determined using the bicinchoninic acid procedure (Pierce) and bovine serum albumin used as a standard.

3.2.14 Circular dichroism spectroscopy

Circular dichroism experiments were recorded at 20 °C using a Jasco J-810 spectropolarimeter fitted with a Peltier heating block using a path length of 0.01 mm. SMALP encapsulated *C. glutamicum* AftA protein was diluted to a final concentration of 0.25 mg/ml of Tris-HCl pH 8.0. Spectra were recorded at a scan speed of 100 nm/min with band width of 1 nm and wavelength range of 190 nm and 260 nm. Spectra were normalised by subtracting the spectrum of buffer alone. Data were processed using the program Spectra Manager I (JASCO).

3.2.15 Thermostability analysis of *C. glutamicum* AftA within SMALP

SMALP encapsulated *C. glutamicum* AftA was diluted to a concentration of 0.5 mg/ml in 10 mM Tris-HCl pH 8.0. Denaturation of AftA was measured by monitoring the temperature dependent changes of ellipticity at 222 nm in a 0.2 mm path length Quartz cuvette with a temperature gradient of 1 °C/min.

3.2.16 Tryptophan fluorescence binding assay

Intrinsic tryptophan fluorescence spectroscopy experiments were performed using a PTI QuantaMaster 40 spectrofluorimeter and recorded using the FeliX32 software package. The excitation wavelength was set to 283 nm and the fluorescence emission ($F_{emission}$) was recorded between 300-380 nm for each ligand aliquot added to a total of 300 μ l solution containing 2 μ M of SMALP encapsulated *C. glutamicum* AftA in 50 mM Tris-HCl pH 8.0. The change in fluorescence emission ($\Delta F_{emission}$) intensity was plotted against ligand concentration [L] (3 independent experiments) and fitted to the saturation binding equation using GraphPad Prism Software.

$$\Delta F_{emission} = F_{max} \times \frac{[L]}{(K_d + [L])}$$

3.2.17 Analytical ultracentrifugation

SMALP encapsulated *C. glutamicum* AftA was diluted to a concentration of 0.5 mg/ml in 10 mM Tris-HCl pH 8.0 and analysed by sedimentation velocity analytical centrifugation. The sample was centrifuged at 129,000 g at 4 °C using Beckman XL-1 analytical ultracentrifuge. The protein motion was monitored at absorbance of 280 nm with 1780 datasets. Data was analysed using SEDFIT and SENDTERP software.

3.3 Results

3.3.1 Molecular and biochemical characterisation of AftA

3.3.1.1 Evidence suggesting that AftA from *M. smegmatis* is an essential gene

In order to delete *aftA* from *M. smegmatis* and study possible consequences in LAM, we generated a conditional *M. smegmatis* Δ *aftA*::pMV306-*aftA* mutant using CESTET. Deletion of chromosomal *aftA* was achieved only in the presence of a rescue plasmid carrying a second copy of *aftA*, strongly suggesting that *aftA* is essential to viability of *M. smegmatis*. The replacement of the native chromosomal copy of *aftA* by a hygromycin resistance cassette in the transductants was confirmed by a Southern blot (Figure 2). These results are comparable to those by Shi et al. (2008), where *aftA* disruption in *M. smegmatis* was only possible with a rescue plasmid present.

Previous studies demonstrated that conditional depletion of essential enzymes in CESTET-derived conditional *M. smegmatis* mutants resulted in cell lysis (Bhatt *et al.*, 2005, Brown *et al.*, 2007, Rana *et al.*, 2012). The growth of *M. smegmatis* Δ *aftA*::pMV306-*aftA* in liquid medium with and without acetamide was monitored over 24 hours. While the strain grew normally in the medium containing acetamide, the culture lacking acetamide showed a decrease in OD₆₀₀ with time, resulting in a clearly lysed culture after 24 hours of incubation (Figure 3A). However, the presence of *M. smegmatis* cells that persisted throughout depletion of AftA has also been noticed (Figure 3A). Therefore, *M. smegmatis* Δ *aftA*::pMV306-*aftA* cells were grown for 24 hours with and without acetamide and subsequently plated onto appropriate agar plates (Figure 3B). Although a significant decrease in cell numbers was observed in the media lacking

acetamide, colonies formed by persisters were still visible (Figure 3B). Often conditional mutant strains acquire mutations in the acetamide promoter region resulting in continuous production of the protein of interest, thus rescuing the cell.

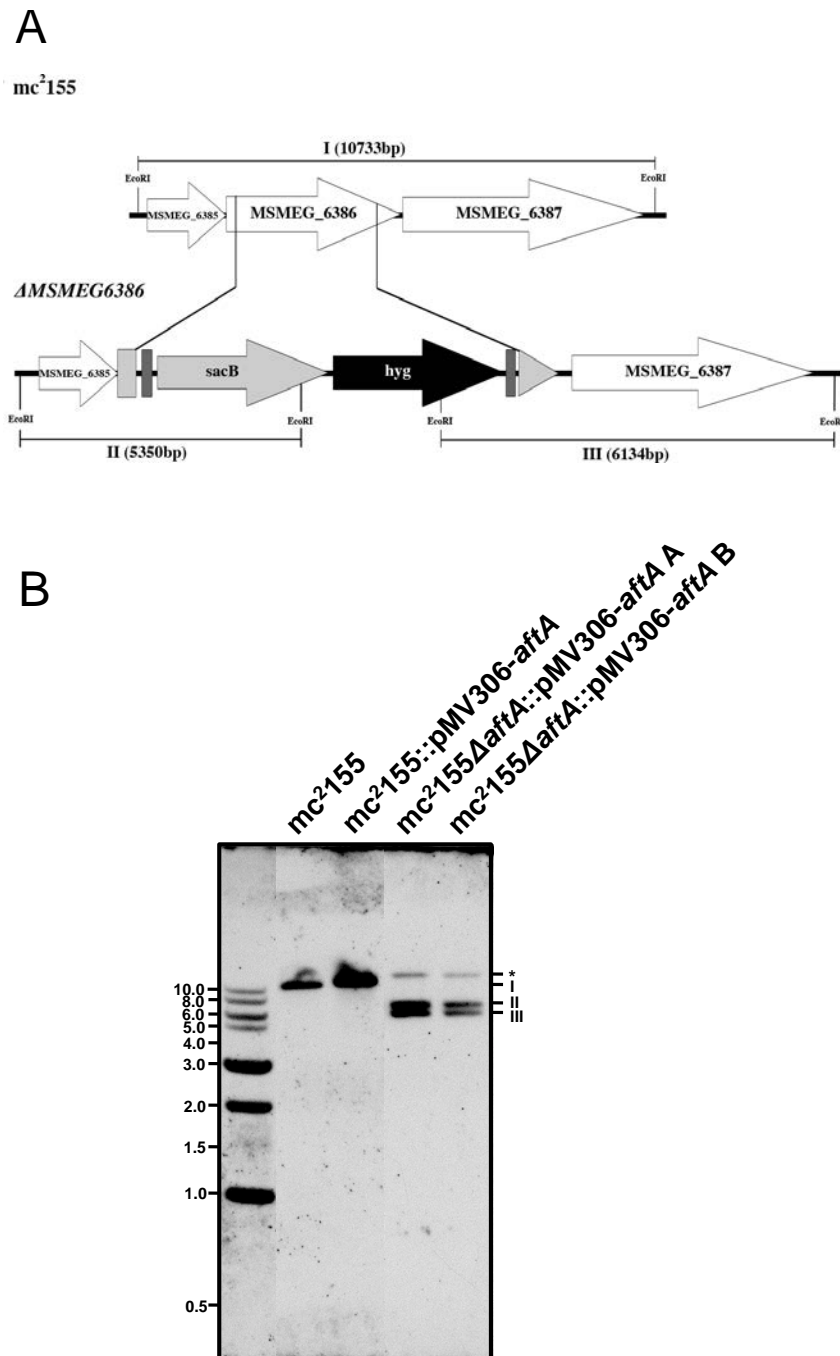


Figure 2 Generation of conditional *M. smegmatis*Δ*aftA*::pMV306-*aftA* mutant. **A** The map of the *aftA* (*MSMEG_6386*) region in the parental *M. smegmatis* strain and its corresponding region in the *M. smegmatis*Δ*aftA*::pMV306-*aftA* mutant; *hyg* – hygromycin resistance gene from *Streptomyces hygroscopicus*, *sacB* – sucrose counter selectable gene from *Bacillus subtilis*. EcoRI-digested bands expected in a Southern blot are depicted as I, II and III. **B** Southern blot analysis of EcoRI-digested genomic DNA from the wild type *M. smegmatis*, merodiploid and conditional mutant strains with expected bands I, II and III. The asterisk indicates a band appearing as a result of a second *aftA* copy integration into mycobacterial chromosome.

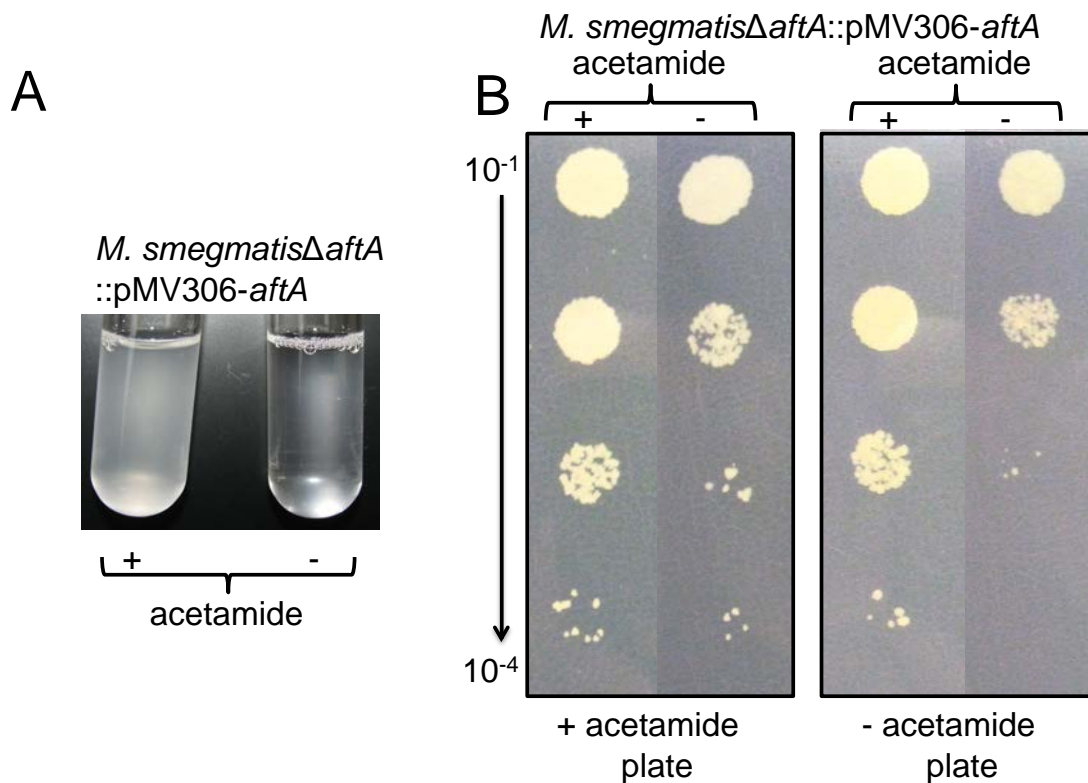


Figure 3 Growth of *M. smegmatis* Δ *aftA* pMV306-*aftA* with and without acetamide. **A** *M. smegmatis* strains to be tested were grown in minimal media supplemented with 0.05 % Tween-80 and 0.2 % acetamide to an OD₆₀₀ of 0.5. Cells were washed twice to remove traces of acetamide and resuspended in the original volume of media supplemented with 0.05 % Tween-80. This cell culture was used as a 20 % inoculum in minimal media and grown for 12 hours to deplete intracellular AftA. The depleted sub-culture served as inoculum (5 %) for cultures grown with or without 0.2 % acetamide for at least 24 hours. **B** Cultures grown with and without acetamide and 10 μ l of 10-fold serial dilutions of cultures were spotted onto \pm acetamide plates and incubated at 37 °C for 3 days.

3.3.1.2 Overexpression and purification of recombinant AftA^{CT}

The hydrophilic C-terminal domain encoded by *aftA*^{CT} of *M. tuberculosis* was cloned into pET expression vectors and transformed into *E. coli* BL21 (DE3), C43 (DE3), Tuner (DE3), and Rosetta (DE3) competent cells. SDS-PAGE and Western blot analyses of bacterial cultures were used to assess expression and solubility of recombinant proteins on a small scale (Figure 4-5). Recombinant AftA^{CT} was expressed in all four *E. coli* cell

lines, but the best expression of soluble protein was observed using *E. coli* BL21 (DE3) cells at either 16 °C or 37 °C using 1mM IPTG or *E. coli* C43 (DE3) at 16 °C using 1 mM IPTG (Figure 4). Notably, production of AftA^{CT} was detected in the culture samples without the addition of the inducer IPTG. This is probably due to incomplete repression of the promoter, commonly seen with *lac* promoters that leads to low levels of expression without induction. Since expression of AftA^{CT} at 16 °C using 1 mM IPTG yielded plenty of soluble recombinant protein, these conditions were chosen for large scale protein expression.

Since expression in BL21 (DE3) cells resulted in a higher amount of protein, this condition was replicated for large protein expression. AftA^{CT} was purified employing HiTrap Ni²⁺-NTA agarose column and during purification eluted at 50 mM - 1 M imidazole fractions (Figure 6). The purity of the AftA^{CT} was determined by a large single band running at ~22 kDa on the SDS-PAGE gel. The 250 mM - 1M imidazole fractions containing pure AftA^{CT} were pooled together and dialysed against Tris buffer (50 mM Tris pH 7.5, 300 mM NaCl). A considerable amount of AftA^{CT} was lost due to precipitation throughout dialysis and only a low concentration of 4 mg/ml was obtained. Thus, different buffers were explored to dialyse purified AftA^{CT} including Tris, HEPES (50 mM HEPES pH 7.5, 300 mM NaCl), and MES (50 mM MES pH 6.5, 300 mM NaCl), however, a high degree of precipitation of the recombinant protein was still observed. As a result, higher volumes of bacterial culture have been used to obtain protein concentrations suitable for protein crystallisation.

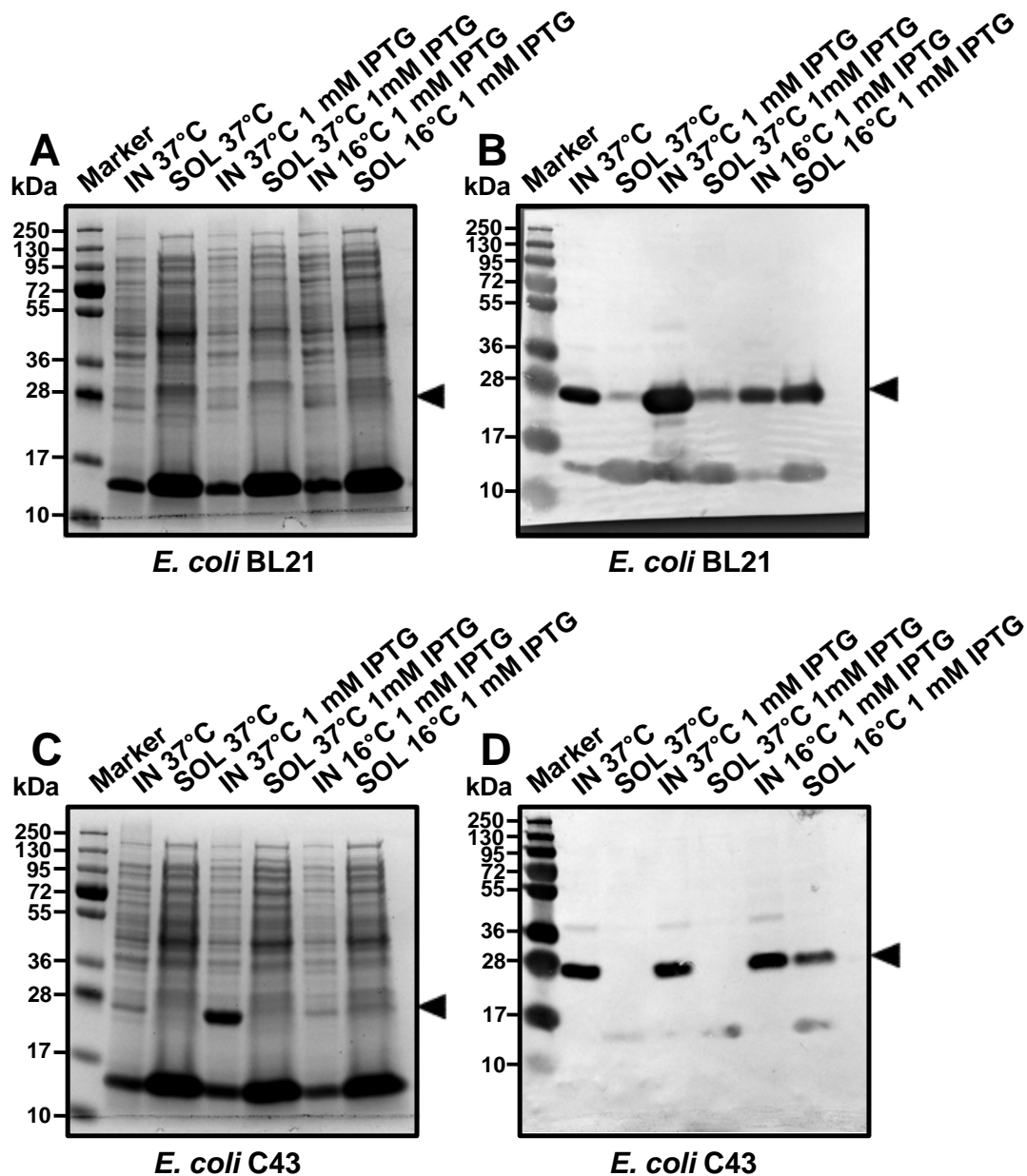


Figure 4 SDS-PAGE (A and C) and Western blot (B and D) analysis of AftA C-terminus from *M. tuberculosis* overexpressed in *E. coli* BL21 (DE3) and C43 (DE3) cells. *E. coli* cells harbouring pET16b-*aftA*^{CT} plasmid were grown at 37°C for 1 h. Expression of AftA C-terminus was induced with 1 mM IPTG and cells were further cultivated at either 37 °C for 4 hours or at 16 °C overnight. Subsequently, cells were harvested, lysed and separated into two fractions: crude cell lysate (IN) and clarified cell lysate (SOL). Samples were analysed by 4-15 % SDS-PAGE gel and stained with Coomassie Blue; the His-tagged AftA^{CT} was detected by a Western blot using a monoclonal anti-His₆ tag antibody. The arrows indicate the overexpressed AftA^{CT} (22 kDa).

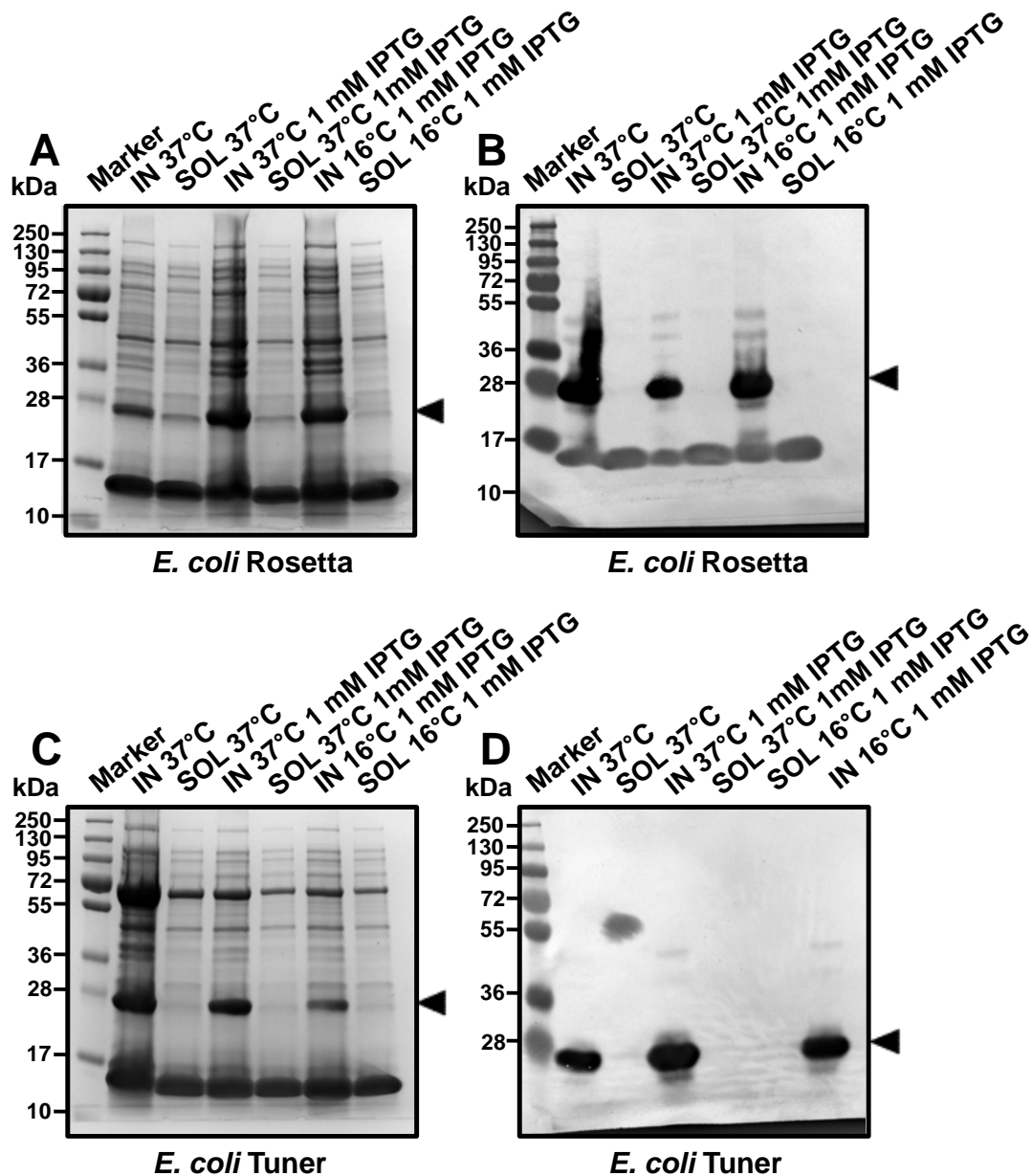


Figure 5 SDS-PAGE (A and C) and Western blot (B and D) analysis of AftA C-terminus from *M. tuberculosis* overexpressed in *E. coli* Rosetta (DE3) pLysS and Tuner (DE3) cells. *E. coli* cells harbouring pET16b-aftA^{CT} plasmid were grown at 37°C for 1 h. Expression of AftA C-terminus was induced with 1 mM IPTG and cells were further cultivated at either 37 °C for 4 hours or at 16 °C overnight. Subsequently, cells were harvested, lysed and separated into two fractions: crude cell lysate (IN) and clarified cell lysate (SOL). Samples were analysed by 4-15 % SDS-PAGE gel and stained with Coomassie Blue; the His-tagged AftA^{CT} was detected by a Western blot using a monoclonal anti-His₆ tag antibody. The arrows indicate the overexpressed AftA^{CT} (22 kDa).

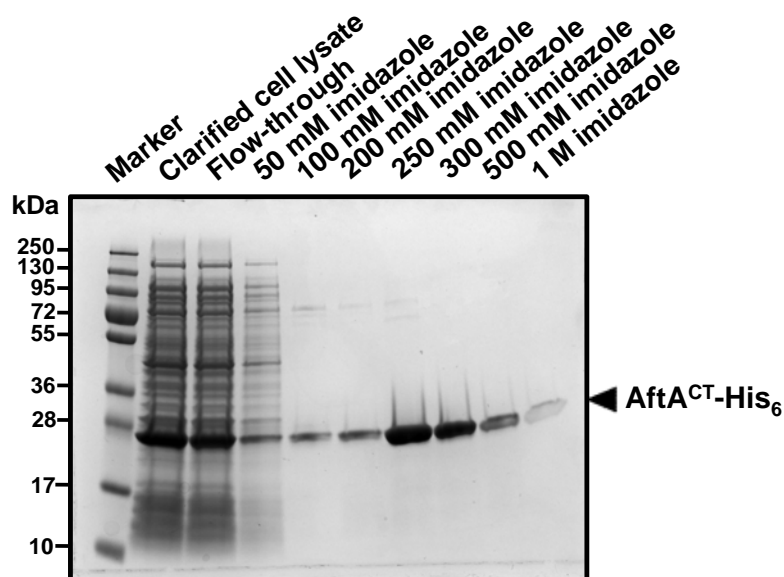


Figure 6 SDS-PAGE analysis of *M. tuberculosis* AftA^{CT} purification. *E. coli* BL21 (DE3) cells harbouring pET16b-*aftA*^{CT} were grown at 37 °C until an OD₆₀₀ 0.5, 1 mM IPTG was used to induce protein expression, and cells were further cultivated for 12 hours at 16 °C. The C-terminus of AftA (22 kDa) was purified using HiTrap Ni²⁺-NTA agarose column *via* a stepwise gradient of 50–1000 mM imidazole. Fractions were analysed by 12% SDS-PAGE and visualised by Coomassie Blue staining.

3.3.1.3 Differential scanning fluorimetry profile of AftA^{CT}

Since AftA^{CT} demonstrated a tendency to precipitate in a number of buffers it was dialysed against, a high throughput thermostability screen employing differential scanning fluorimetry (DSF) was performed. Thermostability assessment identifies solution conditions that facilitate protein conformational integrity and resistance to temperature stress (Niesen *et al.*, 2007). Numerous studies have demonstrated that buffer conditions or ligands that stabilise proteins, reduce their tendency to unfold and aggregate, and improve protein purification and subsequent crystallisation (Murphy, 2001, Elleby *et al.*, 2004, Vedadi *et al.*, 2006, Niesen *et al.*, 2007). DSF measures the fluorescence of an external probe (e.g. SYPRO Orange fluorescent dye), which is

sensitive to hydrophobic areas of the protein. Fluorescence is negligible in solution lacking the probe, however, it increases significantly in the presence of unfolded protein as exposure of protein's inner hydrophobic regions allows the dye to bind. A total of 48 different buffer conditions covering a pH range from 4 to 9 and two different salt concentrations (190 mM and 340 mM NaCl) were screened employing DSF. The typical DSF profile of AftA^{CT} is shown in Figure 7. Effects of pH and buffer composition on the thermostability of AftA^{CT} are also presented as differences in melting temperature (ΔT_m), where T_m obtained in the original buffer (20 mM sodium acetate pH 4.2, 200 mM NaCl) was subtracted from T_m with different buffer content (Figure 8). In more than 70 % of the cases, a condition was identified that stabilised AftA^{CT} compared to the original buffer. Moreover, in the majority of cases higher concentration of salt resulted in an increased stabilisation of AftA^{CT}, however buffers containing sodium acetate, sodium citrate and MES ranging from pH 4 to 6 demonstrated a higher stabilisation effect on AftA^{CT} with lower salt concentration. In several instances, the destabilising effect was more pronounced in conditions containing HEPES and Tris buffers (HEPES pH 7.0, HEPES pH 7.5 and Tris pH 8.6), thus resulting in negative ΔT_m values. Overall, the most significant thermal stabilisation was observed in buffers containing sodium acetate, which was independent of both pH and salt concentration. Therefore, 20 mM sodium acetate pH 6.0, 200 mM NaCl buffer was chosen as a stabilising solution for the purification, concentration and crystallisation of AftA^{CT}.

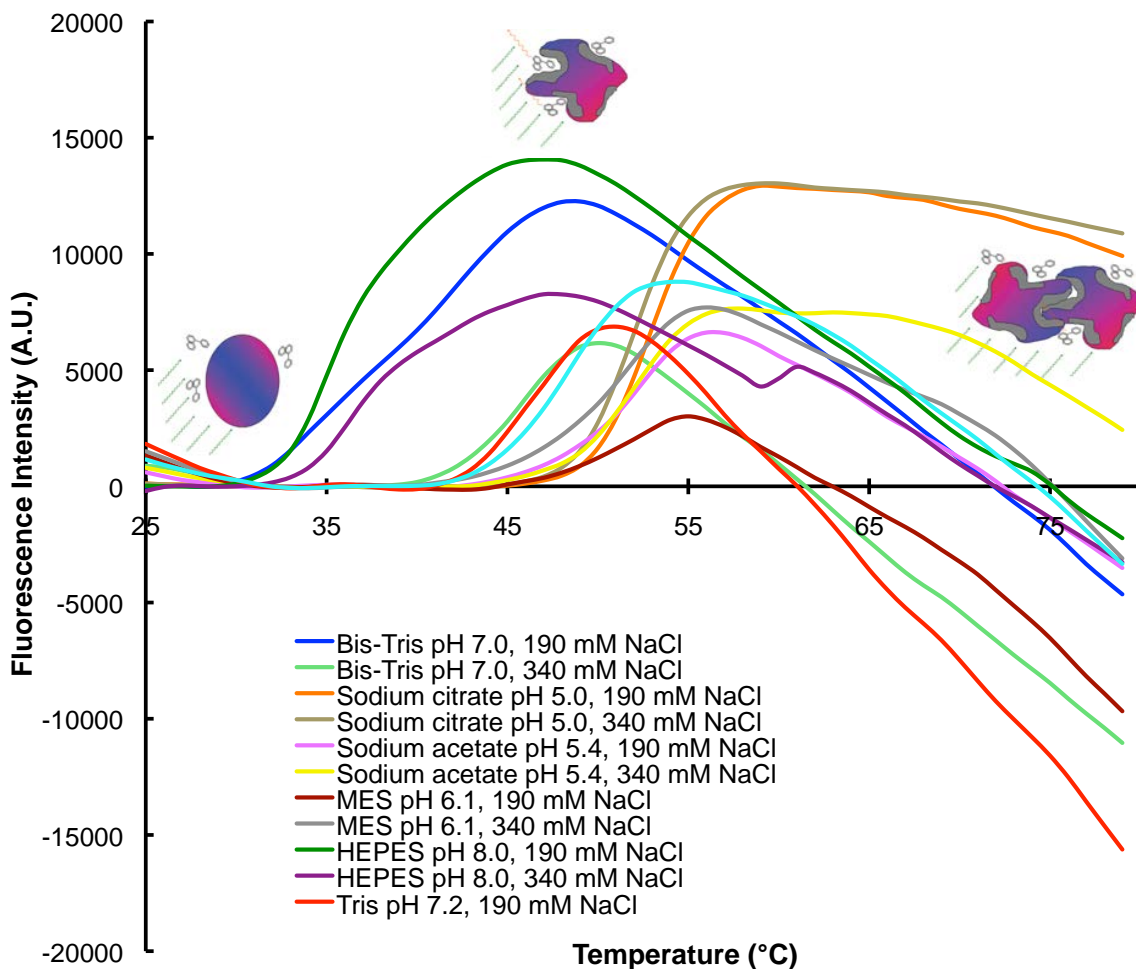


Figure 7 Fluorescence of SYPRO Orange acquired by differential scanning fluorimetry (DSF) in the presence of AftA^{CT} domain as a function of temperature. SYPRO Orange fluorescent dye was diluted into twenty-four buffers differing in their salt concentration and pH to a final dilution of 1:1000. The final concentration of all buffers was 100 mM, while the AftA^{CT} concentration was 100 ng/ml. The detailed DSF procedure is described in Materials and Methods 3.2.10. The SYPRO Orange is depicted here as a three-ring aromatic molecule, whereas AftA^{CT} is represented by a purple sphere. A basic fluorescence intensity is excited by light of 492 nm and is depicted as green strokes. During protein unfolding, hydrophobic areas of the protein are exposed, thus allowing fluorescent dye to bind. As a result, a strong fluorescent light of 610 nm is emitted (orange strokes). Finally, increasing temperature leads to gradual precipitation and aggregation of the protein resulting in a decrease in the fluorescence signal.

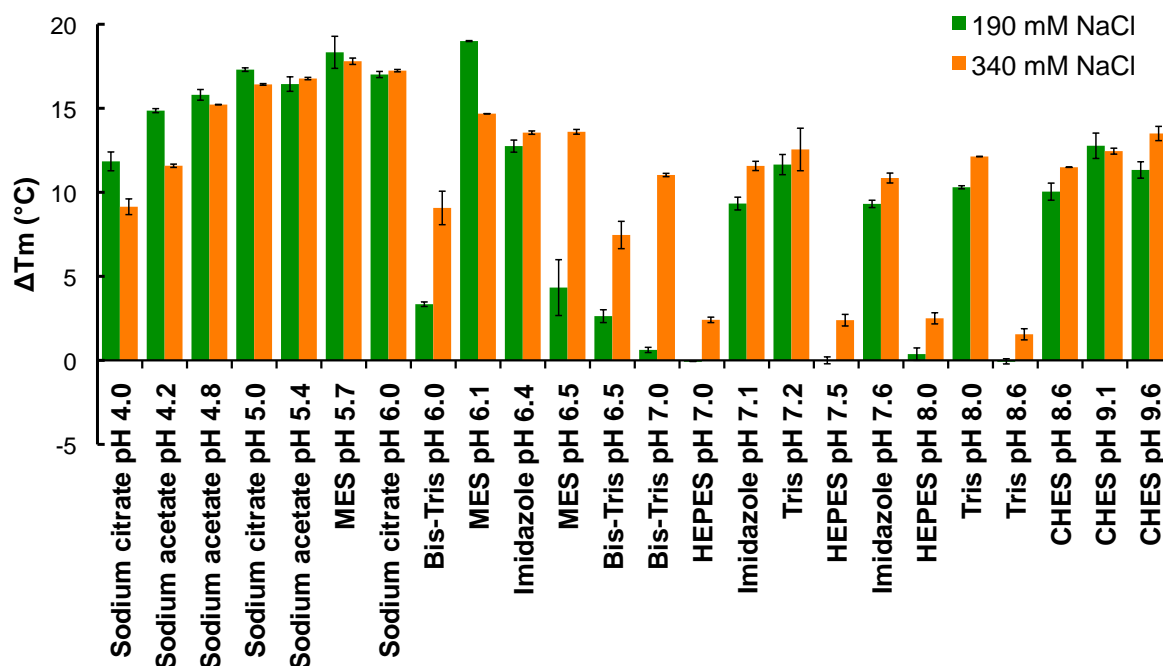


Figure 8 The effects of different buffers and pH expressed as ΔT_m for AftA^{CT} domain. Differential scanning fluorimetry was employed as described in Materials and Methods 3.2.10 with AftA^{CT} concentration of 100 ng/ml. Formulations included twenty-four conditions as indicated in the figure labels. The ΔT_m values were obtained by subtracting T_m value obtained in the reference blank (AftA^{CT} of 100 ng/ml in 100 mM sodium acetate pH 4.2, 200 mM NaCl) from those with different contents. The results are expressed as means \pm S.D. of two independent experiments.

3.3.1.4 Crystallisation trials of *M. tuberculosis* AftA^{CT}

Purified AftA^{CT} was dialysed against Tris buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl) or sodium acetate buffer (20 mM sodium acetate pH 6.0, 200 mM NaCl) and concentrated to 15 mg/ml and 30 mg/ml. The Molecular Dimensions optimised sparse matrix screens Structure screen I & IITM, Pact *premier*TM, JCSG-*plus*TM, MorpheusTM, MIDASTM as well as Hampton Research Crystal screen I & IITM were used to identify conditions that were most suitable for crystal growth. The sitting drop vapour diffusion

format was used in a 96-well plate. Unfortunately, none of the screens that were set up generated crystals.

3.3.1.5 Overexpression studies of recombinant AftA from *M. tuberculosis* and *C. glutamicum*

Full-length AftA from *M. tuberculosis* and *C. glutamicum* encoded by *Rv3792* and *NCgl0185*, respectively, were cloned into both pMSX and pEKEx5 expression vectors and transformed into *C. glutamicum* ATCC 13032. SDS-PAGE and Western blot analyses of bacterial cultures were used to assess expression and solubility of recombinant proteins employing four different conditions: 16 °C for 12 hours using 0.1 mM IPTG, 16 °C for 12 hours using 1 mM IPTG, 30 °C for 4 hours using 0.1 mM IPTG and 30 °C for 4 hours using 1 mM IPTG. Appreciable amounts of AftA in soluble form were generated at 16 °C for 12 hours using 1 mM IPTG, 30 °C for 4 hours using 0.1 mM IPTG and 30 °C for 4 hours using 1 mM IPTG (Figure 9). Since a greater quantity of protein seemed to be expressed at 16 °C using 0.1 mM IPTG, this condition was used for further experiments. Notably, a ladder effect was observed in the samples overexpressing *C. glutamicum* AftA at approximately 150 kDa and above. It is tempting to speculate that this may be the evidence of a multi-protein complex formation, where *C. glutamicum* AftA is forming a dimer or tetramer.

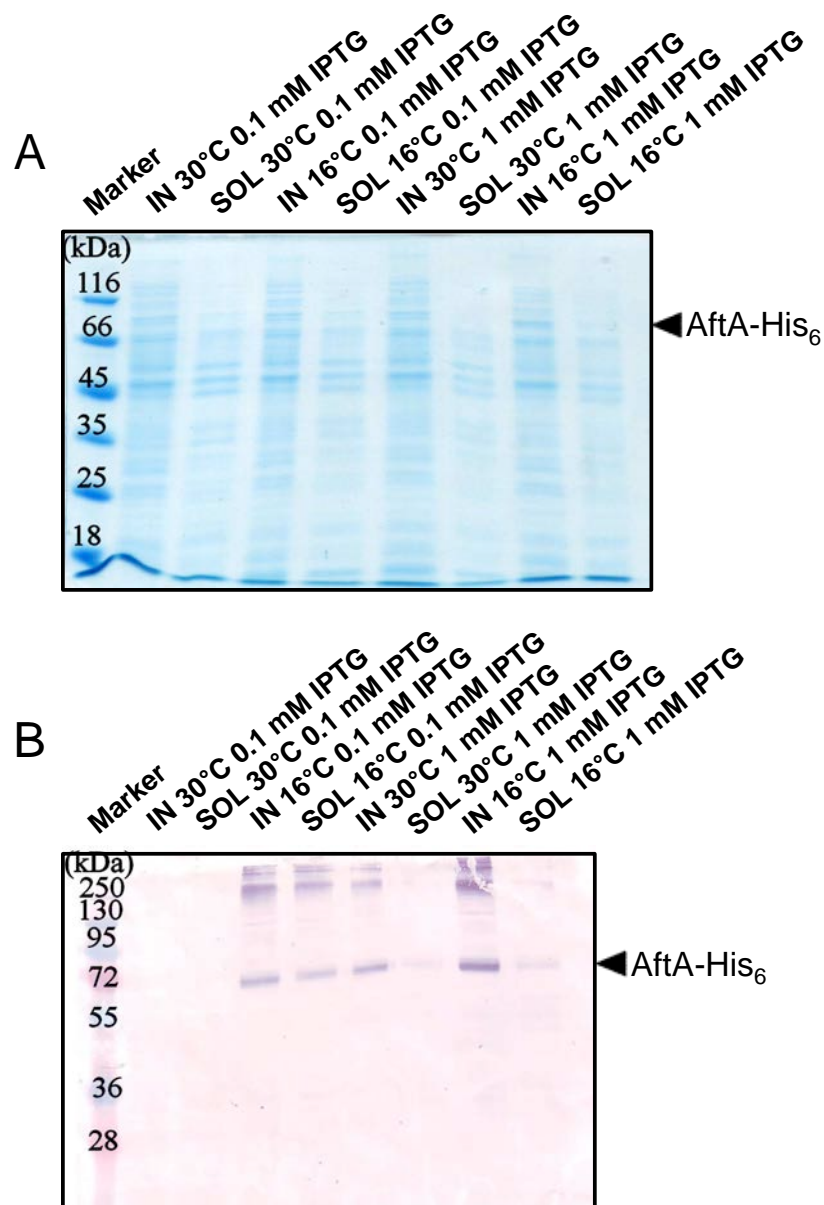


Figure 9 SDS-PAGE (A) and Western blot (B) analysis of full-length AftA from *C. glutamicum* overexpressed in *C. glutamicum* ATCC 13032 cells. *C. glutamicum* cells harbouring pMSX-*aftA* plasmid were grown at 30 °C for until an early growth phase. Expression of AftA was induced using either 0.1 mM or 1 mM IPTG and cells were further cultivated at either 30 °C for 4 hours or at 16 °C overnight. Subsequently, cells were harvested, lysed and separated into two fractions: crude cell lysate (IN) and clarified cell lysate (SOL). Samples were analysed by 12 % SDS-PAGE gel and stained with Coomassie Blue; the His-tagged AftA was detected by a Western blot using a monoclonal anti-His₆ tag antibody. The arrows indicate the overexpressed AftA (74 kDa).

3.3.1.6 Complementation of *C. glutamicum* Δ *aftA* mutant with *C. glutamicum* AftA to verify its intact enzyme function

Since the expression vector in use – pMSX – carries a C-terminal His₆ tag sequence and the C-terminal region of AftA may play a crucial role in enzyme function, we complemented *C. glutamicum* strain lacking chromosomal *aftA* with recombinant AftA-His₆. The pMSX-*aftA* construct was electroporated into *C. glutamicum* Δ *aftA* and overexpressed at 16 °C for 12 hours using either 0.1 mM IPTG or 1 mM IPTG. SDS-PAGE and Western blot analysis confirmed that *C. glutamicum* AftA was successfully expressed under both chosen conditions in the *C. glutamicum* Δ *aftA* mutant (Figure 10). Importantly, empty pMSX plasmid was electroporated into *C. glutamicum* Δ *aftA* and served as a negative control (Figure 10). To reaffirm that recombinant *C. glutamicum* AftA is fully functioning, cell walls from wild type *C. glutamicum*, *C. glutamicum* Δ *aftA*, and *C. glutamicum* Δ *aftA*::pMSX-*aftA* were analysed using glycosyl compositional analysis. As expected, the presence of rhamnose, arabinose and galactose was observed in alditol acetates from *C. glutamicum* mAGP (Figure 11). In contrast, alditol acetates prepared from *C. glutamicum* Δ *aftA* lacked arabinose and only contained rhamnose and galactose residues as reported previously (Alderwick *et al.*, 2006). Complementation of the *C. glutamicum* Δ *aftA* mutant with plasmid encoded *aftA* from *C. glutamicum* restored the glycosyl composition to that of wild type *C. glutamicum* (Figure 11), thus, demonstrating that recombinant *C. glutamicum* AftA is fully active and the His₆ tag does not interfere with its enzymatic function.

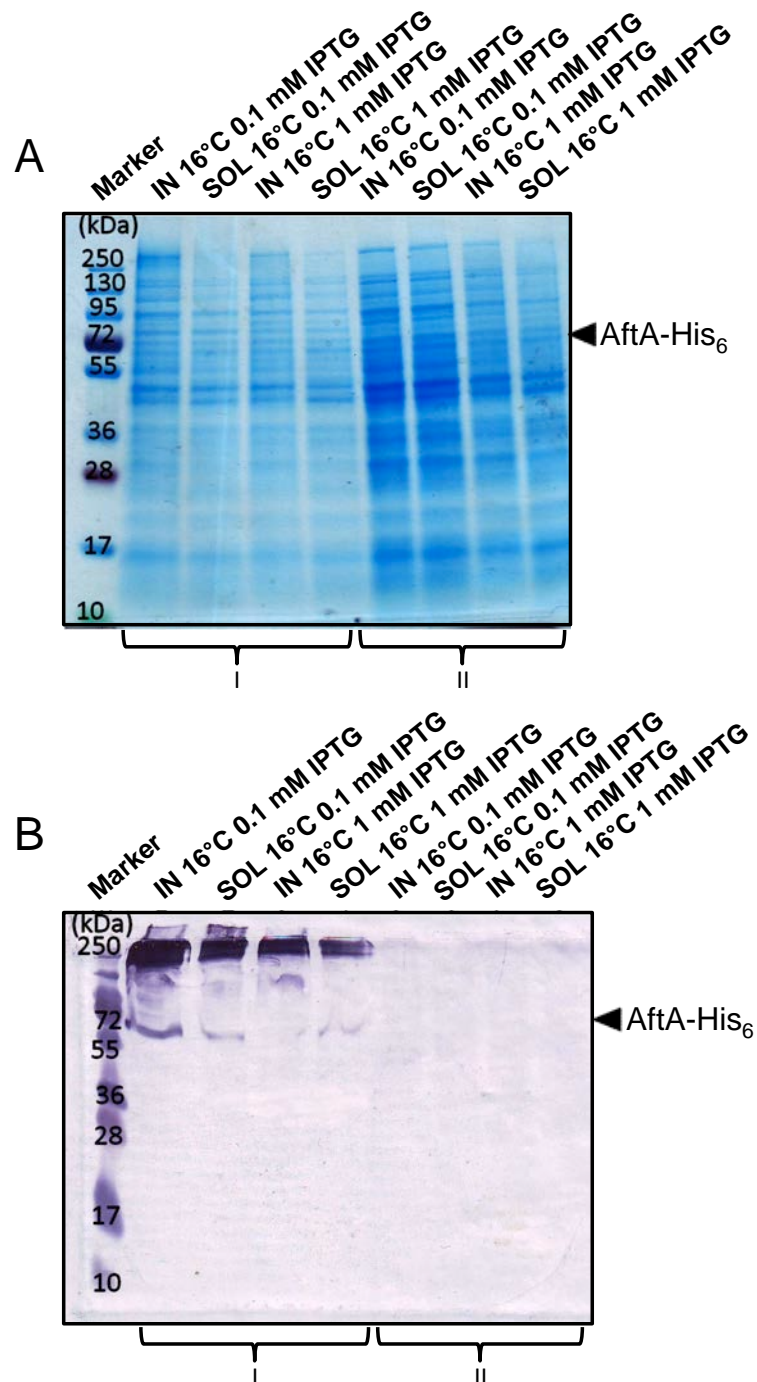


Figure 10 SDS-PAGE (A) and Western blot (B) analysis of *C. glutamicum* Δ *aftA* complemented with either pMSX-*aftA* (I) or empty pMSX plasmid (II). *C. glutamicum* cells harbouring appropriate plasmid were grown at 30 °C for until an early growth phase. Expression of AftA was induced with either 0.1 mM or 1 mM IPTG and cells were further cultivated at either 16 °C for 12 hours. Subsequently, cells were harvested, lysed and separated into two fractions: crude cell lysate (IN) and clarified cell lysate (SOL). Samples were analysed by 12 % SDS-PAGE gel and stained with Coomassie Blue; the His-tagged AftA was detected by a Western blot using a monoclonal anti-His₆ tag antibody. The arrows indicate the overexpressed AftA (74 kDa).

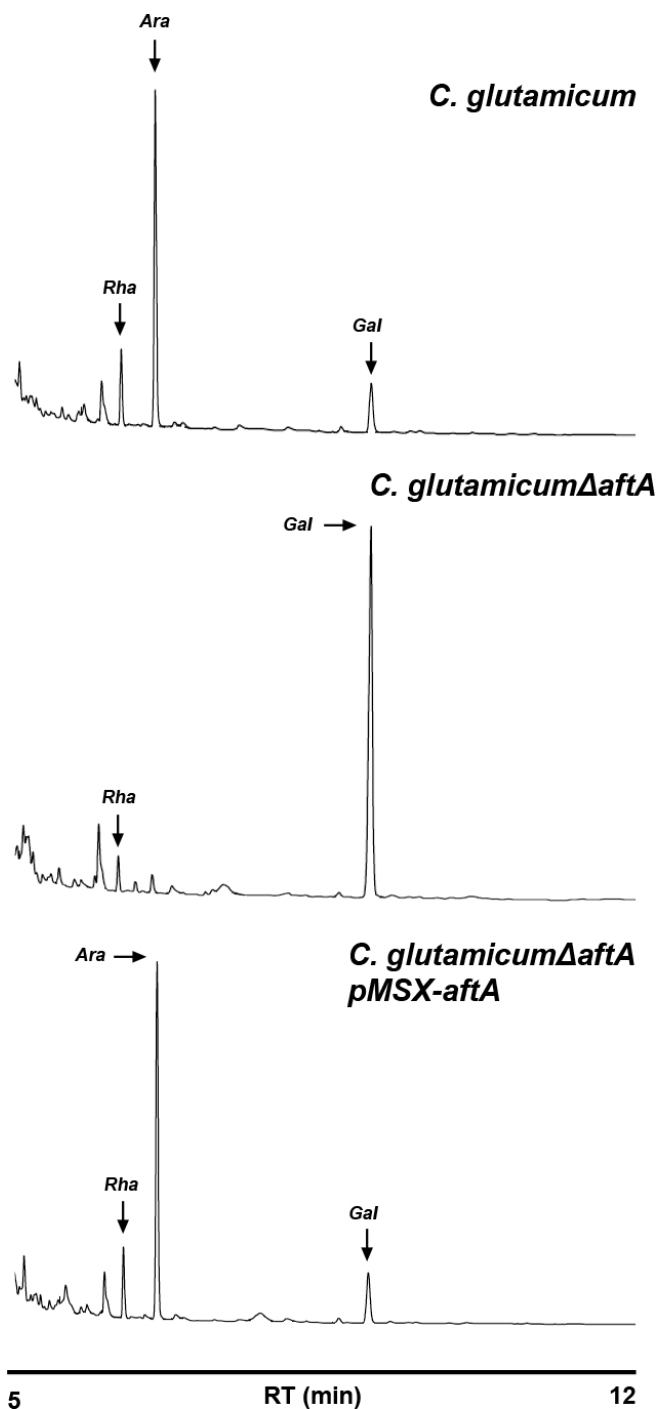


Figure 11 Glycosyl compositional analysis of cell walls of *C. glutamicum*, *C. glutamicum* Δ *aftA*, and *C. glutamicum* Δ *aftA*::*pMSX-aftA*. Samples of purified mAGP were hydrolysed with 2 M trifluoroacetic acid, reduced, per-*O*-acetylated, and subjected to GC analysis.

3.3.1.7 Purification of SMALP encapsulated AftA from *C. glutamicum*

Since AftA expression at 16 °C for 12 hours using pMSX-*aftA* at 0.1 mM IPTG resulted in the highest amount of recombinant protein, this condition was replicated for large protein expression. Cultures were used to prepare membrane fractions that were subsequently mixed with styrene maleic acid polymer and incubated at 30 °C for 12 hours to allow encapsulation of transmembrane proteins. Insoluble material was removed by centrifugation and the supernatant was subjected to HiTrap Ni²⁺-NTA agarose gravity column in order to purify SMALP encapsulated His₆ tagged AftA. Recombinant protein eluted at 50-200 mM imidazole fractions that were pooled, dialysed against 50 mM Tris-HCl pH 8.0 and concentrated to 2 mg/ml yielding for the first time a preparation of high purity full-length *C. glutamicum* AftA.

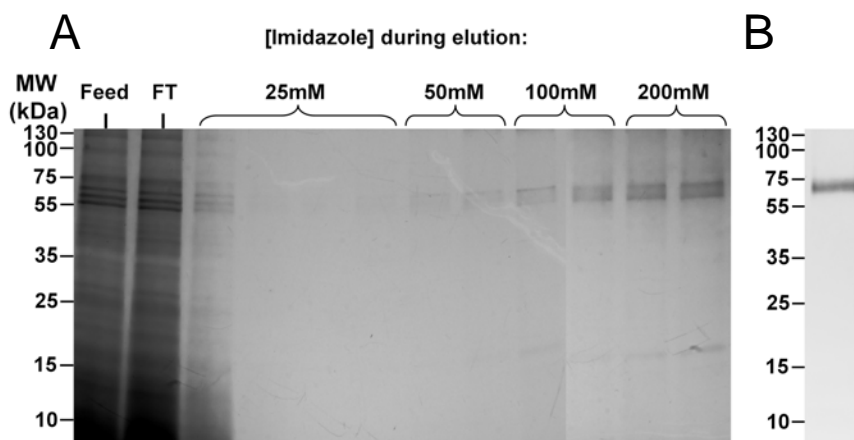


Figure 12 SDS-PAGE (A) and Western blot (B) analysis of SMALP encapsulated AftA purification. **A** *C. glutamicum* cells harbouring pMSX-*aftA* were grown at 30 °C until an OD₆₀₀ of 0.5, protein expression was induced using 0.1 mM IPTG and cells were further cultivated for 12 hours at 16 °C. Harvested *C. glutamicum*::pMSX-*aftA* cultures were used to prepare membrane fractions as described in General Materials and Methods 6.5.13. Membranes were homogenised to a final concentration of 60 mg/ml in 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 5 mM imidazole and 2.5 % (w/v) styrene maleic acid. Mixture was incubated at 30 °C for 12 hours with gentle agitation. Once formed, SMALP encapsulated AftA was purified using Ni²⁺-NTA linked agarose via a stepwise gradient of 50-200 mM imidazole. Fractions were analysed by 12% SDS-PAGE and visualised by Coomassie Blue staining. **B** The polyhistidine tagged AftA was detected by Western blotting using a monoclonal anti-His tag antibody.

3.3.1.8 Biophysical characterisation of SMALP encapsulated AftA from *C. glutamicum*

Further, we wanted to characterise the biophysical properties of AftA encapsulated within SMALPs. Firstly, we analysed sample homogeneity by analytical centrifugation, which revealed a single major peak, thus indicating stable, monodispersed species (Figure 13A). A circular dichroism (CD) spectrum was obtained for *C. glutamicum* AftA within the SMALP to estimate the structural state of the transmembrane protein in the intact native membrane environment (Figure 13B). Since no protein is required to form the nanoparticles they have negligible CD absorbance, as previously demonstrated by Knowles et al. (2009). The AftA structure was maintained within SMALPs: far-UV CD spectra show shoulders at 222 nm and 208 nm, indicating the expected α -helical structures that are characteristic of the folded state of the protein (Figure 13B). Finally, we hypothesised that the lipid environment surrounding transmembrane AftA protein would retain and possibly increase protein stability when compared to protein purification without SMALPs. Thus, we examined the thermostability of SMALP encapsulated AftA, where the folded state was monitored over time with increasing temperature (Figure 13C). Excellent stability of AftA was observed with unfolding temperature of approximately 65 °C.

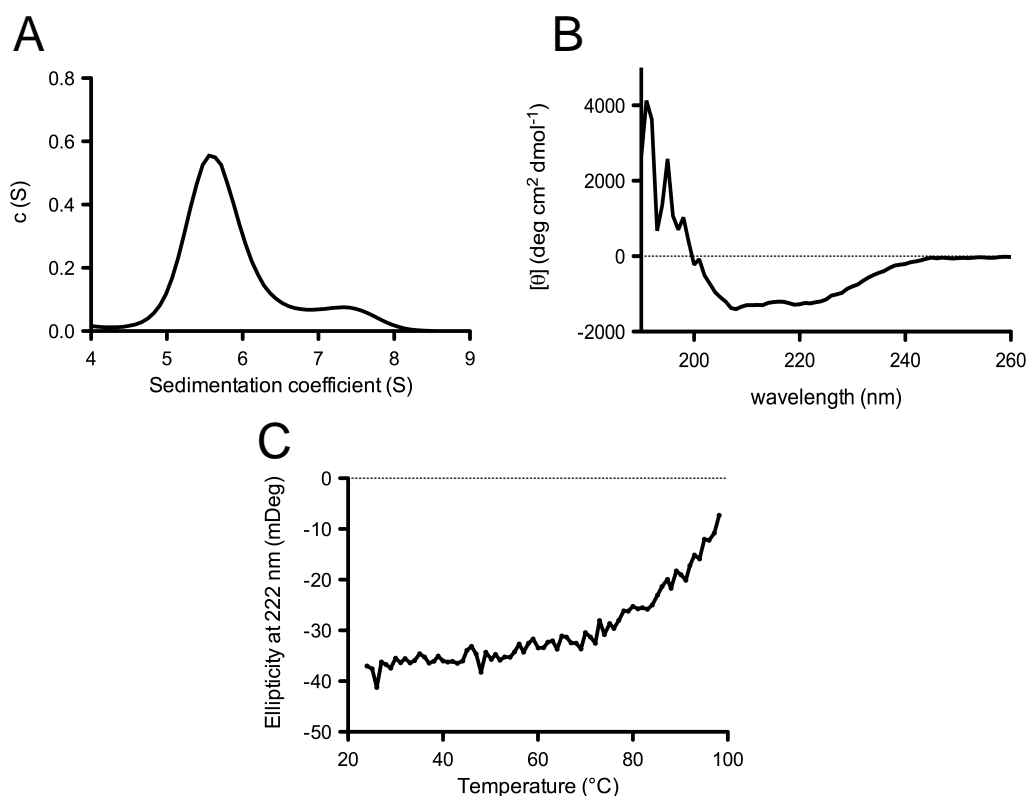


Figure 13 Biophysical characterisation of SMALP encapsulated AftA from *C. glutamicum*. **A** Analytical ultracentrifugation analysis of AftA within SMALPs. **B** Far-UV circular dichroism analysis of AftA within SMALPs. **C** Circular dichroism thermal denaturation studies of AftA within SMALPs.

3.3.1.9 Ligand binding assays of AftA from *C. glutamicum*

The binding properties of SMALP encapsulated AftA from *C. glutamicum* were examined with three synthetic substrates β -D-Galf(1 \rightarrow 5)- β -D-Galf-*O*-(CH₂)₇CH₃ (G5G), β -D-Galf(1 \rightarrow 6)- β -D-Galf-*O*-(CH₂)₇CH₃ (G6G) and (β -D-Galf(1 \rightarrow 6)- β -D-Galf(1 \rightarrow 5))6- β -D-Galf(1 \rightarrow 4)- α -L-Rhap-GlcNAc-*O*-(CH₂)₇CH₃ (Acc13). Nonlinear regression analysis of $\Delta F_{\text{emission}}$ against G5G showed a typical saturation ligand-binding curve with a calculated K_d of 1.372 mM (Figure 14A). The same analysis was performed where G6G acceptor was used as a substrate for AftA and demonstrated a higher specificity with K_d

of 0.9959 mM than compared to G5G (Figure 14B). This is unsurprising, since it was previously shown that AftA can transfer an arabinose residue to the C-5 position of the internal 6-linked galactose and, therefore, would bind the G6G neoglycolipid acceptor more strongly. Finally, a much longer Acc13 acceptor was used as a substrate for *C. glutamicum* AftA enzyme and resulted in a saturation ligand-binding curve with K_d of 7.314 μ M (Figure 14C). The data show that AftA ArafT binds Acc13 with higher affinity than compared to both G5G and G6G.

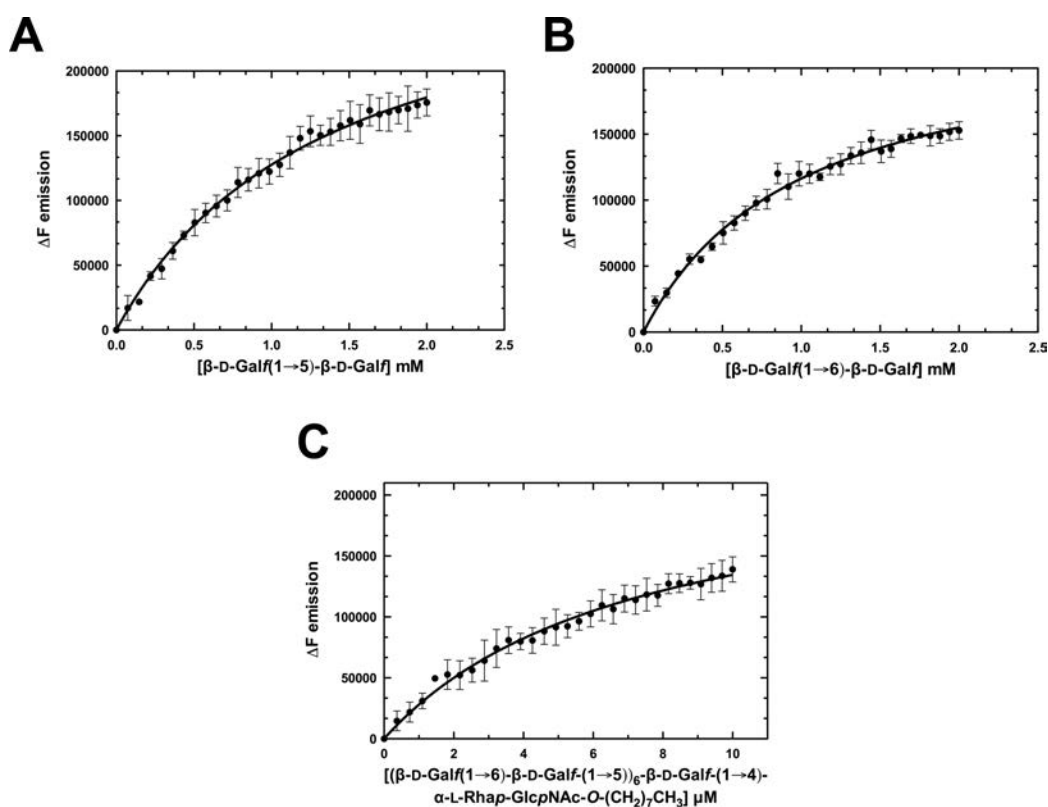


Figure 14 Binding kinetics of SMALP encapsulated AftA from *C. glutamicum* to different substrates. Ligand binding was analysed by intrinsic tryptophan fluorescence spectroscopy employing three substrates: **A** β -D-Galf-(1 \rightarrow 5)- β -D-Galf-(CH₂)₇CH₃; **B** β -D-Galf-(1 \rightarrow 6)- β -D-Galf-(CH₂)₇CH₃; **C** $(\beta$ -D-Galf-(1 \rightarrow 6)- β -D-Galf-(1 \rightarrow 5))₆- β -D-Galf-(1 \rightarrow 4)- α -L-Rhap-GlcNAc-O-(CH₂)₇CH₃. The excitation wavelength was set to 283 nm and the fluorescence emission was recorded between 300-380 nm for each ligand aliquot added to a total of 300 μ l solution containing 2 μ M SMALP encapsulated *C. glutamicum* AftA in 50 mM Tris-HCl pH 8.0. The change in fluorescence emission intensity was plotted against ligand concentration (3 independent experiments) and fitted to the saturation binding equation using GraphPad Prism Software.

3.3.2 Molecular and biochemical characterisation of AftB

3.3.2.1 Evidence suggesting that AftB from *M. smegmatis* is an essential gene

We constructed a knockout phage phMSMEG6400 designed to replace the chromosomal *aftB* in *M. smegmatis* with a hygromycin resistance cassette. Several attempts to obtain an *aftB* null mutant in *M. smegmatis* were unsuccessful suggesting that *aftB*, unlike its homologue in *C. glutamicum*, is an essential gene (Seidel *et al.*, 2007a). We then employed CESTET to generate an *aftB* conditional deletion mutant by transducing a meridiplod strain containing a second, inducible copy of *aftB* (Bhatt and Jacobs, 2009). The correct replacement of the native chromosomal copy of *aftB* with a hygromycin cassette in the transductants was confirmed by Southern blot analysis (Figure 15).

In order to confirm *aftB* essentiality and study the fate of mycobacterial cells depleted of AftB we monitored the growth of *M. smegmatis* Δ *aftB*::*pMV306-aftB* in liquid medium with and without acetamide (second copy of *aftB* is on and off, respectively) over 28 hours (Figure 16). While the mutant strain grew normally in the presence of acetamide, the culture without the inducer showed a decrease in OD₆₀₀ with time, resulting in a clearly lysed culture after 28 hours of incubation (Figure 16). Since growth of the conditional *aftB* mutant is dependent on the acetamide inducible expression of the second recombinant copy of *aftB*, cell lysis in the culture lacking acetamide demonstrated the essentiality of *aftB* gene to *M. smegmatis* viability.

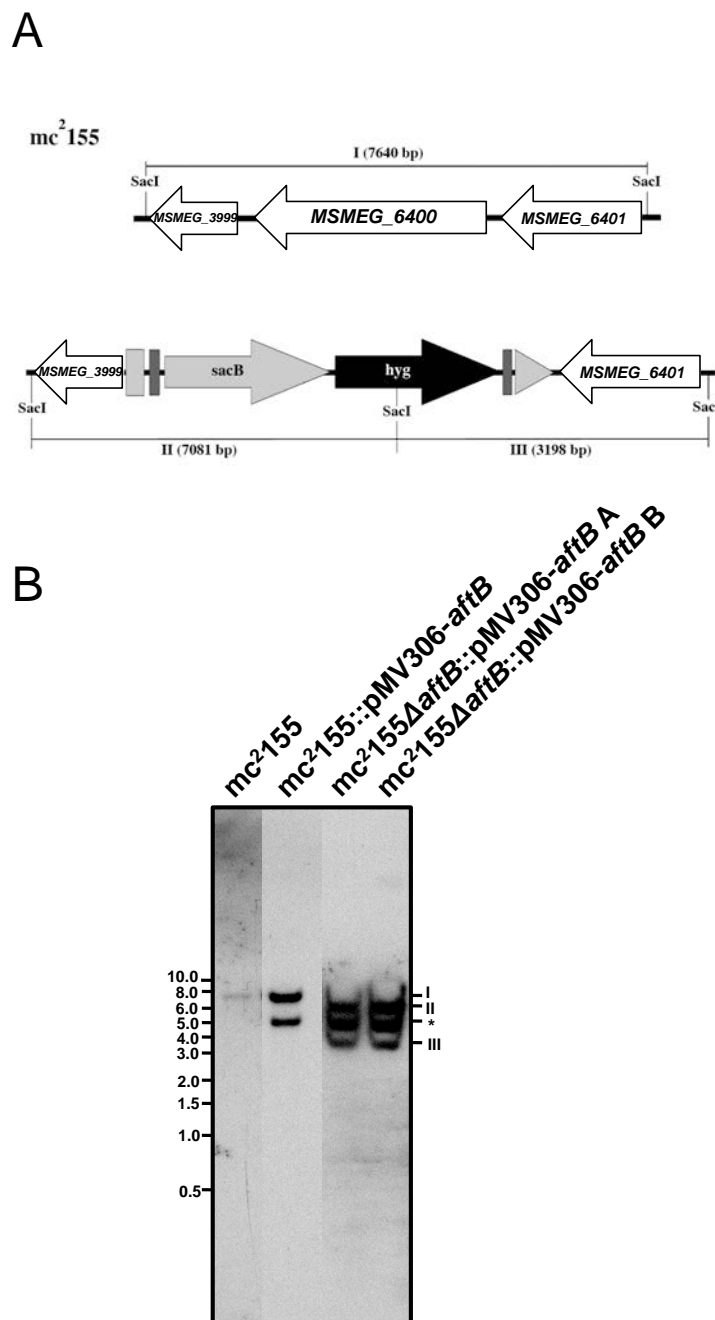


Figure 15 Generation of conditional *M. smegmatis*Δ*aftB*::*pMV306-aftB* mutant. **A** The map of the *aftB* locus (*MSMEG6400*) region in the parental *M. smegmatis* strain and its corresponding region in the *M. smegmatis*Δ*aftB*::*pMV306-aftB* mutant; *hyg* – hygromycin resistance gene from *Streptomyces hygroscopicus*, *sacB* – sucrose counter selectable gene from *Bacillus subtilis*. SacI-digested bands expected in a Southern blot are indicated as I, II and III. **B** The Southern blot of SacI-digested genomic DNA from from the wild type *M. smegmatis*, meridiplod and conditional mutant strains with expected bands I, II and III. The asterisk indicates a band appearing as a result of a second *aftB* copy integration into mycobacterial chromosome.

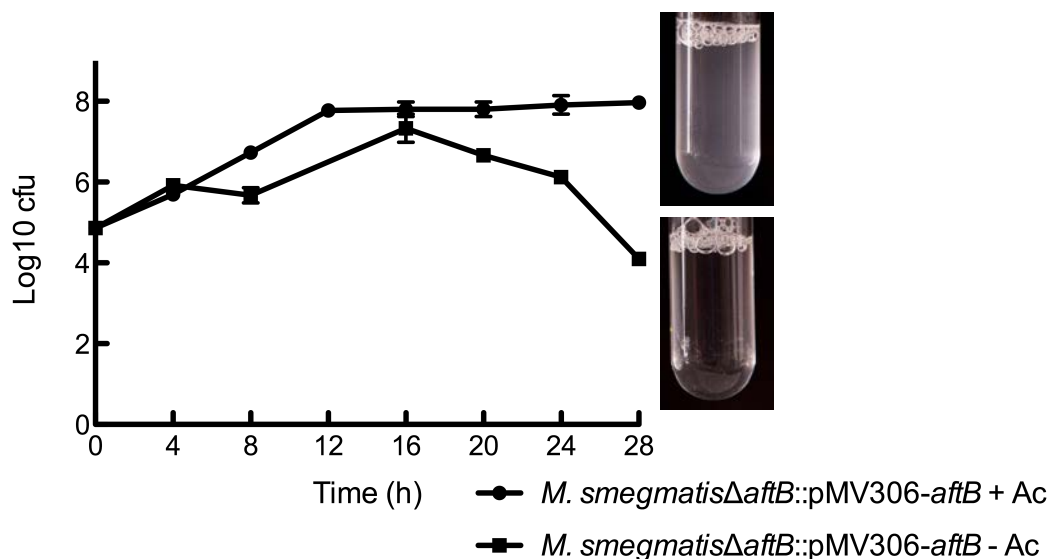


Figure 16 Growth of *M. smegmatis*Δ*aftB*::pMV306-*aftB* in liquid minimal medium with and without acetamide (Ac). *M. smegmatis* strains to be tested were grown in minimal media supplemented with 0.05 % Tween-80 and 0.2 % acetamide to an OD₆₀₀ of 0.5. Cells were washed twice to remove traces of acetamide and resuspended in the original volume of media supplemented with 0.05 % Tween-80. This cell culture was used as a 20 % inoculum in minimal media and grown for 12 hours to deplete intracellular AftB. The depleted sub-culture served as inoculum (5 %) for cultures grown with or without 0.2 % acetamide for at least 24 hours.

3.3.2.2 Lipid characterisation of the *M. smegmatis*Δ*aftB*::pMV306-*aftB* conditional mutant

To study the function of mycobacterial *aftB* deletion, its cellular composition was analysed. Firstly, the wild type *M. smegmatis*, merodiploid strain and *M. smegmatis*Δ*aftB*::pMV306-*aftB* grown with and without acetamide and analysed for AG esterified mycolic acids and cell wall associated lipids, from an equivalent starting amount of biomass for each strain due to differences in growth rate. As expected, wild type *M. smegmatis*, *M. smegmatis*::pMV306-*aftB* and the conditional mutant grown with acetamide exhibited a typical profile of cell wall bound α -, α' - and epoxy-mycolic acid

methyl esters (MAMES) (Figure 17B). In contrast, *M. smegmatis* Δ *aftB*::pMV306-*aftB* cells cultivated lacking acetamide demonstrated drastically reduced amounts of all three types of mycobacterial MAMES (Figure 17B). Analysis of the cell wall associated lipids highlighted an apparent increase in TMMs and TDMs in the mutant strain depleted of AftB (Figure 17A). *M. smegmatis* strains producing AftB exhibited the known free lipid profile containing TDMs (on average approximately 685 CNT/mm²) and TMMs (337 CNT/mm²). In contrast, *M. smegmatis* Δ *aftB*::pMV306-*aftB* grown without acetamide included a significant 1.5 fold increase in TMMs (563 CNT/mm²) and a smaller 1.3 fold increase in TDMs (935 CNT/mm²). These lipid profiles compared well with the data observed in the *C. glutamicum* Δ *aftB* mutant and suggest that AftB is involved in key aspects of arabinan biosynthesis in AG.

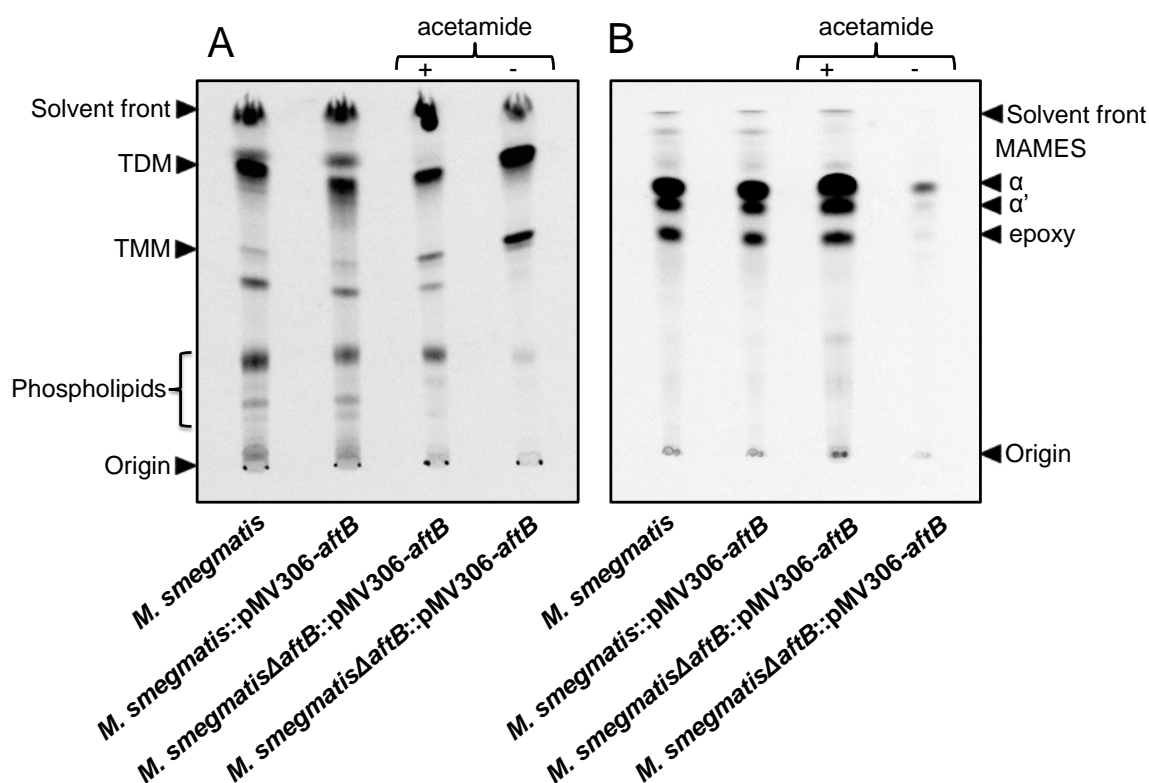


Figure 17 Analysis of cell wall lipids from *M. smegmatis*, merodiploid and *M. smegmatis* Δ *aftB*::pMV306-*aftB* strains grown with and without acetamide. **A.** Analysis of cell wall associated [14 C]-lipids from *M. smegmatis*, merodiploid and *M. smegmatis* Δ *aftB*::pMV306-*aftB* strains grown with and without acetamide. Lipids were extracted from cells by a series of organic washes as described in General Materials and Methods 6.5.14. Equivalent aliquots (25,000 cpm) from each strain were subjected to TLC using silica gel plates (5725 silica gel 60F254, Merck) developed in chloroform/methanol/water (60:16:2, v/v/v) and visualised by autoradiography using Kodak BioMax MR films. **B.** Analysis of cell wall bound [14 C]-lipids from *M. smegmatis*, merodiploid and *M. smegmatis* Δ *aftB*::pMV306-*aftB* strains grown with and without acetamide. An equivalent amount of cells (100 mg) were delipidated using three extractions of chloroform/methanol/water (10:10:3, v/v/v) at 50 °C for 4 hours, and bound mycolic acids were released by addition of *tetra*-butyl ammonium hydroxide at 100 °C for 12 hours and subsequently methylated. An equal amount of sample from each strain was subjected to TLC using silica gel plates (5725 silica gel 60F254, Merck) and developed thrice in petroleum ether/acetone (95:5, v/v). MAMES were visualised by autoradiography using Kodak BioMax MR films.

3.3.2.3 Lipoglycan analysis of *M. smegmatis* Δ *aftB*::pMV306-*aftB* conditional mutant

Three mycobacterial Ara f Ts have been identified so far to play a role in LAM biosynthesis: an α (1 \rightarrow 5) Ara f T EmbC and α (1 \rightarrow 3) branching Ara f Ts AftC and AftD (Goude *et al.*, 2008, Skovierova *et al.*, 2009, Birch *et al.*, 2010, Mishra *et al.*, 2012). Since AftB attaches terminal β (1 \rightarrow 2) Ara f residues to AG, we hypothesised that AftB plays a dual role and adds the terminal Ara f residues to the tetra-arabinoside and hexa-arabinoside motifs of LAM. Therefore, cultures of wild type *M. smegmatis*, *M. smegmatis*::pMV306-*aftB* and *M. smegmatis* Δ *aftB*::pMV306-*aftB* (grown with and without acetamide) were labelled with [14 C]-glucose, radioactive lipoglycans extracted and equal amount of radioactive samples analysed employing 15 % SDS-PAGE. Extracts from all strains showed presence of [14 C]-LAM, -LM and -PIMs (Figure 18A). Although densitometry analysis revealed a degree of variation in terms of [14 C]-LAM amounts extracted from wild type, merodiploid and the conditional mutant cultivated with and without acetamide strains (963 CNT/mm 2 , 1201 CNT/mm 2 , 1255 CNT/mm 2 and 875 CNT/mm 2 , respectively), no significant difference between lipoglycan samples was observed. To further assess the effects of AftB depletion on incorporation of arabinose into the LAM, radioactive lipoglycan extracts were analysed for their sugar content (Figure 18B). The [14 C]-arabinose content remained similar in *M. smegmatis* Δ *aftB*::pMV306-*aftB* mutants with and without depletion of AftB (Figure 18B). It is worth noticing that a considerable amount of other sugar residues were present in the samples, most likely due to contamination with α -glucan or intermediate molecules of mAGP.

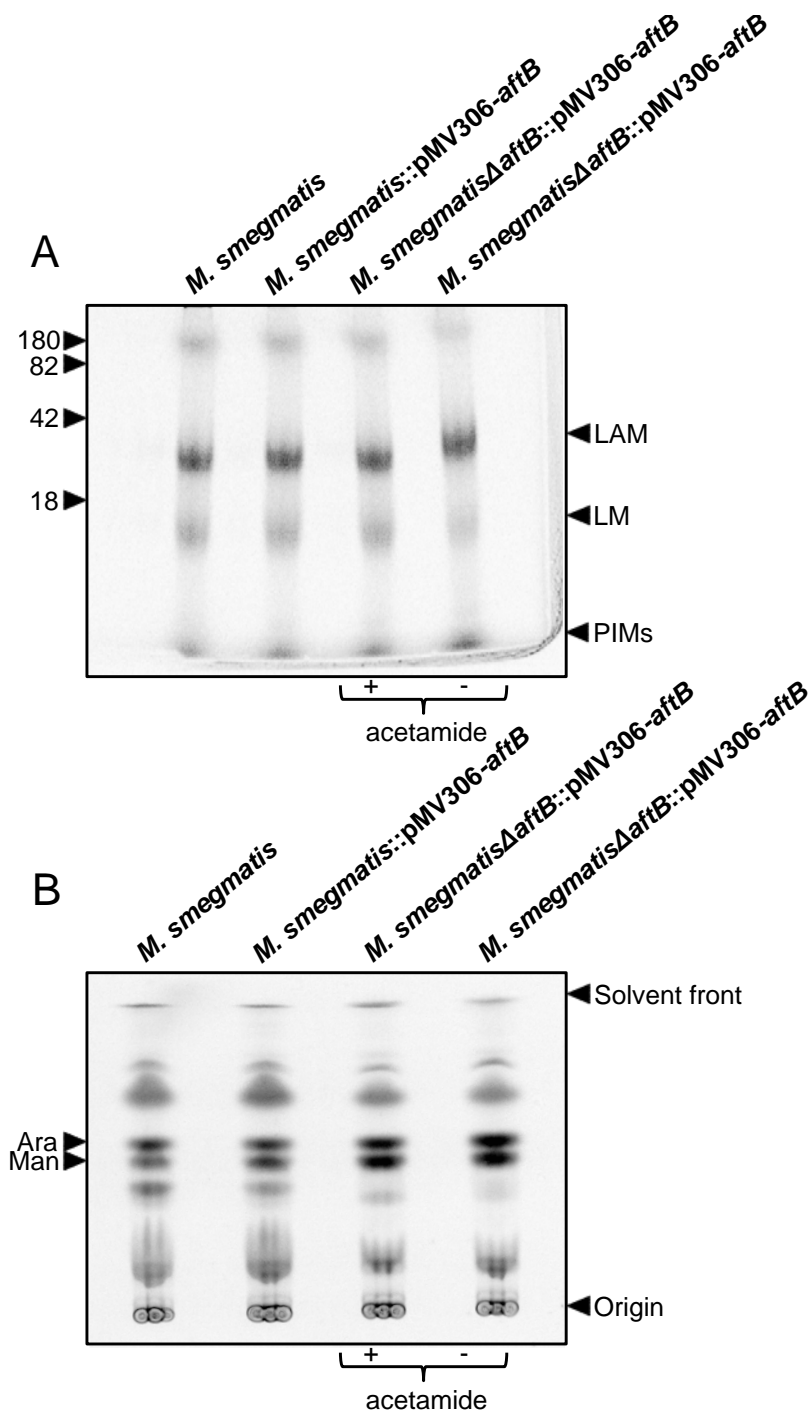


Figure 18 SDS-PAGE analysis of lipoglycans extracted from *M. smegmatis*, merodiploid and *M. smegmatis*Δ*aftB*::pMV306-*aftB* strains grown with and without acetamide. **A** Bacterial cells were refluxed with 50 % ethanol to extract crude lipoglycans. Samples were then subjected to hot phenol treatment and dialysed against water to yield a pure lipoglycan fraction. Equivalent aliquots (20,000 cpm) from each strain were analysed by SDS-PAGE and visualised by autoradiography employing Kodak BioMax MR film. **B** Lipoglycans were hydrolysed and samples of equivalent aliquots (20,000 cpm) loaded onto HPTLC-Cellulose TLC (Merck) and developed thrice in formic acid/water/tert-butanol/methyl ethyl ketone (3:3:8:6, v/v/v/v). The ^{14}C -containing sugars were visualised by autoradiography employing Kodak BioMax MR films.

3.3.2.4 Role of AftB in the synthesis of hexa-arabinan motif of LAM

Extracted lipoglycans from wild type *M. smegmatis*, *M. smegmatis*::pMV306-*aftB* and *M. smegmatis* Δ *aftB*::pMV306-*aftB* cells grown with and without acetamide were further analysed by immunoblot employing a monoclonal antibody F30-5, which specifically recognises the hexa-arabinan motif of AraLAM (Kolk *et al.*, 1984, Appelmelk *et al.*, 2008). All strains reacted with the antibody F30-5, with an exception of *M. smegmatis* Δ *aftB*::pMV306-*aftB* strain depleted of AftB (Figure 19). This data suggest that loss of AftB enzymatic activity in the conditional mutant results in a variant of LAM, that lacks terminal $\beta(1\rightarrow2)$ Ara_f residues in the classic hexa-arabinan motif and, thus is no longer recognised by F30-5 monoclonal antibody.

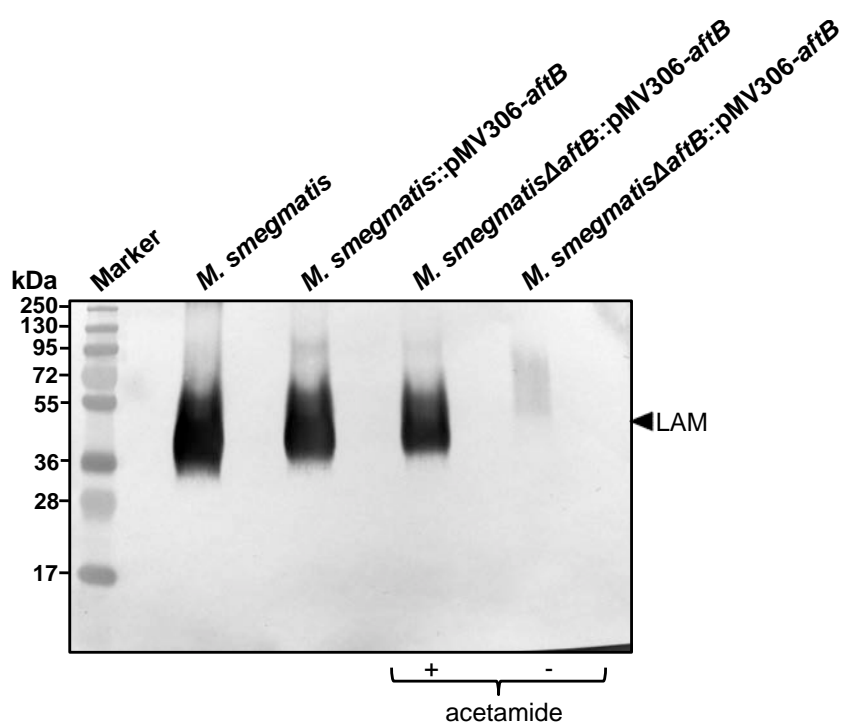


Figure 19 SDS-PAGE immunoblot of mycobacteria probed with anti-arabinan monoclonal antibody F30-5. Bacterial cells were refluxed with 50 % ethanol to extract crude lipoglycans. Samples were then subjected to hot phenol treatment and dialysed against water to yield pure lipoglycan fraction. Equivalent aliquots (50,000 cpm) from each strain were subjected to SDS-PAGE and subsequently immunoblot employing monoclonal antibody F30-5, which recognises the hexa-arabinan motif of LAM as described in Materials and Methods 6.6.2.

3.3.2.5 Overexpression studies of recombinant AftB from *M. tuberculosis* and *C. glutamicum*

In a similar fashion to AftA expression studies, the hydrophilic C-terminal domain encoded by *aftB*^{CT} of *M. tuberculosis* was cloned into pET16b and transformed into *E. coli* BL21 (DE3), C43 (DE3), Tuner (DE3), and Rosetta (DE3) competent cells. SDS-PAGE and Western blot analyses of bacterial cultures demonstrated moderate amounts of AftB^{CT} expressed under various conditions, however none was produced in the soluble form (Figure 20-21). In addition, full-length AftB from *M. tuberculosis* and *C. glutamicum* encoded by *Rv3805c* and *NCgl2780*, respectively, were cloned into both pMSX and pEKE5, and electroporated into *C. glutamicum* to assess expression and solubility of recombinant AftB. Unfortunately, a lack of expression in soluble or insoluble form was observed under all four different conditions: 16 °C for 12 hours using 0.1 mM IPTG, 16 °C for 12 hours using 1 mM IPTG, 30 °C for 4 hours using 0.1 mM IPTG and 30 °C for 4 hours using 1 mM IPTG. Therefore, production of recombinant AftB was not pursued further.

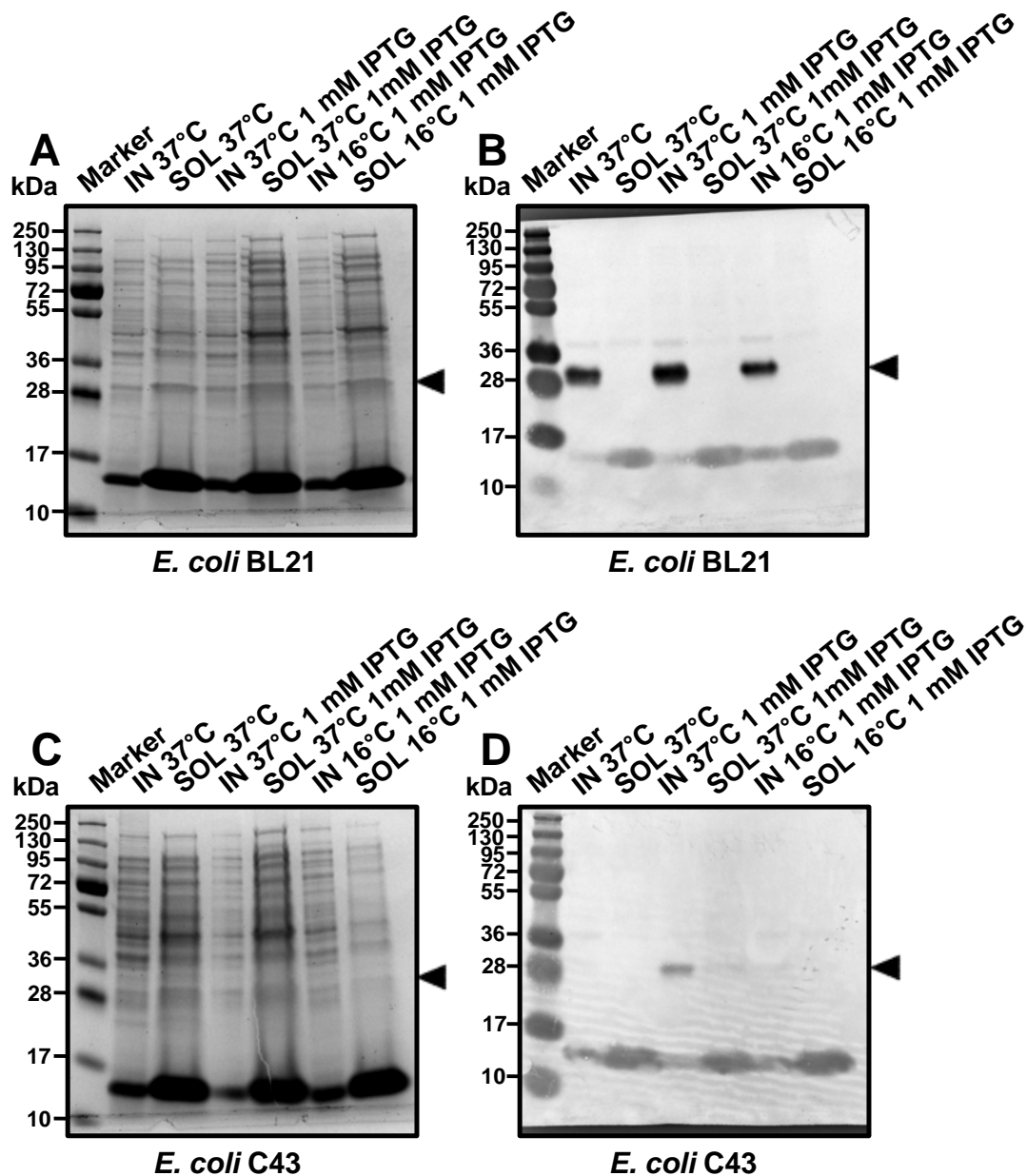


Figure 20 SDS-PAGE (A and C) and Western blot (B and D) analysis of AftB C-terminus from *M. tuberculosis* overexpressed in *E. coli* BL21 (DE3) and C43 (DE3) cells. *E. coli* cells harbouring pET16b-*aftB*^{CT} plasmid were grown at 37°C for 1 hour. Expression of AftB C-terminus was induced with 1 mM IPTG and cells were further cultivated at either 37 °C for 4 hours or at 16 °C overnight. Subsequently, cells were harvested, lysed and separated into two fractions: crude cell lysate (IN) and clarified cell lysate (SOL). Samples were analysed by 12 % SDS-PAGE gel and stained with Coomassie Blue; the His-tagged AftB^{CT} was detected by a Western blot using a monoclonal anti-His₆ tag antibody. The arrows indicate the overexpressed AftB^{CT} (28 kDa).

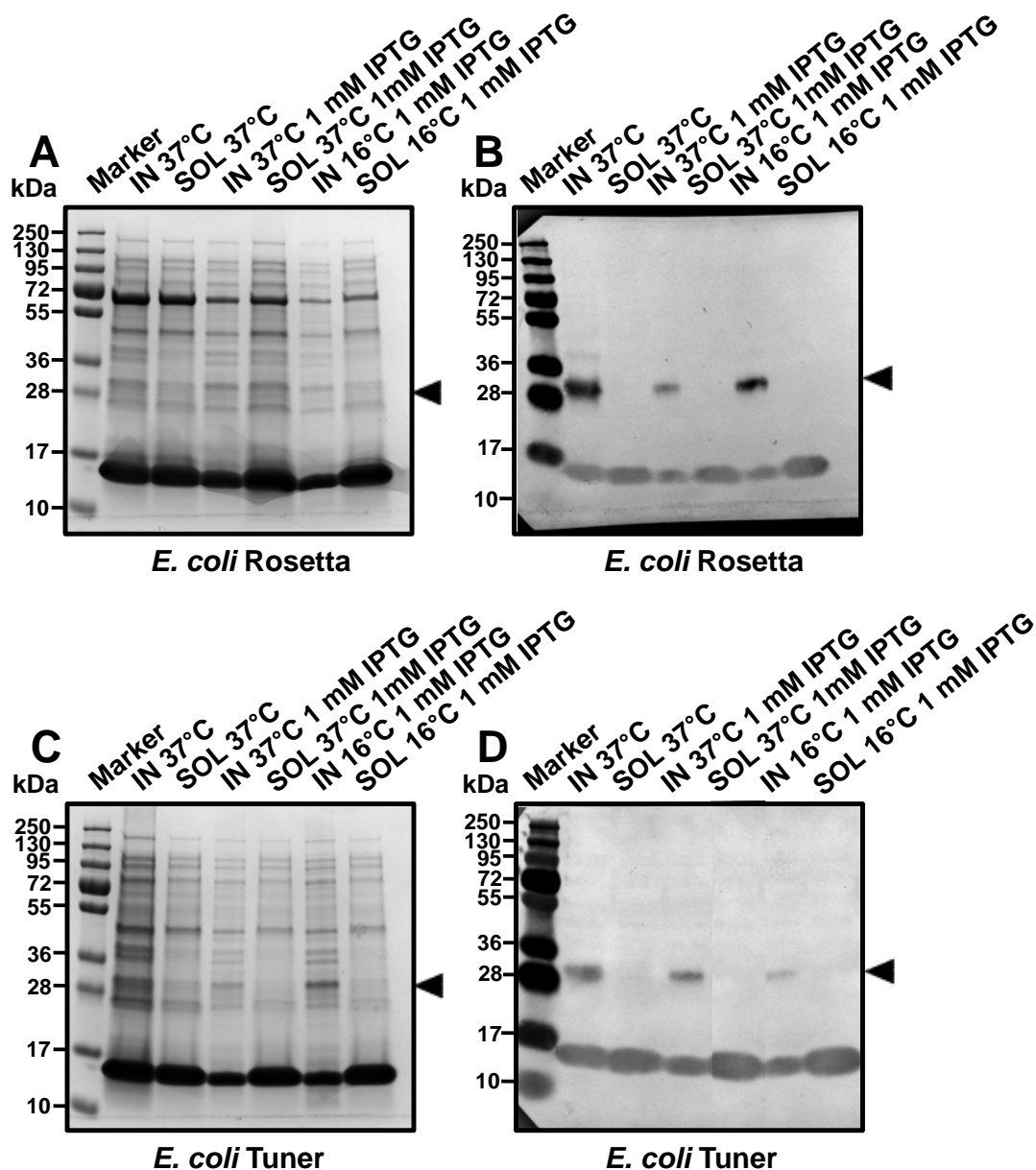


Figure 21 SDS-PAGE (A and C) and Western blot (B and D) analysis of AftB C-terminus from *M. tuberculosis* overexpressed in *E. coli* Rosetta (DE3) pLysS and Tuner (DE3) cells. *E. coli* cells harbouring pET16b-*aftB*^{CT} plasmid were grown at 37 °C for 1 hour. Expression of AftB C-terminus was induced with 1 mM IPTG and cells were further cultivated at either 37 °C for 4 hours or at 16 °C overnight. Subsequently, cells were harvested, lysed and separated into two fractions: crude cell lysate (IN) and clarified cell lysate (SOL). Samples were analysed by 12 % SDS-PAGE gel and stained with Coomassie Blue; the His-tagged AftB^{CT} was detected by a Western blot using a monoclonal anti-His₆ tag antibody. The arrows indicate the overexpressed AftB^{CT} (28 kDa).

3.4 Discussion

Both LM and LAM display immunoregulatory and anti-inflammatory properties that affect the host immune response (Mishra, 2011, Jankute *et al.*, 2012). Species specific capping moieties of the arabinan domain in LAM including manno-oligosaccharides (ManLAM) in slow growers, such as *M. tuberculosis* (Chatterjee *et al.*, 1993), phosphoinositide units (PILAM) in fast growers, such as *M. smegmatis* (Khoo *et al.*, 1995), and uncapped LAM known as AraLAM in *M. chelonae* (Guerardel *et al.*, 2002) have been demonstrated to play a role in modulating host response during infection (Khoo *et al.*, 1995, Gilleron *et al.*, 1997, Nigou *et al.*, 1997, Nigou *et al.*, 1999, Khoo *et al.*, 2001). Although structures of the arabinan domain and its capping motifs in LAM are well defined, the biosynthesis of both remains somewhat incomplete. Nevertheless, mycobacterial ArafTs with dual functionalities that are involved in LAM biosynthesis in addition to AG have been successfully identified, for example, AftC and AftD that both act as $\alpha(1\rightarrow3)$ ArafTs in the assembly of AG and LAM (Skovierova *et al.*, 2009, Birch *et al.*, 2010). We have investigated the potential role of known ArafTs AftA and AftB from *M. smegmatis* in the biosynthesis of LAM.

Previous study by Shi *et al.* (2008) reported that *aftA* (*MSMEG_6386*) from *M. smegmatis* is essential gene for AG biosynthesis and survival of mycobacteria. In order to analyse the effects of *aftA* deletion on LAM biosynthesis, we constructed a knockout phage designed to replace *MSMEG_6386* in *M. smegmatis* with a hygromycin cassette. We were unsuccessful in obtaining any null mutant transductants and therefore employed CESTET to construct *aftA* deletion in the presence of a second copy of inducible *aftA*. Conditional depletion of AftA in derived *M. smegmatis* Δ *aftA*::*pMV306-aftA* mutants

resulted in cell lysis. However, depletion of AftA protein in the mutant strain also yielded persisters, mycobacterial cells that acquired mutations in the acetamide promoter region resulting in continues production of AftA. Thus, a careful selection from generated transductants has to be performed in order to segregate true *aftA* conditional mutants in *M. smegmatis*. Further *in vitro* cell free assays and chemical characterisation of lipoglycans extracted from *aftA* conditional mutant would provide evidence of AftA role or the lack of it in the biosynthesis of LM/LAM.

The biosynthesis of mycobacterial AG and LAM requires a panel of glycosyltransferases including transmembrane ArafTs members whose structures are largely unknown (General Introduction 1.8.2). Due to inherent hydrophobicity of integral membrane proteins their expression and purification can be extremely difficult. Here we employed a strategy to express and purify the hydrophilic C-terminal domain of AftA from *M. tuberculosis*. Overexpression of glycosyltransferase AftA^{CT} from *M. tuberculosis* was investigated in standard microbial expression systems. Soluble pure recombinant AftA^{CT} was obtained in suitable amounts for crystallisation trials. However, a significant degree of precipitation was observed during protein purification. High throughput thermostability screening employing DSF was completed for AftA^{CT} in order to identify buffer conditions that reduce the protein's tendency to unfold and precipitate. Several solution compositions containing sodium acetate, varying in pH and salt concentration were determined as stabilising and, thus, were included in AftA^{CT} purification and crystallisation processes. However, none of the 96-well plate screens generated protein crystals.

The hydrophilic C-terminal domain of AftA glycosyltransferase can be further

characterised using a diverse range of methods. Binding assays employing purified AftA^{CT} as well as synthetic acceptor analogues would provide information on ligand preferences and extent of their interactions. Recombinant AftA^{CT} carrying single residue substitutions of conserved residues used in ligand-binding assays would identify amino acid residues critical for the specific ligand-binding interaction. Moreover, several other solution conditions identified by DSF could be tested for improved stability of AftA^{CT}, and result in higher protein yield and possibly facilitate further crystallisation efficacy. Finally, NMR spectroscopy could be applied in determining the structure of AftA^{CT}. Although sensitivity and resolution of NMR spectroscopy is inferior compared to that of X-ray crystallography, a number of successfully solved protein structures utilising NMR have been reported (Stehle *et al.*, 2012, Priya *et al.*, 2013, Kim *et al.*, 2014).

Recent studies have demonstrated that SMALPs offer a universal method for extracting, purifying and studying membrane proteins in their native lipid environment (Jamshad *et al.*, 2011, Orwick-Rydmark *et al.*, 2012, Gulati *et al.*, 2014, Paulin *et al.*, 2014). Here, we employed the SMALP method and purified AftA membrane protein directly from *C. glutamicum* membranes without the need of detergent. The non-peptide nature of the styrene maleic acid reagent ensured that spectroscopic studies of AftA structure and function remained unaffected allowing biochemical characterisation of the protein. In addition, ligand binding studies utilising neoglycolipid acceptors and recombinant AftA within SMALP demonstrated a markedly stronger affinity for the longer Acc13 (β -D-Galf(1 \rightarrow 6)- β -D-Galf(1 \rightarrow 5))6- β -D-Galf(1 \rightarrow 4)- α -L-Rhap-GlcNAc-*O*-(CH₂)₇CH₃) acceptor compared to G5G (β -D-Galf(1 \rightarrow 5)- β -D-Galf-*O*-(CH₂)₇CH₃) and G6G (β -D-Galf(1 \rightarrow 6)- β -D-Galf-*O*-(CH₂)₇CH₃). It is tempting to speculate that by having a higher affinity to longer linear Galf acceptor, AftA may more conveniently locate the acceptor for further

addition of first three *Araf* residues to the galactan chain. Overall, we have shown that detergent-free SMALP solubilisation and extraction of membrane proteins does work for membrane proteins from *Corynebacterineae*. This advance may remove a major bottleneck in the study of mycobacterial membrane proteins and opens new opportunities for structural and functional studies on some of the least studied molecules in mycobacterial research.

Similarly to AftA, we have demonstrated that the homologue of *Rv3805c* in *M. smegmatis* (*MSMEG_6400*) is an essential gene. The failure to obtain *aftB* null mutant in the absence of second inducible copy of *aftB*, coupled with inability of *M. smegmatis* to grow in the media without acetamide, an inducer that activates *aftB* expression, showed that AftB enzymatic function is crucial for the viability of *M. smegmatis*. Although *aftB* homologue in *C. glutamicum* has been reported to be non-essential (Seidel *et al.*, 2007a), high density mutagenesis studies by Sasseti *et al.* (2003) suggest that *Rv3805c* in *M. tuberculosis* is an essential gene. AftB has been reported to act as $\beta(1\rightarrow2)$ ArafTs, which attaches *Araf* residues to the nonreducing end of the arabinan domain of AG in *C. glutamicum* (Seidel *et al.*, 2007a). SDS-PAGE and sugar analyses of lipoglycans extracted from both parental and conditional mutant strains did not exhibit significant difference between each other, however monoclonal antibody F30-5 generated against hexa-arabinan motif of LAM no longer recognised LAM extracted from the *aftB* mutant, thus suggesting a loss of terminal $\beta(1\rightarrow2)$ *Araf* residues. Further chemical characterisation of the LAM from *aftB* conditional mutant by glycosyl linkage analysis or NMR is necessary to confirm our predictions.

Expression of the full length AftB and its hydrophilic C-terminal domain from *M.*

tuberculosis and *C. glutamicum* was also examined. Production of full length AftB was not observed under any conditions, whereas AftB^{CT} was produced exclusively as insoluble inclusion bodies in all *E. coli* cell lines tested. The overexpression of recombinant proteins in *E. coli* host is commonly utilised to generate proteins in large amounts. However, approximately 70 % of recombinant proteins expressed in *E. coli* either degrade or form insoluble inclusion bodies (Yang *et al.*, 2011). A number of methods have been developed to obtain soluble protein from inclusion bodies, but a general strategy involves solubilising aggregated protein in denaturant, followed by subsequent refolding resulting in soluble and biologically active protein (Cabrita and Bottomley, 2004, Singh and Panda, 2005, Burgess, 2009). This strategy could be employed with recombinant AftB^{CT} that has generated in large amounts, but formed insoluble inclusion bodies. Moreover, reducing the rate of protein synthesis in *E. coli* host by lowering inducer concentration or using a lower copy number plasmid could potentially improve protein solubility. Finally, alterations of the host strain, changes of the construct encoding AftB^{CT} or co-expression of proteins with molecular chaperones could lead to higher levels of soluble protein.

Chapter 4

Molecular and biochemical
characterisation of Emb
arabinosyltransferase from
Corynebacterium glutamicum

4 Molecular and biochemical characterisation of Emb arabinosyltransferase from *Corynebacterium glutamicum*

4.1 Introduction

AG is a key structural component of the mAGP complex that plays an important role in covalently anchoring the mycolic acid layer to the inner PG, and is unique in that all of its sugars are present in the furanose ring form (McNeil *et al.*, 1987). Moreover, unlike most bacterial polysaccharides, AG lacks repeating units and is instead composed of a few distinct structural motifs (Daffe *et al.*, 1990). Previous research has demonstrated that the whole mAGP complex is covalently attached to PG *via* a specific linker unit, which connects the linear galactan chain to selected MurNGly residues of PG (McNeil *et al.*, 1990). These cell wall components are crucial for the growth, viability and virulence of *M. tuberculosis* and, therefore, are often the targets of effective chemotherapeutic agents, such as EMB.

The galactan component of AG is present in the form of approximately 30 alternating $\beta(1\rightarrow5)$ and $\beta(1\rightarrow6)$ Galf residues combined in a linear fashion. Three arabinan chains, each containing approximately 30 Araf residues, are linked to this linear galactan at C-5 of some of $\beta(1\rightarrow6)$ Galf residues (Daffe *et al.*, 1990, Bhamidi *et al.*, 2008). The arabinan domain, in contrast to the linear galactan, is highly branched. Since the AG component is essential to mycobacteria, its structure was also investigated in the related *Corynebacterium* genus, in which AG is non-essential. Deletion studies in *C. glutamicum* coupled with mass spectrometry analysis determined that the three arabinan chains are attached at positions 8, 10 and 12 of the galactan chain (Alderwick *et al.*, 2006). The

majority of *Araf* residues form a backbone containing α -5-linked α -D-*Araf* residues with branching introduced at the C-3 hydroxyl forming 3,5-*Araf* linked residues (Daffe *et al.*, 1990). At the non-reducing end of AG, the highly branched arabinan domain is covalently attached to the mycolic acids of the outer membrane, thus completing the whole mAGP complex (Kaur *et al.*, 2009). Previous research has shown that the non-reducing termini of the arabinan structure consists of a distinct Ara₆ motif [β -D-*Araf*-(1 \rightarrow 2)- α -D-*Araf*]₂-3,5- α -D-*Araf*-(1 \rightarrow 5)- α -D-*Araf*], with the mycolic acids attached to position 5 of both the terminal β -D-*Araf* units and the penultimate 2- α -D-*Araf* units (McNeil *et al.*, 1991).

Earlier studies demonstrated that EMB, a front-line anti-TB drug, specifically inhibits AG biosynthesis with the precise molecular target being the *embCAB* locus in *M. tuberculosis* (Telenti *et al.*, 1997) and *embRAB* locus in *M. avium* (Belanger *et al.*, 1996). Individual disruption of *embA*, *embB* or *embC* in *M. smegmatis* resulted in three slow growing viable mutants. In comparison to wild type *M. smegmatis*, the Ara₆ motif of AG was altered in both *embA* and *embB* mutants, whereas the *embC* deletion resulted in diminished LAM synthesis (Escuyer *et al.*, 2001, Goude *et al.*, 2008). These studies suggest that both EmbA and EmbB are involved in the formation of AG, specifically Ara₆, while EmbC is involved in the biosynthesis of LAM. Interestingly, in comparison with *M. tuberculosis*, *C. glutamicum* has been shown to have only one *emb* gene, which is not essential to this organism (Alderwick *et al.*, 2005). Nevertheless, the deletion of this gene caused severe reduction of arabinose in the cell wall, resulting in a truncated AG structure with terminal *Araf* residues on the galactan chain (Alderwick *et al.*, 2005). It is important to note that although AG from *C. glutamicum* is comparable to that of *M.*

tuberculosis, the LAM structure is significantly different in terms of arabinosylation. Specifically, LAM from *C. glutamicum* lacks high-branching arabinose domain and contains single Ara_f residues attached to the mannan core by AftE (*NCgl2096*).

Another Ara_fT from the *emb* locus, AftA (*Rv3792*), has been demonstrated to catalyse the addition of the very first Ara_f residue from the DPA onto the galactan chain at positions 8, 10 and 12, thus priming the galactan for further extension by unknown $\alpha(1\rightarrow5)$ Ara_fT(s) (Alderwick *et al.*, 2006). Recent knockout studies in *M. smegmatis* demonstrated *aftA* to be essential for viability, however an *aftA* deletion in *C. glutamicum* resulted in a slow growing mutant with a cell wall lacking arabinose, as well as cell wall bound mycolates (Alderwick *et al.*, 2006). Knockout studies in *M. smegmatis* identified AftC (*MSMEG_2785*, orthologue of *Rv2673*), which is responsible for the α -1,3-branching of the internal arabinan domain at the non-reducing end of AG and LAM (Birch *et al.*, 2008, Birch *et al.*, 2010). A functional Ara_fT encoded by *aftD* (*Rv0236c*) in *M. tuberculosis* was assigned as possessing α -1,3-branching activity (Skovierova *et al.*, 2009). Due to its size, AftD has been suggested to have other additional functions in AG and LAM biosynthesis. For example, it may act as a scaffold for the multi-enzyme machinery involved in arabinan synthesis (Skovierova *et al.*, 2009). The terminal $\beta(1\rightarrow2)$ Ara_f residues are transferred onto the arabinan domain by AftB (*Rv3805c*), thus marking the end-point of AG biosynthesis before decoration with mycolic acids (Belisle *et al.*, 1997, Seidel *et al.*, 2007).

In this study we examined effects of an *aftA* and *emb* double deletion mutant on AG biosynthesis in *C. glutamicum*. Analysis of its cell wall revealed a complete absence of arabinose resulting in a truncated cell wall containing only galactan accompanied with

the loss of cell wall bound mycolates. In addition, by employing an *in vitro* cell free assay using various *C. glutamicum* mutants (*C. glutamicum* Δ *aftA*, *C. glutamicum* Δ *emb*, *C. glutamicum* Δ *aftB* Δ *aftD*, *C. glutamicum* Δ *aftA* Δ *emb*) and use of synthetic oligosaccharide acceptors, β -D-Galf-(1 \rightarrow 5)- β -D-Galf-(1 \rightarrow 6)[α -D-Araf-(1 \rightarrow 5)]- β -D-Galf and β -D-Galf-(1 \rightarrow 6)-[α -D-Araf-(1 \rightarrow 5)]- β -D-Galf-(1 \rightarrow 5)- β -D-Galf, we demonstrated that the transfer of Araf residue from DPA to the primed galactan in *C. glutamicum* is catalysed by Emb ArafT with α (1 \rightarrow 5) activity.

4.2 Materials and Methods

4.2.1 Construction of plasmids and strains

C. glutamicum Δ *aftB* Δ *aftD* and *C. glutamicum* Δ *aftA* Δ *emb* strains were kindly provided by Dr Lothar Eggeling, Jülich Research Centre, Germany. Briefly, to construct the deletion vector pK19mobsacB Δ *aftD* (*NCgl2757*) PCR was applied with primer pairs AB (A, 5'-CGC TTC TAG ACC ACG TCA TGG CAT ACG AAA CTG-3'; B, 5'-CCC ATC CAC TAA ACT TAA ACA CAC CAC AAA ACC CAG CAC-3') and CD (C, 5'-TGT TTA AGT TTA GTG GAT GGG GCT GCG CTA TTT GCT GG-3'; D, 5'-GCG GGA ATT CGC ATG GCA AGC CGG TAA G-3') and genomic DNA of the wild type of *C. glutamicum* ATCC 13032 used as template. Both amplified products were used in a second PCR with primer pairs AD to generate a 620 bp fragment consisting of sequences adjacent to *C. glutamicum* *aftD*, which was ligated with XbaI-EcoRI-cleaved pK19mobsacB.

All plasmid inserts were confirmed by sequencing. The chromosomal deletion of *C. glutamicum* *aftD* was performed as described previously using two rounds of positive

selection (Schafer *et al.*, 1994), and its successful deletion was verified by use of two different primer pairs hybridising in the genome. In order to construct the *aftB/aftD* double mutant *C. glutamicum* Δ *aftB* Δ *aftD*, the deletion vector pK19mobsac Δ *aftD* was also used and introduced into the previously reported strain *C. glutamicum* Δ *aftB* (Figure 1) (Seidel *et al.*, 2007). Similarly, the deletion vector pK19mobsac Δ *aftA* constructed previously (Alderwick *et al.*, 2006) was employed and introduced into previously reported strain *C. glutamicum* Δ *emb* (Alderwick *et al.*, 2005), thus generating *C. glutamicum* Δ *aftA* Δ *emb* (Figure 1).

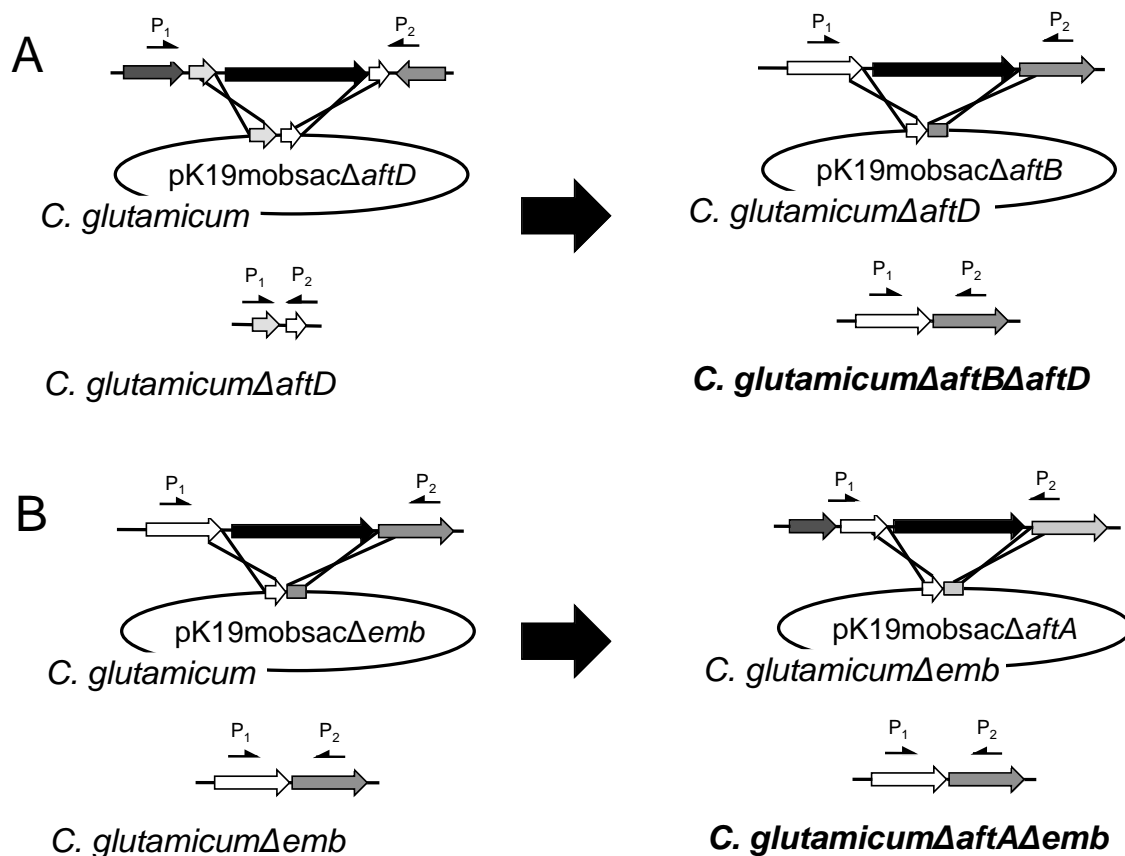


Figure 1 Construction strategy of *C. glutamicum* Δ *aftB* Δ *aftD* (A) and *C. glutamicum* Δ *aftA* Δ *emb* (B) double mutants.

4.2.2 Extraction and purification of arabinogalactan from *C. glutamicum*

The mAGP samples of wild-type *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb* were prepared as described in General Materials and Methods 6.5.16. Mycolic acids were removed from the mAGP by saponification: using 2 % potassium hydroxide in methanol-toluene (1:1) for 48 hours. The insoluble residue was recovered by centrifugation at 27,000 g. The sample was washed 3-5 times with methanol and the resulting AGP complex freeze-dried. The AGP (2 g) preparation was treated with 75 ml of 2 M sodium hydroxide for 16 hours at 80 °C. The supernatant, which contains solubilised AG, was recovered by centrifugation at 27,000 g for 30 minutes. The crude AG preparation was neutralised with acetic acid, re-centrifuged and dialysed to remove residual salts (MWCO 3,500). A precipitate that was formed during dialysis was removed by centrifugation. The supernatant was diluted in cold ethanol (80 %, v/v) and left at -20 °C overnight to precipitate the base-solubilised AG, which was recovered by centrifugation, lyophilised and stored at 4 °C.

4.2.3 NMR spectroscopic analysis of AG from *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb*

NMR spectra of AG from wild-type *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb* samples were recorded using Bruker DMX-500. Samples were repeatedly exchanged in deuterium oxide (99.9 atom % D, Sigma Aldrich) with intermediate lyophilisation, and finally dissolved in 0.5 ml deuterium oxide and analysed at 313 K. The ^1H and ^{13}C NMR chemical shifts were referenced relative to internal acetone at 2.225 and 34.00 ppm,

respectively. Data was collected and analysed using Bruker XWIN-NMR software. Details concerning NMR sequences used and experimental procedures were described previously (Daffe *et al.*, 1990).

4.2.4 Extraction and purification of lipoglycans from *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb*

Wild-type *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb* cells (10 ml) were grown to an early exponential growth phase (OD₆₀₀ of 0.6), labelled with [¹⁴C]-glucose 1 μ Ci/ml and incubated for a further 12 hours at 30 °C with shaking. The labelled cells were harvested and washed twice in phosphate buffered saline. The pellet was resuspended in 4 ml of 50 % ethanol in water (v/v) and the mixture refluxed at 85 °C for 6 hours. The supernatant was recovered by centrifugation and transferred to a fresh tube. The extraction process was repeated five times and the supernatant fractions pooled and dried. The crude lipoglycans were subjected to a 90 % phenol treatment at 65 °C for 1 hour. After cooling, the sample was centrifuged and the upper aqueous layer recovered and dialysed against water (MWCO 3500 kDa). An equal amount of radioactivity was subjected to 15 % SDS-PAGE and [¹⁴C]-labelled purified lipoglycans were visualised by autoradiography employing Kodak BioMax MR films.

4.2.5 Preparation of *C. glutamicum* membranes

After cultivation, *C. glutamicum* cells (wet weight 10 g) were resuspended in 50 mM MOPS pH 7.9, 5 mM β -mercaptoethanol and 10 mM MgCl₂. Cell slurry was sonicated (MSE Soniprep 150, 12 micron amplitude, 60 seconds ON, 90 seconds OFF for 10 cycles) and centrifuged at 27,000 g for 20 minutes at 4 °C. The resultant supernatant was

centrifuged again at 100,000 g for 90 minutes at 4 °C and isolated cell membranes were collected and concentrated to 15-20 mg/ml in 50 mM MOPS pH 7.9, 10 mM MgSO₄, 5 mM β-mercaptoethanol.

4.2.6 Preparation of *C. glutamicum* P60 cell wall material

C. glutamicum cells (wet weight 10 g) were resuspended in 50 mM MOPS pH 7.9, 5 mM β-mercaptoethanol and 10 mM MgCl₂. The cell slurry was sonicated (MSE Soniprep 150, 12 micron amplitude, 60 seconds ON, 90 seconds OFF for 10 cycles) and centrifuged at 27,000 g for 20 minutes at 4 °C. Resultant pellet was resuspended in buffer containing 50 mM MOPS pH 7.9, 10 mM MgSO₄, 5 mM β-mercaptoethanol (24 ml) and Percoll (32 ml) (Sigma Aldrich), mixed thoroughly and centrifuged at 27,000 g for 60 minutes at 4 °C. The upper band, corresponding to *C. glutamicum* cell wall 'P60' material was carefully removed, washed using 50 mM MOPS pH 7.9, 10 mM MgSO₄, 5 mM β-mercaptoethanol to remove Percoll and the final cell wall fraction resuspended in 50 mM MOPS pH 7.9, 10 mM MgSO₄, 5 mM β-mercaptoethanol to a final protein concentration of 10 mg/ml.

4.2.7 [¹⁴C]-pRpp synthesis

Radiolabelled [¹⁴C]-glucose (200 μCi) was dried and resuspended in 500 μl of ice-cold 50 mM KH₂PO₄, 5 mM MgCl₂. Both ATP and NADP were added to the reaction mixture to a final concentration of 1 mM and 4 mM, respectively. Hexokinase/ glucose-6-phosphate dehydrogenase mixture (5 units, Roche) was added to the reaction, gently mixed and incubated at 37 °C for 15 minutes. Subsequently, 6-phosphogluconic dehydrogenase (5 units, Sigma Aldrich) was added to the mixture, gently mixed and

incubated at 37 °C for 15 minutes followed by addition of phosphoriboseisomerase (5 units, Sigma Aldrich) and incubation at 30 °C for 15 minutes. Purified phosphoribosyl pyrophosphate synthetase PrsA from *M. tuberculosis* prepared as described in Alderwick *et al.* (2011) was added to the reaction and incubated at 37 °C for 30 minutes. The reaction mixture was passed through a protein concentrator to remove enzymes used in the assay and the [¹⁴C]-pRpp was purified employing ion exchange LC-SAX column. All unbound material was removed using water (2 ml) and [¹⁴C]-pRpp eluted *via* a gradient of sodium acetate (100 mM-2M). Fractions containing pure [¹⁴C]-pRpp were collected and stored at -20 °C for further use.

4.2.8 Arabinofuranosyltransferase activity using membrane preparations of *C. glutamicum*, *C. glutamicum*Δ*aftA*, *C. glutamicum*Δ*emb*, *C. glutamicum*Δ*aftA*Δ*emb* and *C. glutamicum*Δ*aftB*Δ*aftD*

Membrane and P60 fractions were prepared as described earlier (Materials and Methods 4.2.5 and 4.2.6, respectively). Neoglycolipid acceptors MJ-13-77 and MJ-14-01 (2 µl of 20 mM, stored in ethanol) and decaprenyl phosphate (1 µl of mg/ml, stored in ethanol) were aliquoted into 1.5 ml Eppendorf tubes and dried. IgePalTM (Sigma-Aldrich) was added (0.1%, v/v) with the appropriate amount of buffer (50 mM MOPS pH 7.9, 10 mM MgSO₄, 5 mM β-mercaptoethanol) to a final volume of 80 µl. Samples were sonicated for 15 minutes to resuspend lipid linked substrates and then mixed with the remaining assay components: membrane protein and ‘P60’ fraction (1 mg each) from either *C. glutamicum*, *C. glutamicum*Δ*aftA*, *C. glutamicum*Δ*emb*, *C. glutamicum*Δ*aftA*Δ*emb*, *C. glutamicum*Δ*aftB*Δ*aftD*, 1 mM ATP, 1 mM NADP, p[¹⁴C]Rpp (25,000 cpm) and in some cases EMB (1 mg/ml). Reaction mixtures were incubated for 1 hour at 37 °C, quenched

by the addition of 533 μ l of chloroform/methanol (1:1, v/v) and mixed for 30 minutes. Supernatant was recovered following centrifugation at 27,000 g for 30 minutes and dried under nitrogen. The residue was resuspended in 750 μ l ethanol/water (1:1, v/v) and loaded onto 1 ml SepPak ion exchange columns, pre-equilibrated with ethanol/water (1:1, v/v). The column was washed twice with 2 ml of ethanol (100 %) and the eluate collected and dried. Sample was resuspended in a mixture of water-saturated *n*-butanol (3 ml) and water (3 ml), mixed and the organic phase recovered following centrifugation. The aqueous phase was extracted once again with *n*-butanol (3 ml) and the organic phases pooled. The extracts were further washed using *n*-butanol-saturated water (3 ml). Finally, the *n*-butanol fraction was dried and resuspended in 200 μ l of *n*-butanol. The incorporation of [14 C]Araf was determined by subjecting samples to TLC using silica gel plates (5735 silica gel 60F₂₅₄, Merck) developed in isopropanol:acetic acid:water (8:1:1, v/v/v) and visualised by autoradiography employing Kodak BioMax MR films.

4.2.9 Analysis of arabinofuranosyltransferase reaction products prepared from *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb* membranes

Large-scale reactions were performed containing cold pRpp (25 mM, Sigma Aldrich). Membranes and 'P60' fraction were prepared as described in Materials and Methods 4.2.5 and 4.2.6, respectively. Neoglycolipid acceptors MJ-13-77 and MJ-14-01 (8 μ l of 20 mM, stored in ethanol) and decaprenyl phosphate (4 μ l of mg/ml, stored in ethanol) were aliquoted into glass tubes and dried. IgePalTM (Sigma-Aldrich) was added (0.1%, v/v) with the appropriate amount of buffer (50 mM MOPS pH 7.9, 10 mM MgSO₄, 5 mM β -mercaptoethanol) to a final volume of 320 μ l. Samples were sonicated for 15 minutes to resuspend lipid linked substrates and then mixed with the remaining assay components:

membrane proteins and 'P60' fraction (1 mg each) from either *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb*, 1 mM ATP, 1 mM NADP, cold pRpp (25 mM, Sigma Aldrich) and in some cases EMB (1 mg/ml). The reaction mixture was incubated at 37 °C for 1 hour. The assays were replenished with fresh membranes and 'P60' and re-incubated at 37 °C for 1 hour with the entire process repeated thrice. Reaction products were extracted with *n*-butanol/water phase separation as described earlier (4.2.8). Products together with a radiolabelled reference product were subjected to preparative TLC plates, developed in isopropanol:acetic acid:water (8:1:1, v/v/v) and the reference sample visualised by autoradiography employing Kodak BioMax MR films. The cold products were recovered from the TLC plate by *n*-butanol extraction. The butanol phases were washed with water saturated with *n*-butanol and dried. Samples were analysed by MALDI-MS and MS/MS as previously described (Lee *et al.*, 2006, Shi *et al.*, 2008).

4.3 Results

4.3.1 Lipid characterisation of *C. glutamicum* Δ *aftA* Δ *emb* mutant

To compare the double-deletion *C. glutamicum* Δ *aftA* Δ *emb* mutant to previously reported phenotypes of *C. glutamicum* Δ *aftA* (Alderwick *et al.*, 2006) and *C. glutamicum* Δ *emb* (Alderwick *et al.*, 2005) its cellular composition was studied. Firstly, wild type *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb* strains were analysed for AG esterified corynomycolic acids and cell wall associated lipids, both from an equivalent starting amount of biomass for each strain due to differences in growth rate. As expected, cell wall bound corynomycolic acid methyl esters (CMAMES) were abolished in the *C. glutamicum* Δ *aftA* Δ *emb* mutant indicating a defect in cell wall biosynthesis in comparison

to the wall of *C. glutamicum* (Figure 2B). Analysis of the cell wall associated lipids in the mutant highlighted an apparent increase in trehalose dicorynomycolates (TDCM) and trehalose monocorynomycolates (TMCM), the equivalent of mycobacterial TDMs and TMMs, respectively (Figure 2A). This was verified quantitatively by labelling cultures with [¹⁴C]-acetate and loading equal amounts of radioactive extractable free lipids from *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb*. As expected, wild type *C. glutamicum* exhibited the known free lipid profile containing non-polar lipids (130 CNT/mm²), TDCMs (134 CNT/mm²), TMCMs (97 CNT/mm²), and phospholipids (356 CNT/mm²). In contrast, densitometry based analysis of the *C. glutamicum* Δ *aftA* Δ *emb* mutant demonstrated a significant 3.5 fold increase in TDCMs (503 CNT/mm²) and a 2.5 fold increase in TMCMs (234 CNT/mm²). These results are consistent with previous studies and indicate that AftA and Emb from *C. glutamicum* are involved in key steps of arabinan biosynthesis of AG.

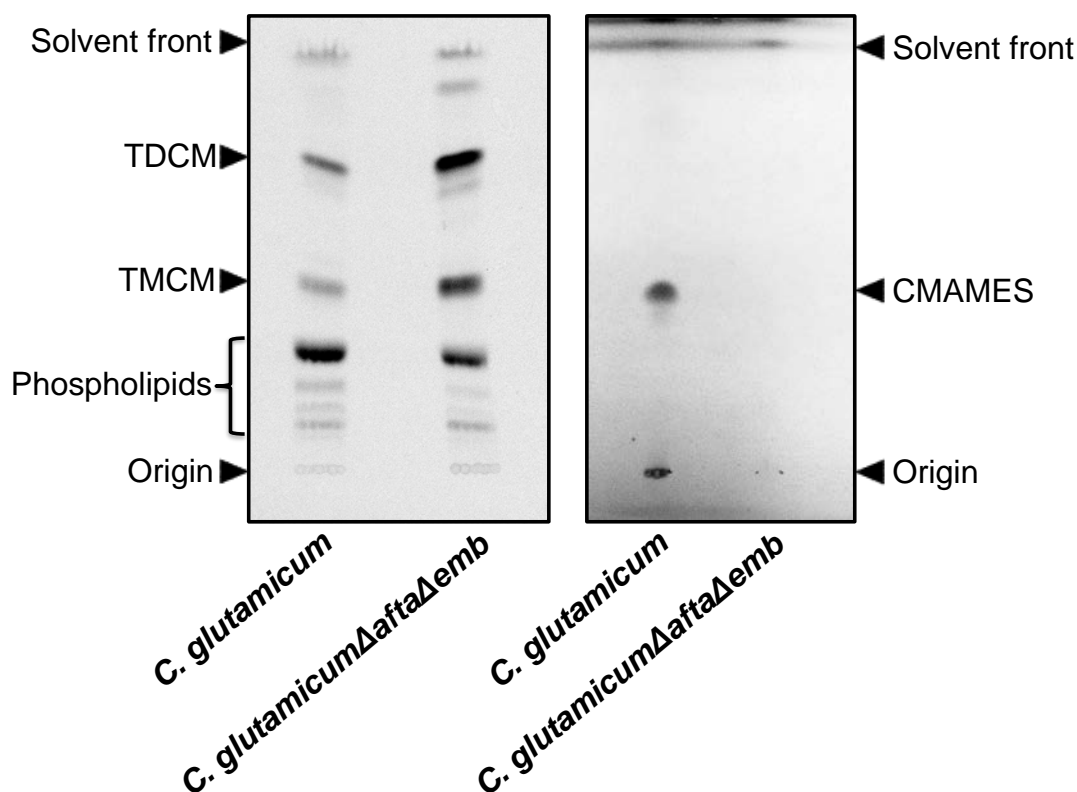


Figure 2 Analysis of cell wall lipids from *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb*. **A.** Analysis of cell wall associated [¹⁴C]-lipids from *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb*. Lipids were extracted from cells by a series of organic washes as described in General Materials and Methods 6.5.14. Equivalent aliquots (25,000 cpm) from each strain were subjected to TLC using silica gel plates (5725 silica gel 60F254, Merck) developed in chloroform/methanol/water (60:16:2, v/v/v) and visualised by autoradiography using Kodak BioMax MR films. **B.** Analysis of cell-wall bound lipids from *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb* cells. An equivalent amount of cells (100 mg) were delipitated using three extractions of chloroform/methanol/water (10:10:3, v/v/v) at 50 °C for 4 hours, and bound mycolic acids were released by addition of *tetra*-butyl ammonium hydroxide at 100 °C for 12 hours and subsequently methylated. An equal amount of sample from each strain was subjected to TLC using silica gel plates (5725 silica gel 60F254, Merck) and developed in petroleum ether/acetone (95:5, v/v). CMAMES were visualised using molybdophosphoric acid (5 % in ethanol), following heating at 100 °C.

4.3.2 Structural characterisation of *C. glutamicum* Δ *aftA* Δ *emb*

GC analysis of alditol acetates prepared from *C. glutamicum* mAGP demonstrates presence of rhamnose, arabinose and galactose sugar residues as reported previously (Alderwick *et al.*, 2005, Alderwick *et al.*, 2006, Seidel *et al.*, 2007). However, analysis of alditol acetates derived from *C. glutamicum* Δ *aftA* Δ *emb* mAGP revealed a total loss of arabinose. Previous independent deletion studies of either *emb* (Alderwick *et al.*, 2005) or *aftA* (Alderwick *et al.*, 2006) in *C. glutamicum* resulted in a drastically truncated structure of AG possessing only t-Araf residues and a total loss of remaining Araf residues, respectively. Consequently, the loss of cell wall arabinose was expected in the *C. glutamicum* Δ *aftA* Δ *emb* mutant.

Further analysis of base solubilised AG from *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb* were performed using NMR spectroscopy. The ^1H -NMR spectrum of wild type *C. glutamicum*-AG (Figure 4A) was highly complex when compared to the anomeric region of *C. glutamicum* Δ *aftA* Δ *emb*-‘AG’ (Figure 4B). Indeed, the cell wall ‘AG’ of the double mutant 1D ^1H spectrum exhibited two major well-defined resonances characterised by several overlapping resonances arising from two different classes of glycosidic residues: 5- β -Gal f (d110.6, d5.18 ppm) and 6- β -Gal f (d111.4, d4.98). Based on our data for wild type *C. glutamicum*-AG and previously published work by Daffe *et al.* (1990), the ^{13}C resonance at d104.2 ppm correlates to an anomeric proton at d5.1 ppm, and was assigned as t- β -Araf \rightarrow 5 linkage. The resonances at d108.9 and d109.2, correlating to protons at d5.2 and d5.12 were assigned as 2- α -Araf \rightarrow 3 and 2- α -Araf \rightarrow 5 linkage, respectively. Spin systems (d111.1 and d5.06 ppm) assigned to 5-a-Araf was observed to overlap with a set of 3,5-a-Araf (d111.1-d5.02) linkages for

wild-type *C. glutamicum*-AG. The ^{13}C resonance at d109.4 ppm, which correlates to a proton at d5.05 ppm was designated as the 5,6- β -Gal f linkage. The well separated spin systems for 5- β -Gal f (d110.6, d5.18 ppm) and 6- β -Gal f (d111.4, d4.98) were visible in both wild type and *C. glutamicum* Δ *aftA* Δ *emb*-‘AG’ spectra. Overall, NMR analysis suggests that, as compared to wild-type AG, the ‘AG’ of the double mutant is lacking arabinan and only possesses unaltered cell wall galactan.

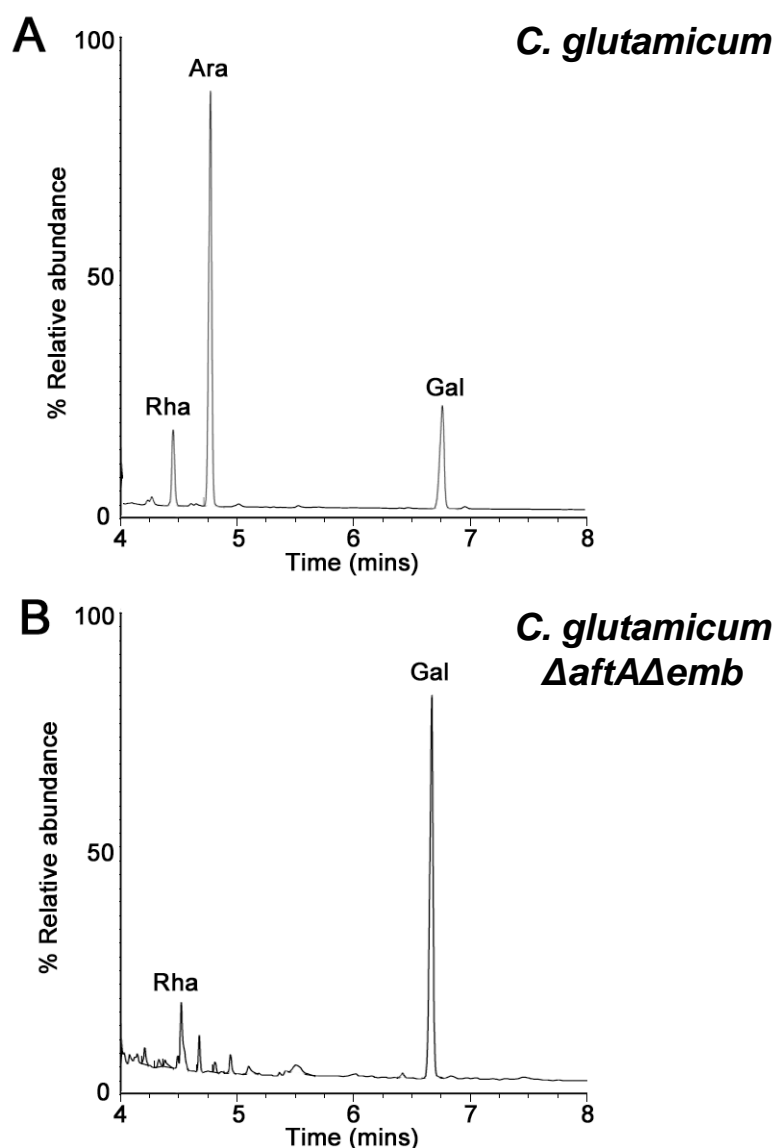


Figure 3 Glycosyl compositional analysis of cell walls (mAGP) of *C. glutamicum* and *C. glutamicum* $\Delta aftA \Delta emb$. Samples of purified cell walls were hydrolysed with 2 M trifluoroacetic acid, reduced with sodium borohydride and per-*O*-acetylated resulting in alditol acetates. Subsequently, samples were examined by gas chromatography as described in General Materials and Methods 6.5.18.

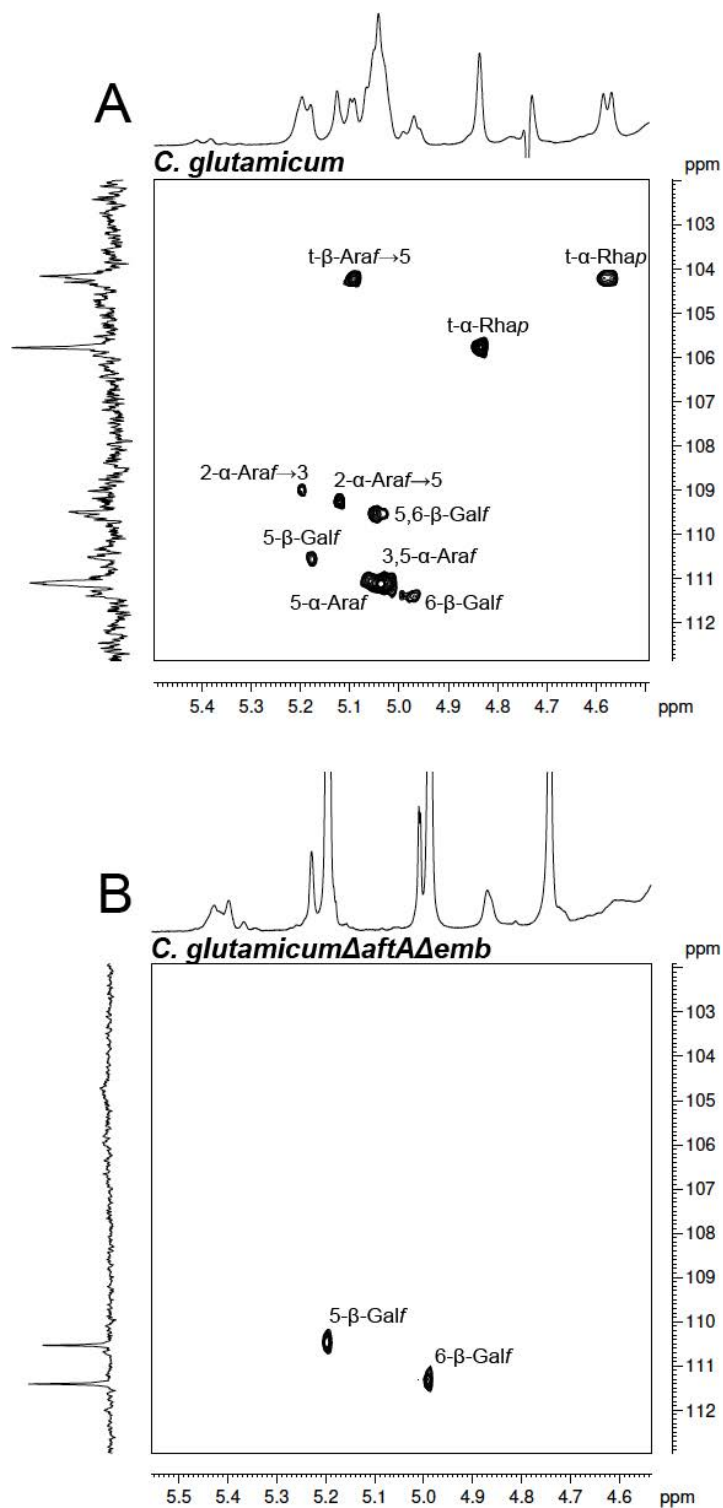


Figure 4 Two-dimensional NMR spectra of base solubilised AG from *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb*. Structural characterisation of wild type *C. glutamicum*-AG (A) and *C. glutamicum* Δ *aftA* Δ *emb*-AG (B). ^1H , ^{13}C HSQC NMR spectra were acquired in deuterium oxide at 313K. Expanded regions (δ ^1H : 4.5-5.5, δ ^{13}C : 102-113) are shown.

4.3.3 Lipoglycan analysis of *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb*

To date, only four ArafT - AftC, AftD, AftE and EmbC - have been demonstrated to play a role in LAM assembly of *Mycobacterium* (Goude *et al.*, 2008, Skovierova *et al.*, 2009, Birch *et al.*, 2010, Mishra *et al.*, 2012). Therefore, we wanted to examine the potential role of AftA and Emb in LAM biosynthesis. Extracted lipoglycans from wild type *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb* were analysed quantitatively by labelling cultures with [14 C]-glucose and loading equal amounts of radioactive lipoglycans to a 15 % SDS-PAGE. Extracts from both strains showed presence of [14 C]-LAM, -LM and -PIMs with no significant difference between the two samples.

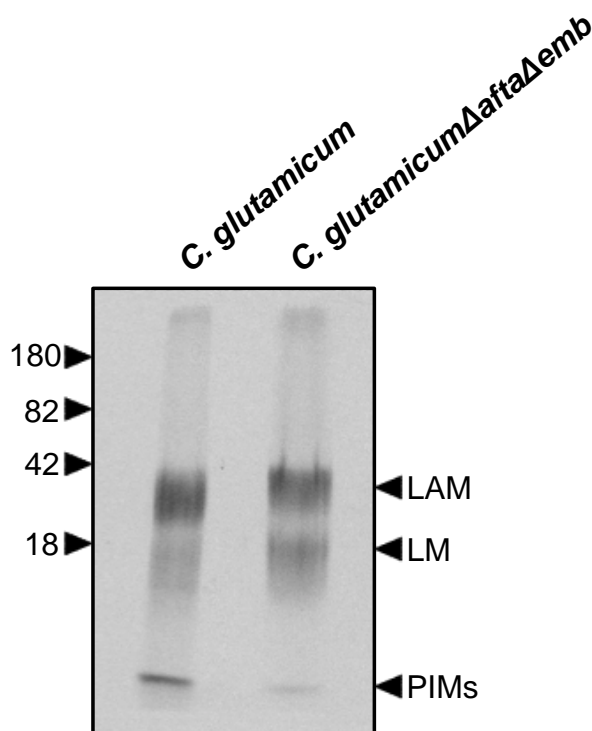


Figure 5 SDS-PAGE analysis of lipoglycans extracted from *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb*. Bacterial cells were refluxed with 50 % ethanol to extract crude lipoglycans. Samples were then subjected to hot phenol treatment and dialysed against water to yield pure lipoglycan fraction from both *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb*. Equivalent aliquots (20,000 cpm) from each strain were analysed by SDS-PAGE and visualised by autoradiography using Kodak BioMax MR film.

4.3.4 *In vitro* arabinofuranosyltransferase activity with extracts of *C. glutamicum*, *C. glutamicum* Δ *aftA*, *C. glutamicum* Δ *emb*, *C. glutamicum* Δ *aftB* Δ *aftD* and *C. glutamicum* Δ *aftA* Δ *emb*

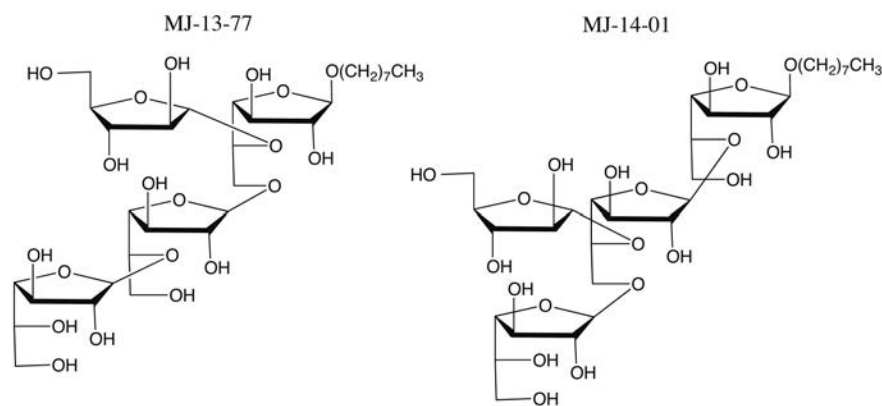


Figure 6 Chemical structures of neoglycolipid acceptors used in this study. MJ-13-77 (β -D-Galf-(1 \rightarrow 5)- β -D-Galf-(1 \rightarrow 6)-[α -D-Araf-(1 \rightarrow 5)]- β -D-Galf-O-(CH₂)₇CH₃) and MJ-14-01 (β -D-Galf-(1 \rightarrow 6)-[α -D-Araf-(1 \rightarrow 5)]- β -D-Galf-(1 \rightarrow 5)- β -D-Galf-O-(CH₂)₇CH₃).

Initial attempts to develop an *in vitro* assay using purified recombinant Emb from *C. glutamicum* have thus far proved unsuccessful. As an alternative, to catalyse ArafT activity in the presence of exogenous neoglycolipid acceptors, MJ-13-77 and MJ-14-01 (Figure 6), was assessed using membrane preparations from *C. glutamicum*, *C. glutamicum* Δ *aftA*, *C. glutamicum* Δ *emb*, *C. glutamicum* Δ *aftB* Δ *aftD* and *C. glutamicum* Δ *aftA* Δ *emb*. The cell free assay format was based on previously established ArafT assays (Seidel *et al.*, 2007, Birch *et al.*, 2008) and was designed to directly examine whether Emb enzyme from *C. glutamicum* acts as an (1 \rightarrow 5) ArafT and transfers Araf residue to the primed galactan chain. Three pentasaccharide products (A, B and C) were synthesised in the presence of synthetic acceptors, when assayed with wild type *C. glutamicum* membranes (Figure 7A). Products A and C were likely to be formed as a

result of $\alpha(1\rightarrow5)$ Araf linkages added to the MJ-13-77 and MJ-14-01 acceptors, respectively. Whereas B pentasaccharide was possibly synthesised by adding either $\alpha(1\rightarrow2)$ or $\alpha(1\rightarrow3)$ Araf linkage to MJ-13-77. Addition of EMB at concentrations of up to 1 mg/ml to the reaction mixture, resulted in an inhibition of A and C product formation, suggesting that the enzyme(s) adding Araf residue onto the acceptors are EMB sensitive, such as Emb. A comparable experiment was repeated using *C. glutamicum* Δ aftA membranes and neoglycolipid acceptors, MJ-13-77 and MJ-14-01, where a similar profile to that of wild-type *C. glutamicum* was observed (Figure 7B). Hence, demonstrating that AftA ArafT is highly unlikely to be involved in addition of Araf residue to either synthetic acceptor. Another experiment utilising *C. glutamicum* Δ emb membranes and synthetic acceptors was performed where both A and C products were not formed (Figure 7C), suggesting that the deleted enzyme Emb is indeed responsible for the generation of these two products. Interestingly, *C. glutamicum* Δ emb strain was still able to produce product B utilising MJ-13-77. Addition of EMB to the reaction mixture did not inhibit the synthesis of product B, therefore, implying that ArafT adding the Araf residue to the MJ-13-77 is EMB insensitive. An *in vitro* experiment was completed using *C. glutamicum* Δ aftB Δ aftD membranes where only A and C products were synthesised (Figure 7D). The data clearly show that AftB or AftD possessing $\alpha(1\rightarrow2)$ and $\alpha(1\rightarrow3)$ ArafT activity, respectively, are most likely to be responsible for the formation of product B. Finally, an *in vitro* ArafT assay was repeated employing *C. glutamicum* Δ aftA Δ emb membranes leading to formation of a single product B. This is consistent with the results obtained utilising *C. glutamicum* Δ emb membranes.

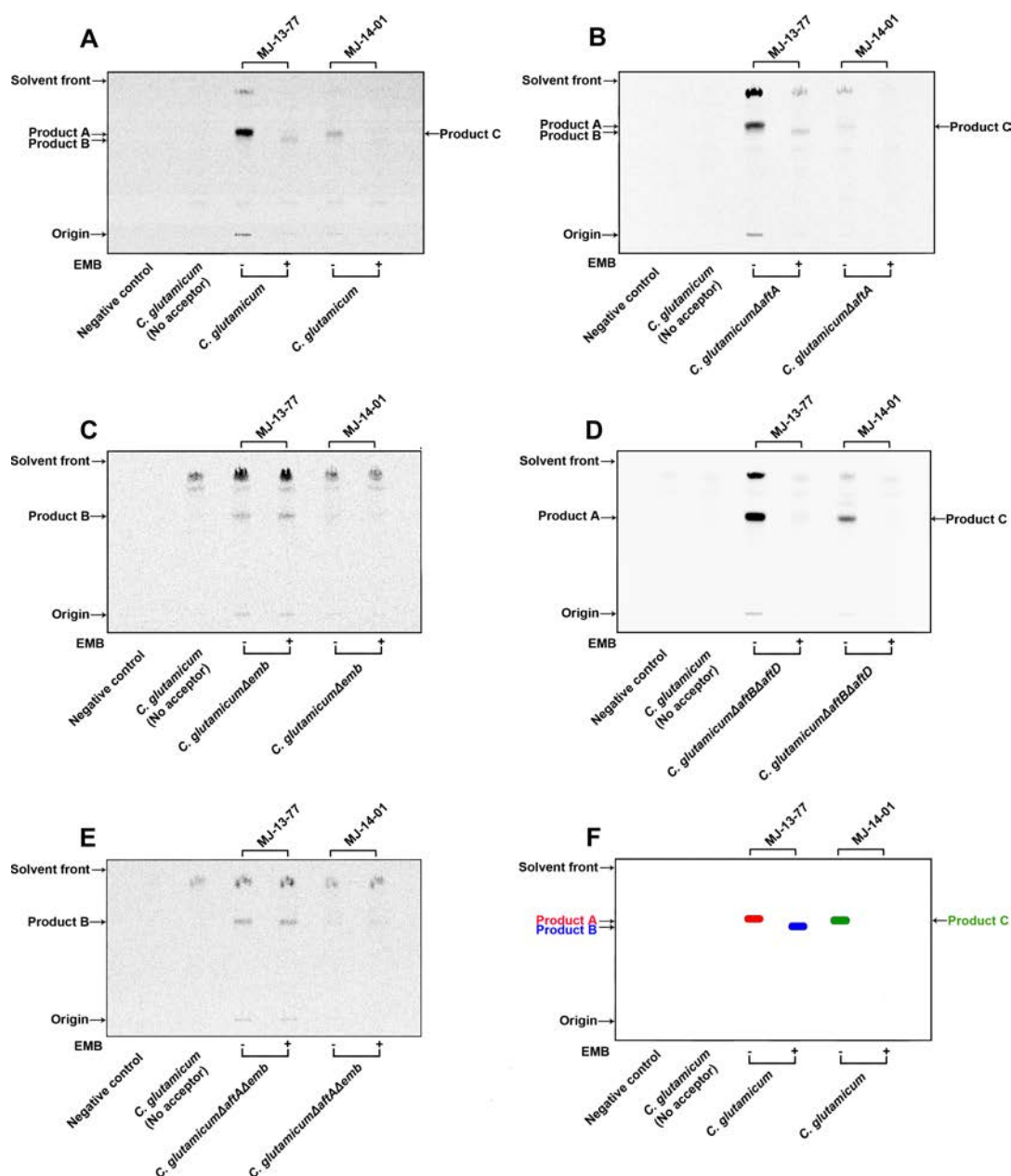


Figure 7 Arabinofuranosyltransferase activity assay utilising MJ-13-77 and MJ-14-01 neoglycolipid acceptors and membranes prepared from *C. glutamicum* (A), *C. glutamicum* Δ aftA (B), *C. glutamicum* Δ emb (C), *C. glutamicum* Δ aftB Δ aftD (D), *C. glutamicum* Δ aftA Δ emb (E). The representative diagram is (F). Arabinofuranosyltransferase activity was determined using synthetic acceptors in a cell-free assay with and without ethambutol (1mg/ml). The products of the assay were subjected to TLC using silica gel plates (5735 silica gel 60F254, Merck) in isopropanol/acetic acid/water (8/1/1, v/v/v) with the reaction products visualised by either autoradiography using Kodak BioMax MR films or phosphorimaging using Molecular Imager FX (Bio-Rad).

4.3.5 Characterisation of arabinosyltransferase reactions products

Products A, B and C, synthesised using synthetic acceptors MJ-13-77 and MJ-14-01, membranes from wild-type *C. glutamicum*, and in some cases EMB were excised from a preparative TLC plate, per-*O*-methylated and subjected to matrix-assisted laser desorption/ionisation-time-of-flight mass spectrometry (MALDI-TOF MS). To rule out co-chromatography of both the starting material and product even after separation by the TLC, both MJ-13-77 and MJ-14-01 acceptors were also analysed and served as a marker for any remaining unreacted acceptor (Figure 8A-B). All three samples containing A, B and C products have provided similar molecular ion profiles (Figure 8 C-E). Molecular ions of m/z 1099.5 were observed and corresponded to the formed pentasaccharide $[\text{Gal}_3\text{Ara}_2\text{-C}_8\text{H}_{17}+\text{Na}]^+$ (Figure 8C-E) containing newly added *Araf*-(1-?). Ions for unreacted MJ-13-77 and MJ-14-01 of m/z 939.5 $[\text{Gal}_3\text{Ara}_1\text{-C}_8\text{H}_{17}+\text{Na}]^+$ were also detected. In addition, all samples including spectra of acceptors contained molecular ions of m/z 969.5, suggesting a possible permethylation artefact. Overall, MALDI-TOF MS analysis confirms the addition of a single *Araf* residue to MJ-13-77 and MJ-14-01 resulting in products A, B and C with composition $\text{Gal}_3\text{Ara}_2\text{-C}_8\text{H}_{17}$.

In order to define the branching pattern of the generated oligomers, the respective ions were subjected to MALDI-TOF MS/MS analysis. The fragmentation patterns of synthetic acceptors and arabinan oligomers have been validated previously (Lee *et al.*, 2006, Zhang *et al.*, 2007, Shi *et al.*, 2008). The original $\text{Gal}_3\text{Ara}_1\text{-C}_8\text{H}_{17}$ acceptors provided the series of ions: $^{\text{O},3}\text{A}$, $^{2,4}\text{A}$, C, E, G, $^{\text{O},2}\text{X}$ and $^{1,4}\text{X}$ (Figure 9A), consistent with expected linkages and cleavages (Figure 9B-C). In the case of product A and B, both provided similar MS/MS spectra with common C and E ions at m/z 463 and 415, respectively, indicating

that in both cases *Araf* residue have been added to either *Araf* or *Galf* residue at the non-reducing end of MJ-13-77 acceptor. Lack of arabinosylation at the C-2 of the *Araf* residue of MJ-13-77 is supported by the absence of ions $^{2,4}A$ and $^{0,2}X$ of m/z 125 and 981, respectively. In addition, the Y ion of m/z 765, coupled with ions at m/z 285, 327 and 863, locate the *Araf* residue at either C-3 or C-5 of the *Araf* unit of the acceptor in both product A and product B. However, ions (m/z 97 and 1025) that would indicate if the residue was attached to the C-3 were not exclusive to both spectra, whereas ions (m/z 257, 865) that would locate the residue specifically to C-5 were not detected most likely due to low abundance. In contrast, for the product C (Figure 11), the Y, $^{0,3}A$ and $^{2,4}A$ ions at m/z 335, 665 and 693, respectively, indicated the newly attached *Araf* residue linked to either *Galf* or *Araf* residue at the non-reducing end of the MJ-14-01 acceptor, whereas the ion of m/z 375 clearly places the *Araf* residue to the $\alpha(1\rightarrow5)$ linked *Araf* of the acceptor. More importantly, the $^{0,3}A$, $^{0,2}X$, E and G ions with m/z of 257, 301, 343 and 863, respectively, defines the *Araf* residue attached to no other position but C-5 (Figure 11).

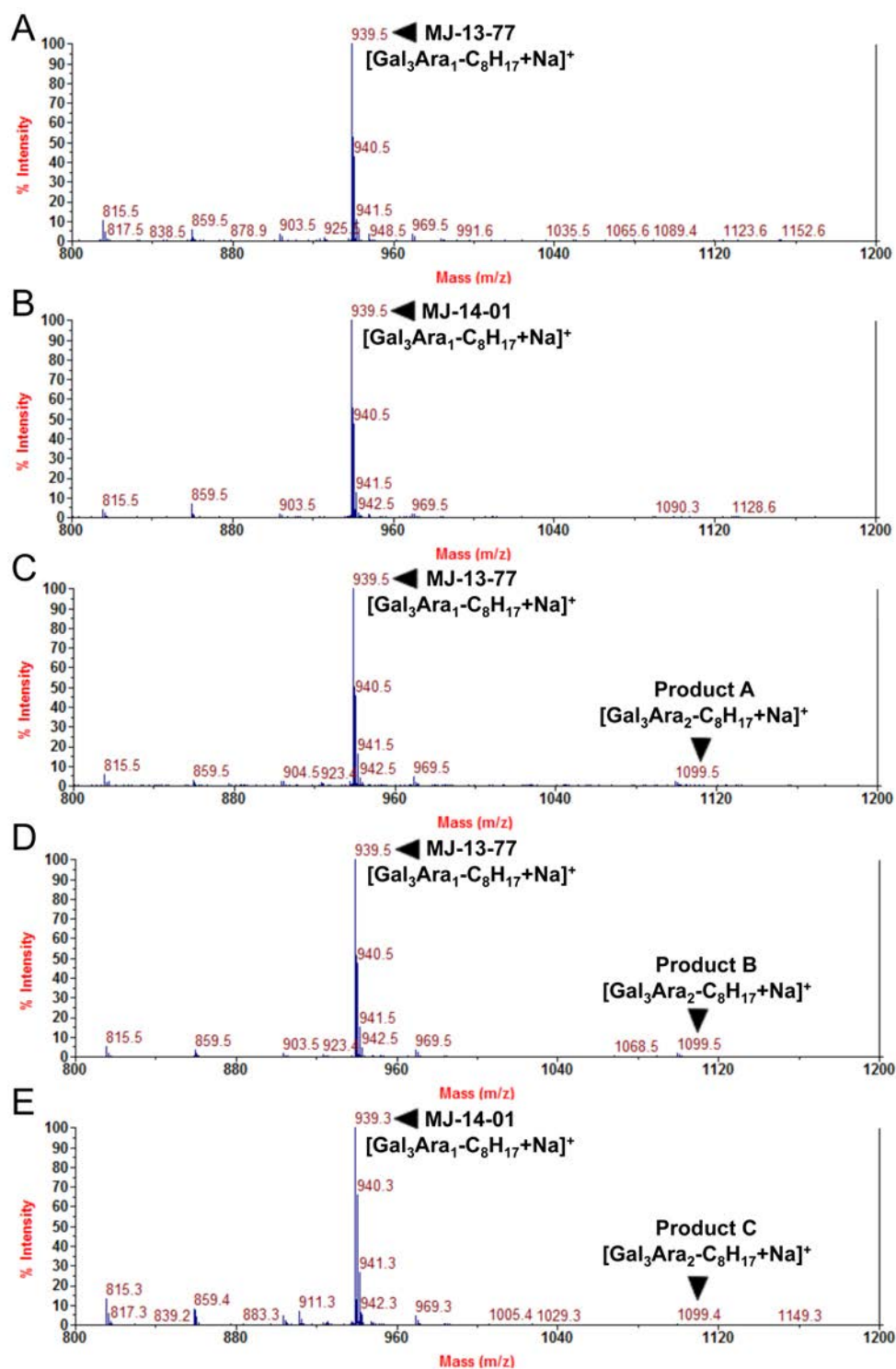


Figure 8 MALDI-TOF MS analysis of acceptors and enzymatic products. Both MJ-13-77 (A) and MJ-14-01 (B) were dried, per-*O*-methylated and subjected to MALDI-TOF MS. (C-E) Products formed utilising Gal₃Ara₁-C₈H₁₇ synthetic acceptors and wild-type *C. glutamicum* membranes were extracted from the preparative TLC and samples prepared for MALDI-TOF MS.

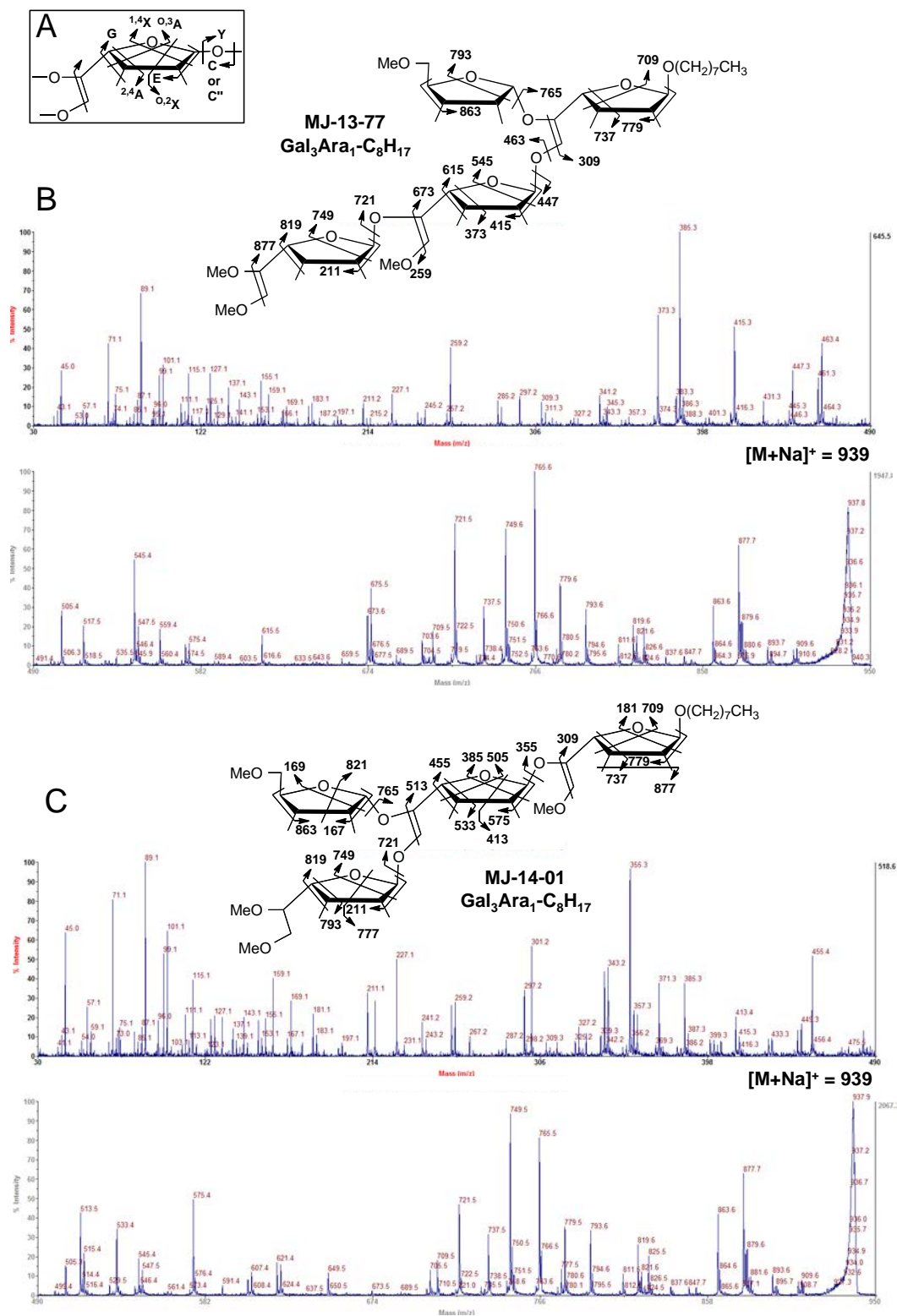


Figure 9 MALDI-TOF MS/MS analysis of per-*O*-methylated MJ-13-77 and MJ-14-01 acceptors. **A** The fragmentation pattern and adopted nomenclature of oligosaccharides. MS/MS spectra related to the expected cleavages and linkages of synthetic acceptors MJ-13-77 (**B**) and MJ-14-01 (**C**).

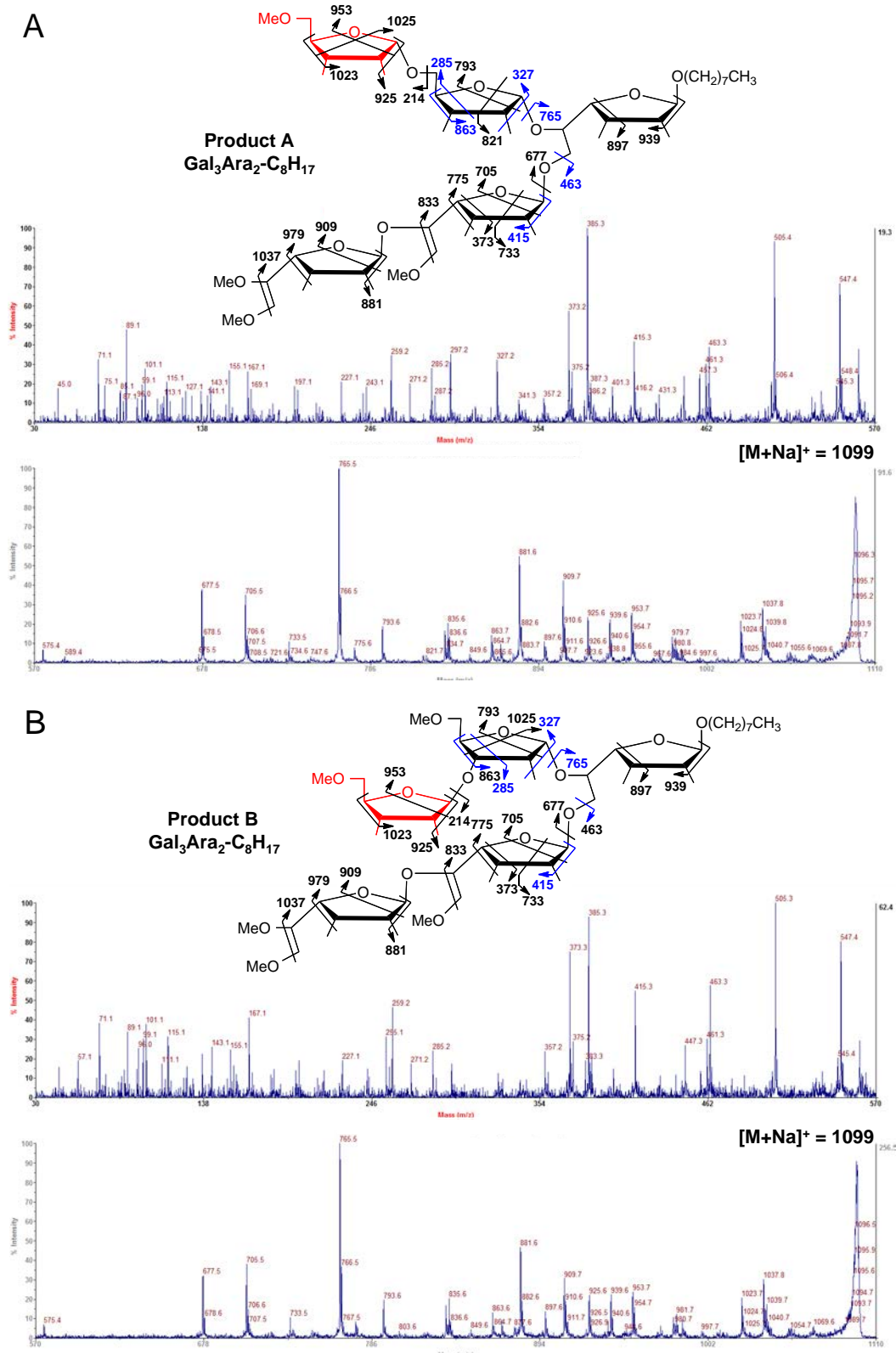


Figure 10 MALDI-TOF MS/MS analysis of per-*O*-methylated product A (A) and product (B). The added Ara_f residue is depicted in red, whereas cleavages and linkages that are diagnostic for determining the structure are depicted in blue.

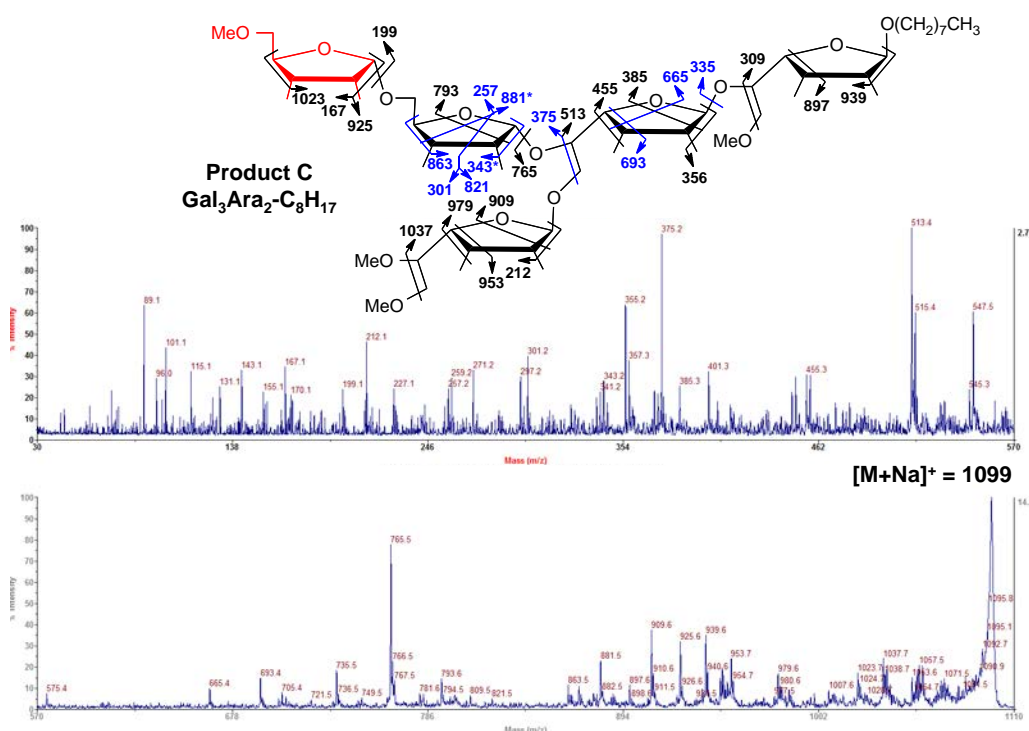


Figure 11 MALDI-TOF MS/MS analysis of per-*O*-methylated product C. The added *Araf* residue is depicted in red, whereas cleavages and linkages that are diagnostic for determining the structure are depicted in blue.

4.4 Discussion

The mAGP complex is a key cell wall component of the *Corynebacteriaceae*, which is essential for the growth and viability of *M. tuberculosis*. This highly complex structure is rich in lipids and sugars that act as a low permeability barrier and contribute to resistance to common antibiotics. It is therefore not surprising that potent anti-TB drugs target the biosynthesis of mAGP, including EMB and INH that inhibit the biosynthesis of AG and mycolic acids, respectively. Nevertheless, cases of drug resistant TB pose a challenge to treatments offered by the chemotherapy, therefore, increasing the need to discover novel drug targets and develop new therapeutic agents against mycobacterial infections. Our current knowledge of *Araf*Ts that synthesise arabinan domains of both AG and LAM

remains somewhat incomplete. For instance, the catalytic mechanisms of how these ArafTs synthesise the bulk of arabinan containing $\alpha(1\rightarrow5)$, $\alpha(1\rightarrow3)$ and $\alpha(1\rightarrow2)$ glycosidic linkages in mycobacteria and corynebacteria remain unknown. Structural and functional information of these enzymes would aid the further exploitation of ArafTs as potential drug targets to inhibit the crucial mAGP complex in *M. tuberculosis*.

In this study, we have characterised a *C. glutamicum* Δ *aftA* Δ *emb* mutant. Analysis of its cell wall revealed AG containing only the galactan domain with no arabinose (Figure 4). In addition, the double deletion of *aftA* and *emb* also led to an increase in non-covalently linked TDCMs and TMCs and a loss of covalently bound cell wall mycolates indicating the loss of mycolylation sites in AG (Figure 2). Individual disruptions of *aftA* and *emb* in *C. glutamicum* obtained previously led to analogous effects – complete absence of AG arabinan and a highly truncated AG with only terminal Araf residues, respectively (Alderwick *et al.*, 2005, Alderwick *et al.*, 2006). In contrast, individual deletions of *embA* and *embB* in *M. smegmatis* (Escuyer *et al.*, 2001) resulted in highly arabinosylated AG with alterations present only in Ara₆ motif. Since previous attempts to construct a double deletion of *embA* and *embB* in *M. smegmatis* were unsuccessful, it is tempting to speculate that EmbA or EmbB may compensate for each other in the individual knockouts in *M. smegmatis*.

In addition, we examined the potential role of AftA and Emb enzymes in LAM biosynthesis by isolating lipoglycans (PIMs, LM and LAM) from both wild type *C. glutamicum* and *C. glutamicum* Δ *aftA* Δ *emb* strains. Since individual inactivation of either AftA or Emb led to significant arabinan loss in AG, deletion of both ArafTs were anticipated to greatly affect both LM and LAM biosynthesis. Surprisingly, lipoglycan

extractions from both wild type and double mutant strains showed similar profiles with no significant difference (Figure 5), implying that neither AftA, nor Emb play a role in the synthesis of LM/LAM in *C. glutamicum*. Nevertheless, the LAM structure is less elaborate in *C. glutamicum* than compared to *M. smegmatis* or *M. tuberculosis* (Mishra, 2011), therefore, it is possible that AftA may still contribute to biosynthesis of LM/LAM in *Mycobacterium*. Previous studies demonstrated that disrupting the synthesis of the only arabinose donor in AG - DPA - through inactivation of UbiA led to a total loss of arabinose (Alderwick *et al.*, 2005). However, *C. glutamicum::ubiA* mutant still produced variants of LAM containing Araf residues, indicating the presence of yet uncharacterised arabinose sugar donor and putative ArafTs that employ it (Tatituri *et al.*, 2007). Notably, AftE has been characterised as an ArafT, which attaches singular Araf residues onto the LAM intermediate in *C. glutamicum* (Mishra *et al.*, 2012). Deletion of *C. glutamicum* AftE resulted in a truncated LAM structure with complete absence of arabinose, but extended mannosylation. One could argue that AftE may employ the uncharacterised Araf donor for the biosynthesis of the arabinan domain in LAM in *C. glutamicum*.

Although, previously described ArafTs including Emb proteins (Escuyer *et al.*, 2001, Alderwick *et al.*, 2005), AftA (Alderwick *et al.*, 2006), AftB (Seidel *et al.*, 2007), AftC (Birch *et al.*, 2008) and AftD (Skovierova *et al.*, 2009) share certain functional relationship, these enzymes are distinct ArafTs with their own individual features. A classic example is the sensitivity of Emb proteins from *Mycobacterium* species and Emb from *C. glutamicum* towards EMB, but insensitivity of other ArafTs including AftA, AftB and AftC towards the same drug. The number of ArafTs involved in the assembly of arabinan domain in AG remains a matter of speculation, however, the current structure of

AG suggests at least six different ArafTs that are involved in its synthesis. Since Emb proteins are the targets of EMB, a number of studies attempted to characterise them. Escuyer et al. (2001) constructed individual EmbA and EmbB mutants in *M. smegmatis* that possessed reduced levels of the disaccharide β -D-Araf-(1 \rightarrow 2)- α -D-Araf in AG resulting in a terminal non-reducing Ara₄ motif instead of the usual Ara₆. It was concluded that EmbA and EmbB are responsible for the 3,5 branching in AG and the synthesis of a distinct disaccharide β -D-Araf-(1 \rightarrow 2)- α -D-Araf (Escuyer *et al.*, 2001). However, taking into account the identification of AftB, AftC and AftD ArafTs, development of *in vitro* studies in combination with mutants and with structural and mutational studies, has promoted the view that EmbA and EmbB possess α (1 \rightarrow 5) ArafT activity.

In this study, we show that Emb from *C. glutamicum* acts as α (1 \rightarrow 5) ArafTs as demonstrated through the use of *in vitro* cell free assays in conjunction with *C. glutamicum* mutant strains. Specifically, we show that incubation of membranes prepared from wild type *C. glutamicum* with pRpp (an indirect donor of Araf residues that is converted to DPA *in situ*) and β -D-Galf-(1 \rightarrow 6)-[α -D-Araf-(1 \rightarrow 5)]- β -D-Galf-(1 \rightarrow 5)- β -D-Galf (MJ-13-77) neoglycolipid acceptor resulted in the synthesis of products A and B (Gal₃Ara₂-C₈H₁₇) (Figure 7). Further characterisation of the products by mass spectrometry established that the acceptor was extended by addition of Araf to α -D-Araf-(1 \rightarrow 5) at either C-3 or C-5 resulting in a further α (1 \rightarrow 5) extension (Figure 10). Although the precise attachment site of the Araf residue was not determined by mass spectrometry, the exploitation of the Emb target EMB and *C. glutamicum* mutant strains assisted in determining the precise location of the *in vitro* attached Araf. Addition of EMB to the

reaction mixtures containing wild type *C. glutamicum* membranes inhibited the formation of product A, thus suggesting that the enzyme adding *Araf* residue onto the acceptor is EMB sensitive, such as Emb. In contrast, the generation of pentasaccharide B was observed even in the presence of EMB, indicating that the *Araf*Ts in action is EMB insensitive, for example, AftA, AftB, AftC or AftD (Figure 7). In addition, reaction mixtures containing membranes lacking Emb activity did not produce product A, indicating that the deleted enzyme Emb *Araf*T is indeed responsible for its formation. *In vitro* assay employing *C. glutamicum* Δ *aftB* Δ *aftD* membranes and β -D-Galf-(1 \rightarrow 6)-[α -D-*Araf*-(1 \rightarrow 5)]- β -D-Galf-(1 \rightarrow 5)- β -D-Galf (MJ-13-77) synthetic acceptor did not form product B implying that either AftB or AftD is responsible for addition of *Araf* to the acceptor. However, considering TLC and MS analysis together, it is clear that pentasaccharide A is formed due to α (1 \rightarrow 5) activity of Emb, whereas pentasaccharide B is generated by AftD acting as an α (1 \rightarrow 3) *Araf*T. Finally, formation of product C was only identified in cell free assays containing membranes with Emb enzymatic activity, p[14 C]Rpp and β -D-Galf-(1 \rightarrow 6)-[α -D-*Araf*-(1 \rightarrow 5)]- β -D-Galf-(1 \rightarrow 5)- β -D-Galf (MJ-14-01) (Figure 7). Further analysis by mass spectrometry confirmed that addition of *Araf* to the MJ-14-01 neoglycolipid acceptor is attached at the position C-5 (Figure 11). Therefore, establishing for the first time that Emb from *C. glutamicum* acts as α (1 \rightarrow 5) *Araf*T.

The discovery of Emb from *C. glutamicum* as α (1 \rightarrow 5) *Araf*T has shed the light on the key *Araf*T involved in the synthesis of AG, which may contribute to a more detailed understanding of pathogenicity and persistence of *M. tuberculosis*. The challenge for the

future will be to define the precise role of Emb enzymes from *Mycobacterium* species as well as the catalytic mechanisms of how these glycosyltransferases synthesise AG.

Chapter 5

Conclusions and Future work

5 Conclusions and Future work

The threat from the rising emergence of drug resistant TB cases worldwide has renewed interest and improved funding into research of the pathogenicity of its causative agent, *M. tuberculosis*. Improved chemotherapeutic treatments and novel drugs that would shorten the duration of TB therapy, would encourage patients to complete their drug regimens and perhaps reduce the trend of drug resistant TB cases. However, the development of new and effective drugs requires a better understanding of the pathogen itself. The cell envelope of *M. tuberculosis* is unique and vital for its persistence and virulence due to the characteristic cell envelope structures, such as AG, LAM and mycolic acids. For many years, proteins involved in the biosynthesis of these cell wall components have been attractive targets for the development of potential anti-TB drugs. The majority of these enzymes are essential for the growth of *M. tuberculosis* and more importantly lack homologues in the mammalian systems, thus forming a focal point for drug discovery. This thesis has focused on characterising genes and their encoded proteins involved in vital AG and LAM biosynthetic processes with the view of discovering new drug targets.

Networks of interacting proteins mediate pathways of biological processes varying from signal transduction to disease pathogenesis and drug resistance (Cui *et al.*, 2009, Wang *et al.*, 2010, Lee *et al.*, 2012). Formation of large multi-protein complexes in bacteria is not uncommon and has been fairly well described in *E. coli*. Bacterial cell division depends on the formation, positioning and contraction of the Z-ring, which in *E. coli* is composed of at least ten different proteins, and includes FtsZ, FtsA, FtsB, FtsI, FtsK, FtsL, FtsN and FtsQ (Buddelmeijer and Beckwith, 2004, Wadenpohl and Bramkamp, 2010, Robichon *et al.*, 2011, Lutkenhaus *et al.*, 2012, Marteyn *et al.*, 2014). In addition, a number of Zap

accessory proteins (ZapA, ZapB, ZapC, ZapD and ZapE) have been shown to contribute to the formation of the Z-ring (Galli and Gerdes, 2010, Hale *et al.*, 2011, Durand-Heredia *et al.*, 2012). These proteins, acting as a large multi-protein complex, which drive cell membrane invagination, inward growth of the cell wall, and, ultimately promote separation of the progenitor cell into two independent daughter cells. However, only very few studies have been conducted to examine protein-protein interactions of enzymes involved in mycobacterial cell wall biosynthesis (Zheng *et al.*, 2011, Larrouy-Maumus *et al.*, 2012, Veyron-Churlet *et al.*, 2004, Veyron-Churlet *et al.*, 2005). In the second chapter, we employed *in vitro* pull-down assays and *in vivo* BACTH system to elucidate the protein-protein interaction network involved in the *C. glutamicum* major cell wall component AG. We have identified twenty-four putative homotypic and heterotypic protein interactions. Our studies demonstrated that WecA and WbbL, proteins that are responsible for the formation of the AG linker unit, interact with decaprenyl-5-phosphoribose synthase UbiA. In addition, WecA and UbiA show evidence for physical interaction between AftB and AftC proteins. Formation of protein complexes was also observed between AftB and the priming enzyme AftA, branching enzyme AftC and arabinan elongation enzyme Emb. Finally, a number of proteins involved in AG biosynthesis were shown to form either dimers or multimers. Our findings demonstrated that enzymes involved in *C. glutamicum* cell wall assembly and precursor formation form sophisticated multi-protein complexes in order to maintain the efficiency and fidelity of AG polymerisation. These findings provide a useful recourse for understanding the biosynthesis and function of the corynebacterial cell wall, as well as providing putative therapeutic targets.

In the third chapter we explored the potential role of AftA and AftB in the biosynthesis of LM/LAM in *M. smegmatis* using CESTET and analysing the cellular composition of the mutants. We also examined full-length AftA and AftB membrane proteins and their C-termini in order to biochemically characterise them. As previously reported by Shi *et al* (2008), we confirmed that AftA is essential to *M. smegmatis* viability and growth. Depletion of AftA in the *M. smegmatis* Δ *aftA*::pMV306-*aftA* strain resulted in a clear lysed cell culture. However, an often drawback of CESTET was observed, where conditional mutant strains acquire mutations in the acetamide promoter region causing continuous production of the protein of interest, thus rescuing the cell from death. Since such persisters mask the effects of the AftA depletion, a careful selection from generated transductants has to be performed in order to segregate true *aftA* conditional mutants in *M. smegmatis*. In order to acquire knowledge on the key ArafT conserved within *Corynebacterineae*, we expressed and purified the hydrophilic C-terminal domain of AftA from *M. tuberculosis*. Attempts to stabilise the recombinant protein employing thermostability screening led to sufficient amount of AftA^{CT}, which was subsequently used for crystallisation screens. In addition, we have employed a recently reported SMALP method to solubilise and purify full-length AftA from *C. glutamicum*. Utilising techniques, such as AUC and CD, we have demonstrated that AftA within SMALP is properly folded and stable. Ligand binding employing various lengths of synthetic acceptors resembling the galactan chain indicated that AftA has a markedly stronger affinity for longer acceptors.

Similarly to AftA, in the third chapter we have demonstrated that *aftB* is an essential gene in *M. smegmatis*. Lipid analysis from the AftB depleted cells demonstrated an increase in cell wall associated lipids TDMs and TMMs and a significant decrease in all three types

of mycobacterial MAMES, suggesting a cell wall defect. SDS-PAGE and sugar analyses of lipoglycans extracted from both parental and conditional mutant strains did not exhibit significant difference between each other, however a monoclonal antibody F30-5 generated against hexa-arabinan motif of LAM no longer recognised LAM extracted from the *aftB* mutant, thus suggesting a loss of terminal $\beta(1\rightarrow2)$ Ara_f residues. Subsequent NMR or glycosidic linkage analysis is required to identify the precise changes of LAM structure resulted upon depletion of AftB. Expression of the full length AftB and its hydrophilic C-terminal domain from *M. tuberculosis* and *C. glutamicum* were also examined, however, this did not yield any soluble recombinant proteins and therefore were not pursued further.

In chapter four, we aimed to characterise the target of front-line drug EMB in *C. glutamicum* - Emb Ara_fT. For that purpose, we constructed and examined effects of an *aftA* and *emb* double deletion on AG biosynthesis. Cell wall analysis revealed a complete absence of arabinose resulting in a truncated cell wall containing only galactan accompanied with the loss of cell wall bound mycolates. In addition, we explored the potential role of AftA and Emb enzymes in *C. glutamicum* LAM biosynthesis, since individual inactivation of either *C. glutamicum* *aftA* or *emb* resulted in a significant arabinan loss in AG (Alderwick *et al.*, 2005, Alderwick *et al.*, 2006). However, lipoglycan extractions from both wild type and *aftA* and *emb* double deletion showed similar profiles of no significant differences, indicating that both enzymes are not involved in LAM biosynthesis in *C. glutamicum*. Finally, by employing a cell free assay using *C. glutamicum*, *C. glutamicum* Δ *aftA*, *C. glutamicum* Δ *emb*, *C. glutamicum* Δ *aftB* Δ *aftD*, *C. glutamicum* Δ *aftA* Δ *emb* strains, neoglycolipid acceptors that resemble primed galactan and analysing the products formed *in vitro*, we demonstrated

that the transfer of Ara f residue from DPA to the primed galactan in *C. glutamicum* is catalysed by Emb Ara f T with $\alpha(1\rightarrow5)$ activity. The findings obtained in the fourth chapter has shed the light on the key Emb Ara f T responsible for the elongation of arabinan domain of AG and the target of EMB in *C. glutamicum*. A clearer knowledge of structure and function of Ara f Ts is key for the further exploitation of putative inhibitors to disrupt the essential mAGP complex in pathogenic species, such as *M. tuberculosis*.

Chapter 6

General Materials and Methods

6 General Materials and Methods

6.1 Chemicals and reagents

All chemicals, solvents and restriction enzymes were purchased from Sigma-Aldrich (UK), Bio-Rad (USA), Thermo Fischer Scientific (USA) or New England BioLabs (USA) unless otherwise stated, and were of highest reagent grade.

6.2 Culture media preparations

6.2.1 Luria-Bertani broth

25 g of Luria-Bertani (LB) powder (10 g/l tryptone, 5 g/l yeast extract and 10 g/l sodium chloride; DifcoTM) was dissolved in 1 l of deionised water and sterilised by autoclaving at 121 °C for 15 minutes.

6.2.2 Luria-Bertani agar

37 g of LB agar powder (10 g/l tryptone, 5 g/l yeast extract, 10 g/l sodium chloride and 15 g/l agar; DifcoTM) was dissolved in 1 l of deionised water and sterilised by autoclaving at 121 °C for 15 minutes.

6.2.3 Brain Heart Infusion broth

37 g of Brain Heart Infusion (BHI) powder (5 g/l beef heart, 12.5 g/l calf brains, 2.5 g/l disodium hydrogen phosphate, 2 g/l glucose, 10 g/l peptone and 5 g/l sodium chloride; DifcoTM) was dissolved in 1 l of deionised water and sterilised by autoclaving at 121 °C for 15 minutes.

6.2.4 Brain Heart Infusion agar

37 g of BHI agar powder (5 g/l beef heart, 12.5 g/l calf brains, 2.5 g/l disodium hydrogen phosphate, 2 g/l glucose, 10 g/l peptone, 5 g/l sodium chloride and 15 g/l agar; Difco™) was dissolved in 1 l of deionised water and sterilised by autoclaving at 121 °C for 15 minutes.

6.2.5 Tryptic soy broth

30 g of Tryptic soy broth (TSB) powder (17 g/l casein peptone, 2.5 g/l dipotassium hydrogen phosphate, 2.5 g/l glucose, 5 g/l sodium chloride, 3 g/l soya peptone; Difco™) was dissolved in 1 l of deionised water and sterilised by autoclaving at 121 °C for 15 minutes.

6.2.6 Tryptic soy agar

40 g of Tryptic soy agar (TSA) powder (15 g/l casein peptone, 5 g/l sodium chloride, 5 g/l soya peptone and 15 g/l agar; Difco™) was dissolved in 1 l of deionised water and sterilised by autoclaving at 121 °C for 15 minutes.

6.2.7 7H9 basal agar

4.7 g of 7H9 powder (0.5 g/l ammonium sulphate, 2.5g/l disodium phosphate, 1 g/l monopotassium phosphate, 0.1 g/l sodium chloride, 0.05 g/l magnesium sulphate, 0.0005 g/l calcium chloride, 0.001 g/l zinc sulphate, 0.001 g/l copper sulphate, 0.04 g/l ferric ammonium citrate, 0.5 g/l glutamic acid, 0.001 g/l pyridoxine and 0.0005 g biotin; Difco™), 15 g/l of agar, and 2.5 ml of glycerol was dissolved in 1 l of deionised water and sterilised by autoclaving at 121 °C for 15 minutes.

6.2.8 7H9 soft agar

4.7 g of 7H9 powder (0.5 g/l ammonium sulphate, 2.5g/l disodium phosphate, 1 g/l monopotassium phosphate, 0.1 g/l sodium chloride, 0.05 g/l magnesium sulphate, 0.0005 g/l calcium chloride, 0.001 g/l zinc sulphate, 0.001 g/l copper sulphate, 0.04 g/l ferric ammonium citrate, 0.5 g/l glutamic acid, 0.001 g/l pyridoxine and 0.0005 g biotin; Difco™), 7.5 g/l bacto agar, and 2.5 ml glycerol was dissolved in 1 l of deionised water and sterilised by autoclaving at 121 °C for 15 minutes.

6.2.9 MacConkey agar

40 g of MacConkey agar powder (17 g/l pancreatic digest of gelatin, 3 g/l peptones, 10 g/l lactose, 1.5 g/l bile salts, 5 g/l sodium chloride, 13.5 g/l agar, 0.03 g/l neutral red and 1 mg/l crystal violet; Difco™) was dissolved in 1 l of deionised water and sterilised by autoclaving at 121 °C for 15 minutes.

6.2.10 CGXII minimal media

CGXII minimal media was prepared by mixing 20 g of ammonium sulphate, 5 g of urea, 1 g of potassium dihydrogen phosphate, 1 g of dipotassium hydrogen phosphate, 0.25 g of magnesium sulphate hepta hydrate, 42 g of MOPS (3-[N-Morpholino] propanesulfonic acid) in 800 mL of deionised water. The media was supplemented with 1 mL of trace elements solution (1 g of ferric sulphate heptahydrate, 1 g of manganese sulphate monohydrate, 0.1 g of zinc sulphate heptahydrate, 0.02 g of copper sulphate, 0.002 g of nickel chloride hexahydrate prepared in 90 mL deionised water, adjusted to pH 1 using concentrated hydrochloric acid), 1 mL calcium chloride solution (1 g in 100 mL deionised water), 50 µL biotin solution (10 mg in 100 mL deionised water), 8 mL

protocatechuate (0.3 g in 8 mL deionised water). The prepared salt media was adjusted to pH 7.0 using sodium hydroxide and its final volume adjusted to 920 mL. Media was filter sterilised and the 0.2-0.4 % carbon source of choice was added to CGXII medium just before use.

6.2.11 M63 minimal media

15.6 g of M63 minimal media powder (2 g/l ammonium sulphate, 13.6 g/l monobasic potassium phosphate, 0.5 mg/l ferrous sulphate heptahydrate; Amresco) was dissolved in 1 l of deionised water, pH adjusted to 7.0 using potassium hydroxide and media sterilised by autoclaving at 121 °C for 15 minutes. After sterilisation, 1 ml of vitamin B1 (5 mg/ml) and the 0.2-0.4 % carbon source was added to the media just before use.

6.2.12 M63 minimal agar

15.6 g of M63 minimal media powder (2 g/l ammonium sulphate, 13.6 g/l monobasic potassium phosphate, 0.5 mg/l ferrous sulphate heptahydrate; Amresco) and 15 g/l of agar was dissolved in 1 l of deionised water, pH adjusted to 7.0 with potassium hydroxide and media sterilised by autoclaving at 121 °C for 15 minutes. After sterilisation, 1 ml of vitamin B1 (5 mg/ml) and the 0.2-0.4 % carbon source was added to the media just before use.

6.2.13 Sauton's minimal media

Sauton's minimal media was prepared by mixing 0.5 g of dipotassium phosphate, 0.5 g of magnesium sulphate, 4 g of asparagine, 0.05 g of ferric ammonium citrate, 60 ml of glycerol and 2 g of citric acid in 1 l of deionised water. After the pH was adjusted to 7.2, the media was sterilised by autoclaving at 121 °C for 15 minutes. The prepared salt media

was supplemented with 1 % (w/v) zinc sulphate and 5 % Triton WR-1339.

6.2.14 Minimal media for *M. smegmatis*

Minimal media for the growth of *M. smegmatis* was prepared by mixing 1.5 g of potassium phosphate, 1 g of ammonium chloride, 0.2 g of magnesium sulphate, 20 mg of calcium chloride, 1.2 mg of ferric ammonium citrate, 0.85 g of sodium chloride and 8.99 g of disodium phosphate in 1 l of deionised water and sterilised by autoclaving at 121 °C for 15 minutes. The 0.2-0.4 % carbon source was added to the media just before use.

6.2.15 Minimal agar for *M. smegmatis*

Minimal agar for the growth of *M. smegmatis* was prepared by mixing 1.5 g of potassium phosphate, 1 g of ammonium chloride, 0.2 g of magnesium sulphate, 20 mg of calcium chloride, 1.2 mg of ferric ammonium citrate, 0.85 g of sodium chloride, 8.99 g of disodium phosphate and 15 g of agar in 1 l of deionised water and sterilised by autoclaving at 121 °C for 15 minutes. The 0.2-0.4 % carbon source was added to the media just before use.

6.3 Bacterial strains and growth conditions

6.3.1 General bacterial strains and growth conditions used in this thesis

M. smegmatis mc²155 strains were grown in TSB or minimal media (Chang *et al.*, 2009) supplemented with 0.05 % Tween-80 (Sigma Aldrich) at 37 °C. Strains of *C. glutamicum* ATCC 13032 were grown on rich BHI or the salt medium CGXII at 30 °C. Samples for lipid analysis, lipoglycan extractions and membrane preparations were prepared by harvesting cells at an OD₆₀₀ of 10-15 followed by subsequent washing using phosphate

buffered saline. Cell pellets were stored at -20 °C until further use. *E. coli* TOP10 (Invitrogen) strain used for plasmid propagation and *E. coli* strains used for protein overexpression including BL21 (F⁻ *ompT hsdS_B(r_B⁻ m_B⁻) gal dcm* (DE3), C43 (F⁻ *ompT gal dcm hsdS_B(r_B⁻ m_B⁻)*(DE3)), Rosetta pLysS (F⁻ *ompT hsdS_B(r_B⁻ m_B⁻) gal dcm* (DE3) pLysSRARE (Cam^R), and Tuner (F⁻ *ompT hsdS_B(r_B⁻ m_B⁻) gal dcm lacYI*(DE3)) were grown in LB media at 37 °C. The pLysS plasmid encoding T7 phage lysozyme (an inhibitor for T7 polymerase) in *E. coli* Rosetta cells was retained with chloramphenicol (25 µg/ml). The concentrations of supplements used were 25 µg/ml of kanamycin, 50 µg/ml of ampicillin, 100 µg/ml of hygromycin, and 0.2 % of acetamide with *M. smegmatis* and *C. glutamicum*; and 50 µg/ml of kanamycin, 100 µg/ml of ampicillin, 150 µg/ml of hygromycin with *E. coli*.

6.3.2 Bacterial strains and growth conditions used for BACTH

All cloning steps were performed in *E. coli* XL1-Blue cells (Invitrogen). XL1-Blue strain has a high level of LacI repressor expression, which is not encoded in pKT25, pKNT25, pUT18 and pUT18c plasmids and, therefore, might prevent accumulation of mutations during cloning procedure if the hybrid protein is toxic to *E. coli* cells. The *E. coli* *cya*⁻ strain BTH101 ((F⁻, *cya*-99, *araD*139, *galE*15, *galK*16, *rpsL*1 (Str^r), *hsdR*2, *mcrA*1, *mcrB*1) was used for the bacterial two-hybrid screen (Euromedex). Plasmids were maintained with ampicillin (100 µg/ml) or kanamycin (50 µg/ml). LB agar reporter plates contained streptomycin (100 µg/ml), ampicillin (100 µg/ml), kanamycin (50 µg/ml), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal; 40 µg/ml) and isopropyl β-D-1-thiogalactopyranoside (IPTG; 0.5 mM). MacConkey plates (DifcoTM) contained streptomycin (100 µg/ml), ampicillin (100 µg/ml), kanamycin (50 µg/ml), IPTG (0.5

mM) and maltose (1%). M63 minimal media plates (Battesti *et al.*, 2012) were supplemented with streptomycin (50 µg/ml), ampicillin (50 µg/ml), kanamycin (25 µg/ml), X-gal (40 µg/ml), IPTG (0.5 mM) and maltose (0.2%).

6.4 Antibiotics and supplements

Table 1 Antibiotics and supplements used in this thesis.

Additives	Concentration	Storage
Acetamide	20 %	at 4 °C, in water
Ampicillin	100 mg/ml	at -20 °C, in water
Chloramphenicol	30 mg/ml	at -20 °C, in ethanol
Hygromycin B	50 mg/ml	at 4 °C, in PBS, protect from light
Kanamycin	50 mg/ml	at -20 °C, in water
Nalidixic acid	30 mg/ml	at -20 °C, in ethanol
Streptomycin	100 mg/ml	at -20 °C, in water
Glucose	20 %	at 4 °C, in water
Maltose	20 %	at 4 °C, in water
Tween-80	10 %	at room temperature, protect from light
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	1 M	at -20 °C, in water
ONPG (o-nitrophenol-β-galactoside)	4 mg/ml	at -20 °C, in PM2 media, protect from light
X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside)	20 mg/ml	at -20 °C, in dimethyl formamide, protect from light

6.5 Molecular biology techniques

6.5.1 Polymerase chain reaction

Phusion DNA polymerase (New England BioLabs) have been employed in all polymerase chain reactions (PCR). Four different reaction mixtures, presented in Table 2, have been exploited to optimise DNA amplification. A temperature gradient PCR was used with two-step PCR procedure presented in Table 3.

Table 2 PCR mixtures used in this thesis.

Components	Concentration	Mix 1 (μ l)	Mix 2 (μ l)	Mix 3 (μ l)	Mix 4 (μ l)
Forward primer	100 pmol/ μ l	0.08	0.08	0.08	0.08
Reverse primer	100 pmol/ μ l	0.08	0.08	0.08	0.08
template DNA	10 ng/ μ l	0.4	0.4	0.4	0.4
dNTP	25 mM	0.4	0.4	0.4	0.4
5 x GC buffer		4	4	4	4
DNA polymerase	10 U/ml	0.2	0.2	0.2	0.2
MgSO ₄	50 mM	-	-	0.4	0.4
DMSO	100 %	-	1	-	1
H ₂ O	-	14.84	13.84	14.44	13.44
Total	-	20	20	20	20

Table 3 PCR cycling conditions.

PCR stage	Temperature	Time	Cycles
Initial denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	35
Annealing	60-75 °C	60 sec	35
Extension	72 °C	60 sec/kb	35
Final extension	72 °C	10 min	1
Hold	4 °C	∞	-

6.5.2 DNA gel electrophoresis

DNA fragments were separated employing agarose gel electrophoresis. Samples were mixed with DNA loading dye (6 x stock: 2.5 % Ficoll 400, 11 mM EDTA, 3.3 mM Tris-HCl, 0.017% SDS and 0.015 % bromophenol blue) and loaded onto 0.8-1 % agarose gel prepared using agarose powder (Bioline) and TAE buffer (50 x stock: 242 g Tris-base, 57.1 ml acetate and 100 ml 0.5 M EDTA). The loaded gel is run at 120–140 V, 400 mA for 45 minutes. DNA is visualised by ethidium bromide stain and analysed under UV light employing Bio-Rad GelDoc XR System.

6.5.3 DNA purification

DNA bands of interest were excised from the agarose gel and purified employing QIAquick Gel Extraction kit (Qiagen). Briefly, excised agarose fragments were dissolved in the provided QG buffer and subjected to a spin column with a silica membrane that binds DNA. Impurities such as salts, agarose and ethidium bromide are removed by several washing steps, whereas the pure DNA is eluted with deionised water.

6.5.4 DNA digestion

Plasmid DNA, PCR products or genomic DNA were single or double digested using a reaction mixture presented in Table 4. Reaction mixtures were incubated at 37 °C for 1-4 hours. After digestion, digestion enzymes were heat inactivated by incubating samples at 65 °C for 5 minutes.

Table 4 General reaction mixture of DNA digestion used in this thesis.

Component	Concentration	Volume (µl)
DNA	0.1-0.2 µg/µl	10
Digestion buffer	10 x	2
Digestion enzyme(s)	50 U/µl	1
BSA	10 x	2
Water	-	5
Total	-	20

6.5.5 DNA ligation

Digested DNA fragments with compatible ends were ligated using a reaction mixture presented in X and incubated at 37 °C for 10 minutes. After digestion, T4 DNA ligase was heat inactivated by incubating samples at 65 °C for 5 minutes.

Table 5 General reaction mixture of DNA ligation used in this thesis.

Component	Concentration	Volume (μl)
Insert DNA	0.020 pmol/ μ l	3
Vector DNA	0.020 pmol/ μ l	1
T4 DNA ligase	400000 U/ml	1
10 x T4 DNA ligase buffer	10 x	2
Water	-	5
Total	-	20

6.5.6 Plasmid DNA extraction

Plasmid DNA purification was performed employing QIAprep Spin Miniprep Kit (Qiagen). *E. coli* cells harbouring plasmid of interest were cultivated in 5 ml of LB media with appropriate antibiotics at 37 °C with shaking overnight. Cells were harvested, resuspended in 250 μ l of P1 buffer and transferred to a 1.5 ml eppendorf. Cell contents were released by addition of 250 μ l of P2 lysis buffer and neutralised by 350 μ l of N3. Subsequently, samples were centrifuged for 15 minutes at 13,000 g and the recovered supernatant containing the DNA subjected to QIAprep spin column. Sample was centrifuged for 30 seconds at 3,000 g and the flow-through discarded. Column was washed with 750 μ l of PE wash buffer before eluting the plasmid DNA with 30-50 μ l of EB buffer or deionised water.

6.5.7 Preparation of competent *E. coli* cells

A single colony of *E. coli* was inoculated into 5 ml LB medium and incubated at 37 °C with shaking overnight. The pre-culture was then used to inoculate 250 ml LB medium with 20 mM MgSO₄. Cells were grown to OD₆₀₀ of 0.4-0.6 and then harvested by centrifugation at 4,500 g for 5 minutes at 4 °C. Cell pellets were gently resuspended in 0.4 volume of ice-cold TFB1 (30 mM potassium acetate, 10 mM CaCl₂, 50 mM MnCl₂, 100 mM RbCl, 15 % glycerol, filter-sterilised and stored at 4 °C) and resuspended cells

were incubated on ice for 5 minutes. Cells were harvested by centrifugation as described previously and were then gently resuspended in 1/25 of the volume of ice-cold TFB2 (10 mM MOPS pH 6.5, 75 mM CaCl₂, 10 mM RbCl, 15 % glycerol, filter-sterilised and stored at 4 °C). Cells were incubated on ice for 15-60 minutes, shock-frozen and stored as 50 µl aliquots at -80 °C.

6.5.8 Preparation of competent *C. glutamicum* cells

A single colony of *C. glutamicum* was inoculated into 5 ml BHI medium and incubated at 30 °C with shaking overnight. The pre-culture was then used to inoculate 100 ml BHI medium and cultivated at 30 °C with shaking until an OD₆₀₀ reached 1.75. Cells were harvested by centrifugation at 4,500 g for 10 minutes at 4 °C. Cell pellets were gently resuspended in 20 ml of ice-cold TG buffer (1 mM Tris-HCl pH 7.5, 12 g 87 % glycerol per 100 ml deionised water, filter-sterilised and stored at 4 °C). Cells were harvested by centrifugation at 4,500 g for 10 minutes at 4 °C and washed twice more in TG buffer. Cell pellet was washed once in ice-cold 10 % glycerol solution and resuspended to a final volume of 2 ml of 10 % glycerol. Aliquots of 100 µl were dispensed, shock-frozen and stored at -80 °C.

6.5.9 Preparation of competent *M. smegmatis* cells

A single colony of *M. smegmatis* was inoculated into 5 ml TSB medium supplemented with 0.05 % Tween-80 and incubated at 37 °C with shaking overnight. The pre-culture was then used to inoculate 50 ml TSB medium with Tween-80 and cultivated at 37 °C with shaking until an OD₆₀₀ reached 0.8. Cells were harvested by centrifugation at 4,500 g for 10 minutes at 4 °C, washed thrice in ice-cold 10 % glycerol and resuspended in 1/10

of the volume 10 % glycerol. Aliquots of 50 μ l were dispensed, shock-frozen and stored at -80 °C.

6.5.10 Transformation of *E. coli* competent cells

Competent *E. coli* cells (50 μ l) were thawed on ice and 50 -100 ng of plasmid DNA was added to the cells. The mixture was transferred into 1.5 ml eppendorf tube and incubated at 4 °C for 30 minutes. Transformation was induced by a heat shock at 42 °C for 30 seconds. LB medium (500 μ l) was added to the sample, cells were allowed to recover at 37 °C for 1 hour and 20-100 μ l aliquots were plated onto LB agar plates supplemented with appropriate antibiotics and plates incubated at 37 °C for 12 hours.

6.5.11 Transformation of *C. glutamicum* competent cells

Competent *C. glutamicum* cells (50 μ l) and 5 μ g of plasmid DNA were transferred to a sterile electroporation cuvette (2 mm) and 0.8 ml of ice-cold 10 % glycerol solution was carefully laid over the cells. Electroporation was performed using Electroporator 2510 at 25 μ F, 2500 V. BHI media was added to the sample (4 ml) followed by a heat shock at 46 °C for 6 minutes. Cells were allowed to recover at 30 °C for 1 hour, harvested by centrifugation and resuspended in 400 μ l of BHI media. Aliquots (100 μ l) were plated onto BHI agar plates supplemented with appropriate antibiotics and plates incubated at 30 °C for 48 hours.

6.5.12 Electroporation of *M. smegmatis* competent cells

Competent *M. smegmatis* cells (100 μ l) were thawed on ice and 5 μ g of plasmid DNA was added to the cells. The mixture was transferred into an electroporation cuvette (1 mm) and electroporated using Electroporator 2510 at 25 μ F, 2500 V. Cells were

transferred into a culture tube, combined with 800 μ l TSB media and allowed to recover at 37 °C for 4 hours. Cell aliquots (50-100 μ l) were plated onto TSB agar plates supplemented with appropriate antibiotics and plates incubated at 37 °C for 48-60 hours.

6.5.13 Membrane fraction preparation

Cells were resuspended in buffer containing 50 mM MOPS pH 7.9, 5 mM β -mercaptoethanol and 10 mM $MgCl_2$ buffer. Cell slurry was sonicated (MSE Soniprep 150, 12 micron amplitude, 60 seconds ON, 90 seconds OFF, for 10 cycles), centrifuged at 27,000 g for 20 minutes at 4 °C, and the resulting supernatant re-centrifuged at 100,000 x g for 90 minutes at 4 °C to isolate cell membranes.

6.5.14 Extraction and analysis of cell wall associated lipids

Cells (10 ml) were grown to an early exponential growth phase (OD_{600} of 0.6), labelled with [1, 2- ^{14}C]-acetate 1 μ Ci/ml and incubated for further 6 hours at 30 °C with shaking. Labelled cells were harvested, washed twice in phosphate buffered saline and dried. The pellets were extracted using 2 ml of chloroform/methanol/water (10:10:3, v/v/v) for 4 hours at 50 °C. The organic extracts were combined with 1.75 ml of chloroform and 0.75 ml of water, mixed, and the two layers separated by centrifugation. The lower organic phase was recovered, washed twice with 2 ml of chloroform/methanol/water (3:47:48, v/v/v), dried and resuspended in 200 μ l of chloroform/methanol/water (10:10:3, v/v/v). An aliquot was subjected to TLC analysis, using silica gel plates (5554 silica gel 60F₂₅₄, Merck) developed in chloroform/methanol/water (60:16:2, v/v/v) and [^{14}C]-labelled lipids were visualised by autoradiography using Kodak BioMax MR films. Analysis of cold cell

wall associated lipids was performed in a similar fashion and visualised using molybdophosphoric acid (5 % in ethanol; w/v), followed by heating at 100 °C.

6.5.15 Extraction and analysis of cell wall bound lipids

Cells were grown, labelled with [1, 2-¹⁴C]-acetate and harvested as described above (Materials and Methods 6.5.14). The pellets were extracted using 2 ml of chloroform/methanol/water (10:10:3, v/v/v) for 4 hours at 50 °C, and bound lipids from the delipitated extracts were released by addition of 2 ml of 5 % (v/v) *tetra*-butyl ammonium hydroxide (TBAH), followed by a 12 hour incubation at 100 °C. After cooling, water (2 ml), dichloromethane (4 ml) and methyl iodide were added and mixed thoroughly for 30 minutes. The organic phase was recovered following centrifugation, washed three times using water (4 ml), dried and resuspended in diethyl-ether (4 ml). After centrifugation, the clear supernatant was dried and resuspended in dichloromethane (100 µl). An aliquot was subjected to the TLC (5554 silica gel 60F₂₅₄, Merck), developed in petroleum ether/acetone (95:5, v/v) and [¹⁴C]-labelled lipids visualised by autoradiography using Kodak BioMax MR films. Analysis of cold cell wall associated lipids was performed in a similar fashion and visualised using molybdophosphoric acid (5 % in ethanol; w/v), followed by heating at 100 °C.

6.5.16 Purification of mAGP complex

Cells (10 g) were resuspended in phosphate buffered saline containing 2 % Triton X-100 and lysed by sonication (MSE Soniprep 150, 12 micron amplitude, 60 seconds ON, 90 seconds OFF, for a total of 10 cycles). The cell slurry was centrifuged at 26,000 g for 30 minutes at 4 °C. The pelleted debris was extracted three times with 2 % SDS in

phosphate buffered saline at 95 °C and washed subsequently with excessive amount of water. The pellet was washed three times in 80% (v/v) acetone in water, and 100% acetone, and finally dried to yield a highly purified cell wall preparation.

6.5.17 Acid hydrolysis and alditol acetate derivatisation

The mAGP preparations (5-10 mg) were hydrolysed in 250 µl of 2 M trifluoroacetic acid at 120 °C for 2 hours. After cooling, samples were dried under compressed nitrogen and sugar residues reduced using 50 µl of sodium borohydride solution (10 mg/ml re-suspended in ethanol : 1 M ammonium hydroxide, 1:1). Alditols were per-*O*-acetylated using 100 µl of acetic anhydride at 100 °C for 1 hour. Samples were cooled, 100 µl of toluene added to the sample and dried. The resultant alditol acetates were particulated between water (2 ml) and chloroform (2 ml). The organic phase containing the alditol acetates were transferred to a new glass tube, dried under compressed nitrogen and further analysed by gas chromatography.

6.5.18 Gas chromatography analysis

GC analysis was performed using a CE Instruments ThermoQuest Trace GC 2000. Samples were injected in the splitless mode. The column used was a DB225 (Supelco) ID 0.10 mm, length 10 m and df 0.05 µm. The oven was programmed to hold at an isothermal temperature of 275 °C for a run time of 15 minutes. Data was collected and analysed using Xcaliber (v.1.2) software.

6.5.19 Extraction and purification of lipoglycans

Cells (10 ml) were cultivated as described above (General Materials and Methods 6.3.1), labelled with [1, 2-¹⁴C]-glucose (1 µCi/ml) and incubated for further 12 hours at 37 °C

with shaking. The labelled cells were harvested and washed twice in phosphate buffered saline. Pellet was resuspended in 4 ml of 50 % ethanol in water (v/v) and the mixture was refluxed at 85 °C for 6 hours. The supernatant was recovered by centrifugation and transferred to a fresh tube. The extraction process was repeated five times, the supernatant fractions were pooled and dried. Crude lipoglycans were subjected to 90 % phenol treatment (w/v in phosphate buffered saline) at 65 °C for 1 hour. After cooling, the sample was centrifuged and the upper aqueous layer recovered and dialysed against water (MWCO 3,500 kDa). An equal amount of radioactivity was subjected to 15 % SDS-PAGE and [¹⁴C]-labelled purified lipoglycans were visualised by autoradiography using Kodak BioMax MR films.

6.5.20 Acid hydrolysis of LAM for sugar analysis

LAM samples from wild type *M. smegmatis* and *M. smegmatis*Δ*aftB*::pMV306-*aftB* were hydrolysed using 400 µl of 2 M trifluoroacetic acid at 120 °C for 1 hour. Mixtures were dried under a stream of nitrogen to remove acid, residue resuspended in water (2 ml) and washed in chloroform (2 ml). The upper aqueous phase was recovered by centrifugation and dried. Samples were loaded onto HPTLC-Cellulose TLC (Merck) and developed thrice in formic acid/water/tert-butanol/methyl ethyl ketone (3:3:8:6, v/v/v/v). The [¹⁴C]-containing sugars were visualised by autoradiography employing Kodak BioMax MR films.

6.6 Protein biochemistry techniques

6.6.1 SDS-PAGE analysis

Mini-PROTEAN TGX pre-cast gels (Bio-Rad) were used for protein electrophoresis. Protein samples were mixed 1:1 with the loading dye (0.125 M Tris-HCl pH 6.8, 20 % glycerol, 4 % SDS and 10 % β -mercaptoethanol) and boiled at 95 °C for 10 minutes and approximately 10-20 μ g of protein was loaded in each well of the pre-cast gel. Electrophoresis was run at 300 mV and 25 mA and 300 mV and 45 mA for one and two gels, respectively. Protein bands were visualised by staining the gel for 30 minutes using Instant Blue (Expedeon).

6.6.2 Western blot analysis

Gels were processed for Western blot analysis. Briefly, protein transfer to nitrocellulose membrane (Bio-Rad) was carried out using the Mini Trans-blot electrophoretic transfer cell (Bio-Rad). The nitrocellulose membranes were probed with monoclonal penta-His antibody (1:1000, QIAGEN) and subsequently incubated with secondary anti-mouse antibody (1:20000, GE Healthcare) for His₆-tag detection. Protein bands were visualised using SIGMA FAST™ BCIP/NBT (5-Bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium; 1 tablet per 10 ml of dH₂O) system. For detection of T25 or T18 domains, nitrocellulose membranes were probed with either Cya (3D1) monoclonal (SantaCruz), or Cya (b-300) polyclonal antibodies (SantaCruz) (1:1,300) and subsequently incubated with either goat anti-mouse or goat anti-rabbit antibodies (1:25,000), respectively, and developed with chemiluminescence reagents (SantaCruz).

6.6.3 Protein expression and purification

Recombinant cells harbouring the construct of interest were used to inoculate an overnight culture of 5 ml media supplemented with appropriate antibiotic(s). This pre-culture was then used to inoculate 1 l of media supplemented with appropriate antibiotics and incubated at 30 °C (*C. glutamicum*) or 37 °C (*E. coli*) with shaking until an OD₆₀₀ of 0.5-0.8. IPTG was then added to a final concentration of 0.1-1 mM and growth continued for additional 4 hours at 30 °C/37 °C or 12 hours at 16 °C with shaking. Cells were harvested by centrifugation, washed with phosphate buffered saline, and cell pellets frozen until further use. Cell pellets were thawed and resuspended in lysis buffer (50 mM KH₂PO₄ pH 7.9, 300 mM NaCl, 20 % glycerol and 20 mM imidazole) containing protease inhibitor pellet (Roche). Cells were sonicated at a pulse rate of 60 seconds ON, 90 seconds OFF for a total of 10 cycles (MSE Soniprep 150, 12 micron amplitude) and centrifuged at 27,000 g for 30 minutes at 4 °C. The supernatant was collected and passed over a 1 ml HiTrap Ni²⁺-NTA agarose column (GE Healthcare), which was previously equilibrated with lysis buffer. Protein elution occurred *via* a stepwise gradient of 50-2,000 mM imidazole in 50 mM KH₂PO₄ pH 7.9, 300 mM NaCl. Pure protein fractions were determined by SDS-PAGE analysis.

| Appendix

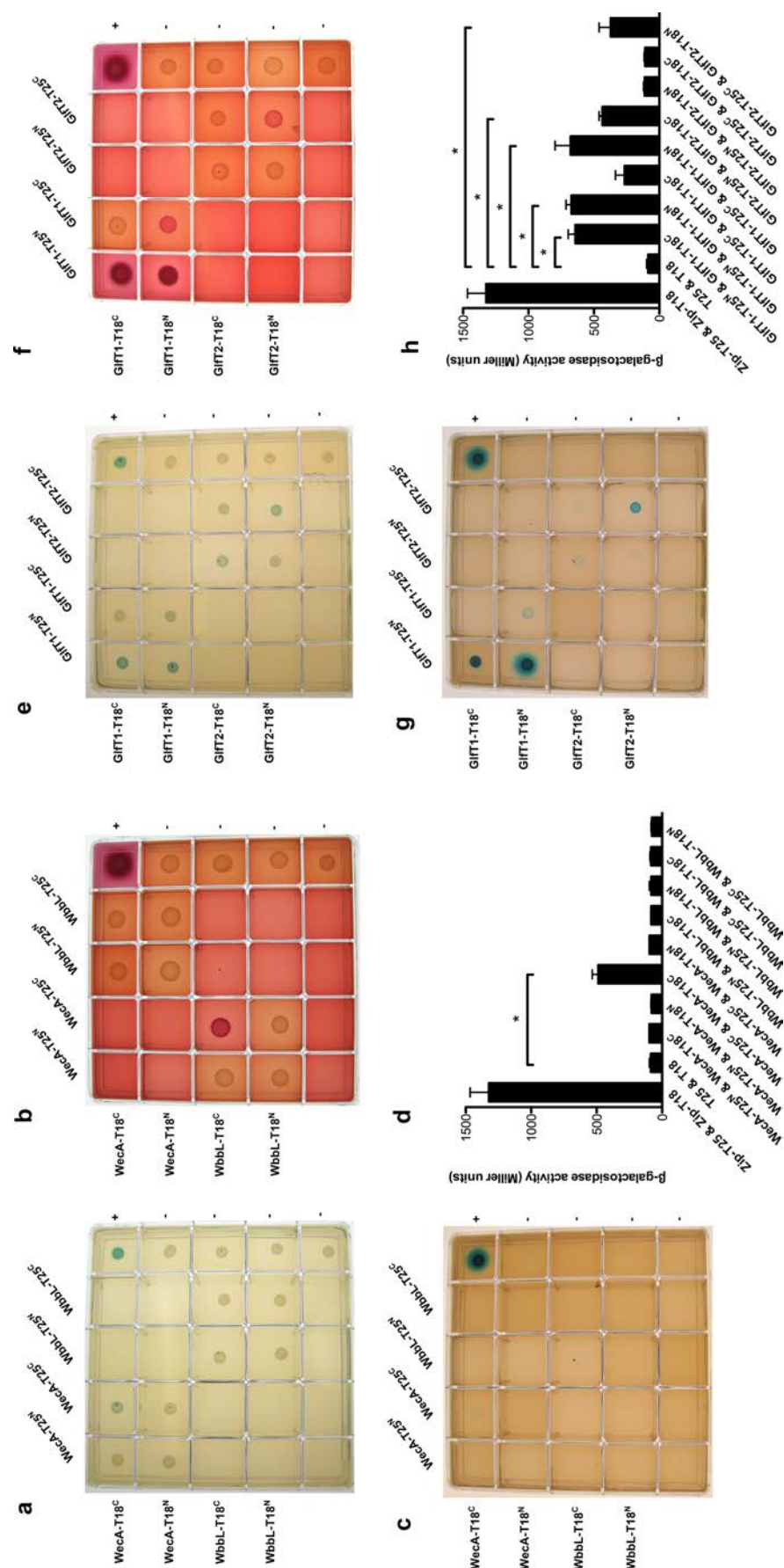


Figure 1 BACTH analysis of self-interactions of WecA, WbbL, GltT1, and GltT2 from *C. glutamicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya⁻* BTH101. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18, pKT25-pUT18c, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). d, h Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).

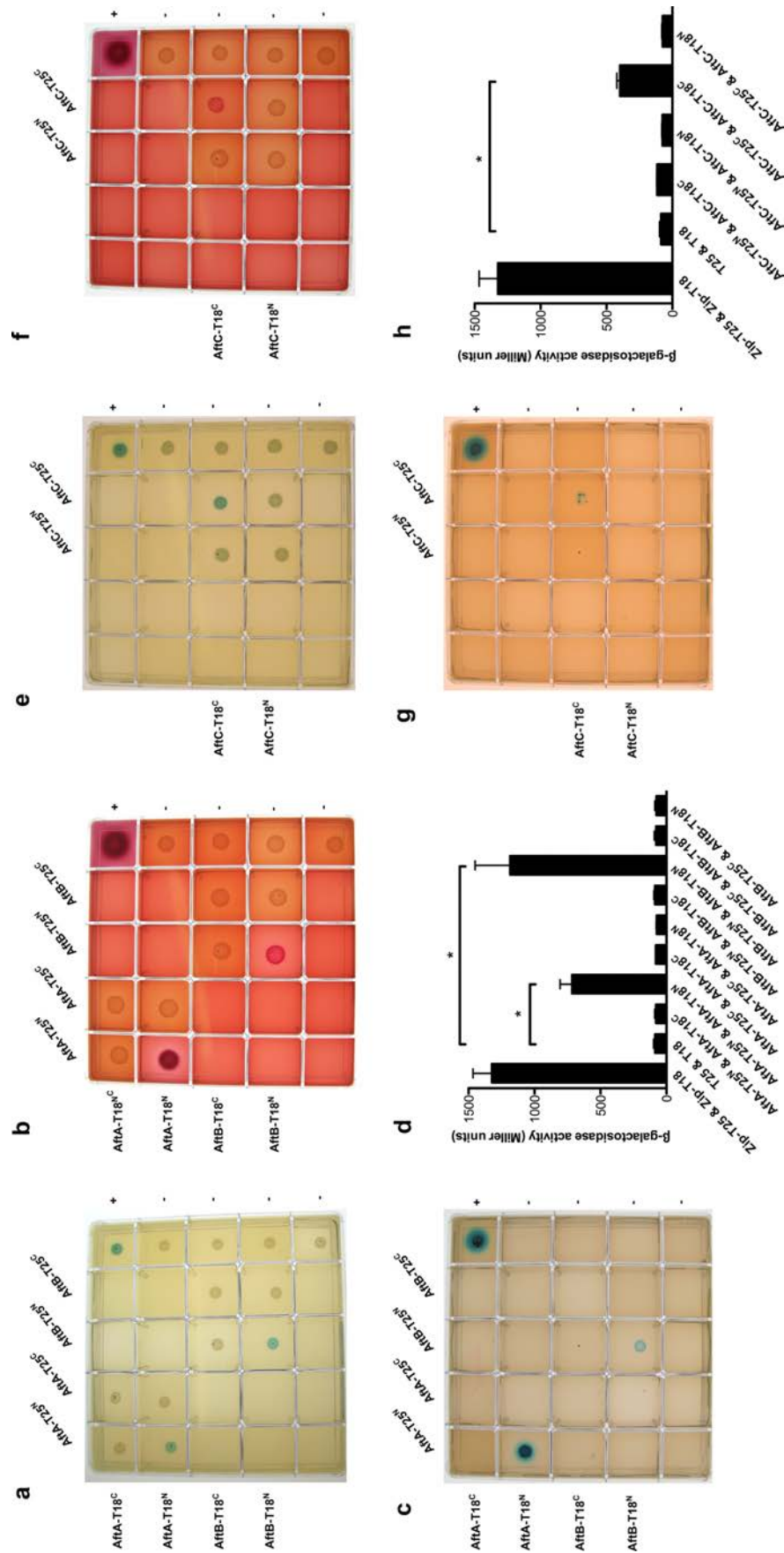


Figure 2 BACTH analysis of self-interactions of AftA, AftB and AftC from *C. glutamicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya* BTH101. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18c, pKT25-pUT18, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). d, h Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).

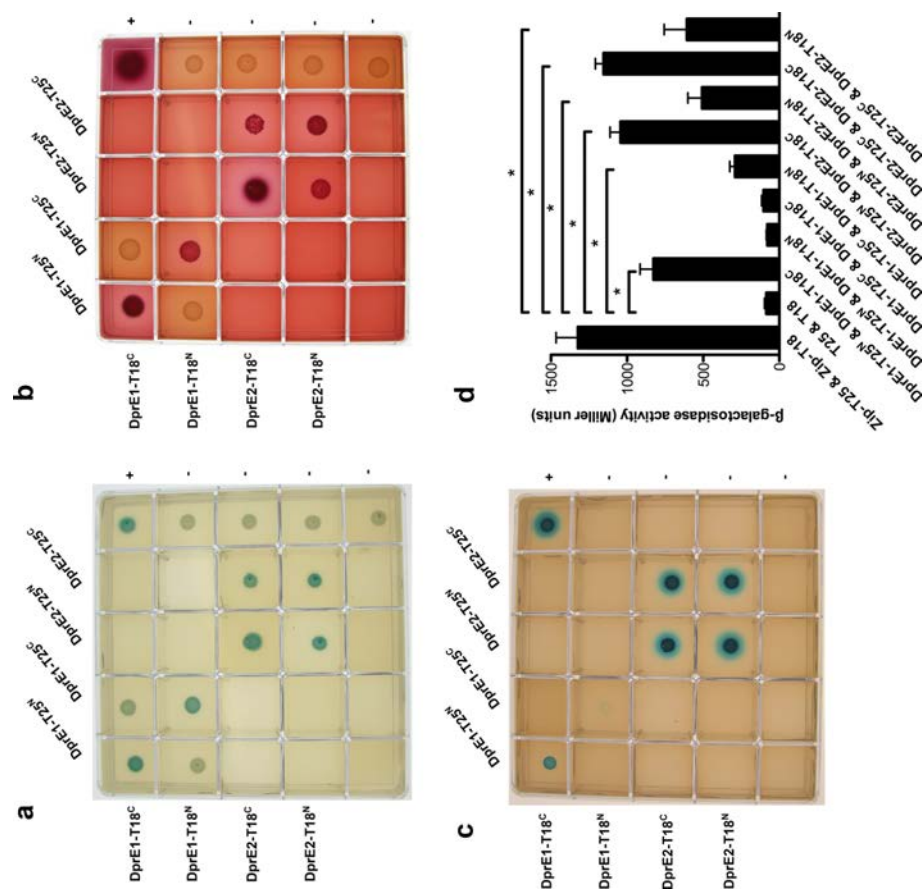
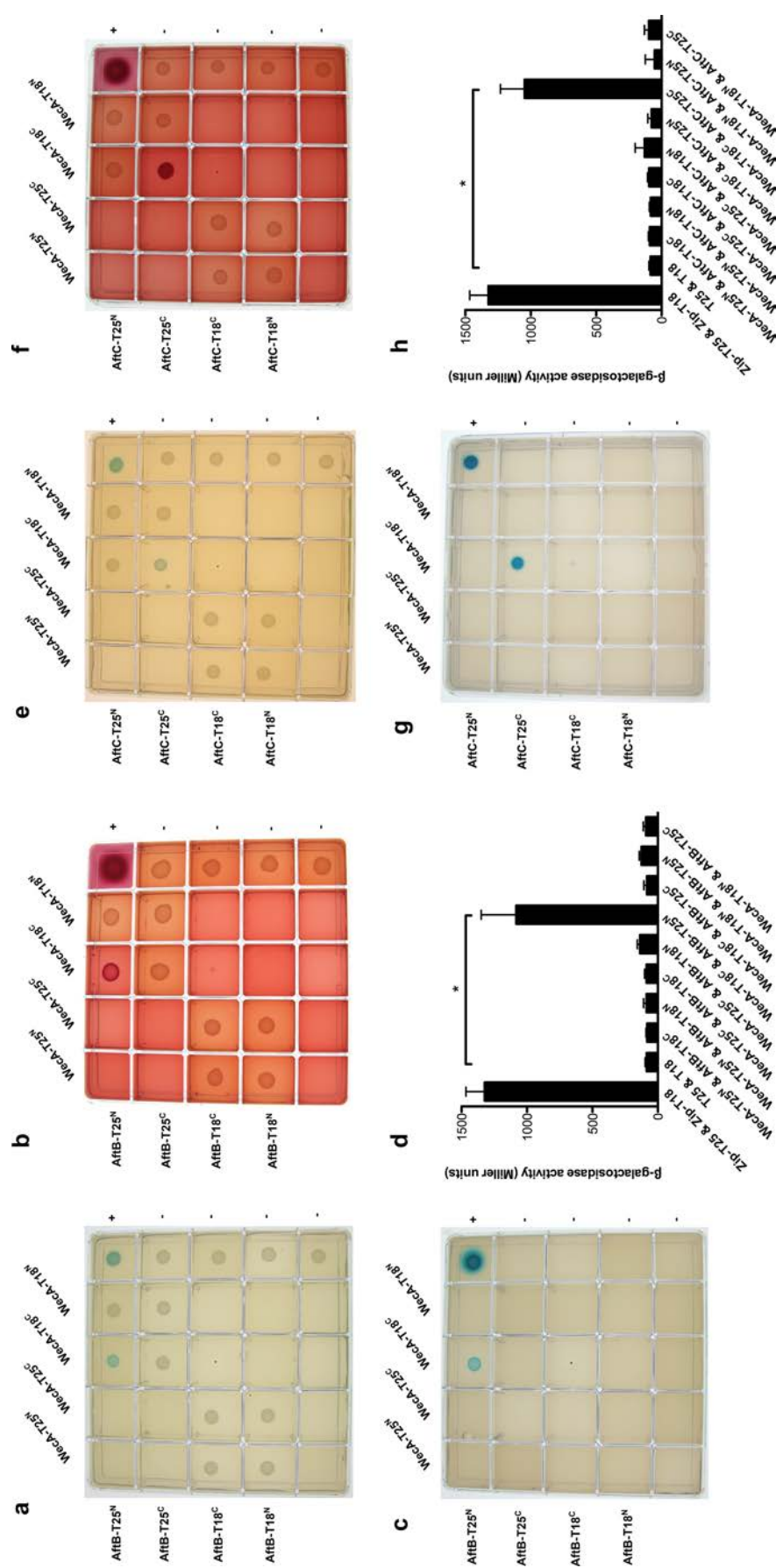


Figure 3 BACTH analysis of self-interactions of DprE1 and DprE2 from *C. glutanicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya*⁻ BTH101. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18, pKT25-pUT18c, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). d, h Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).



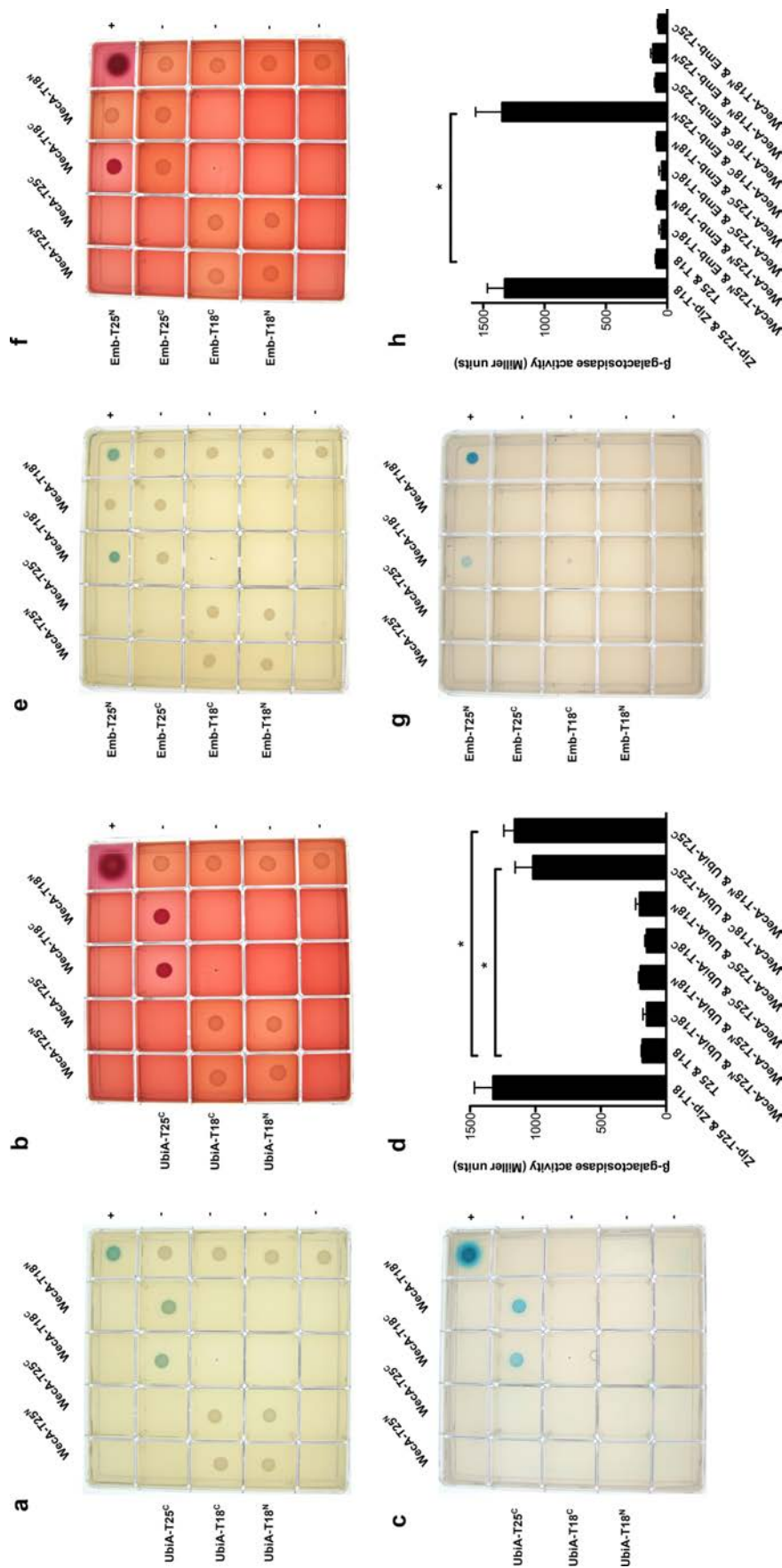


Figure 5 BACTH analysis of interactions between WecA-UbiA and WecA-Emb from *C. glutamicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya⁻ BTH101*. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18, pKT25-pUT18c, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). **d, h** Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).

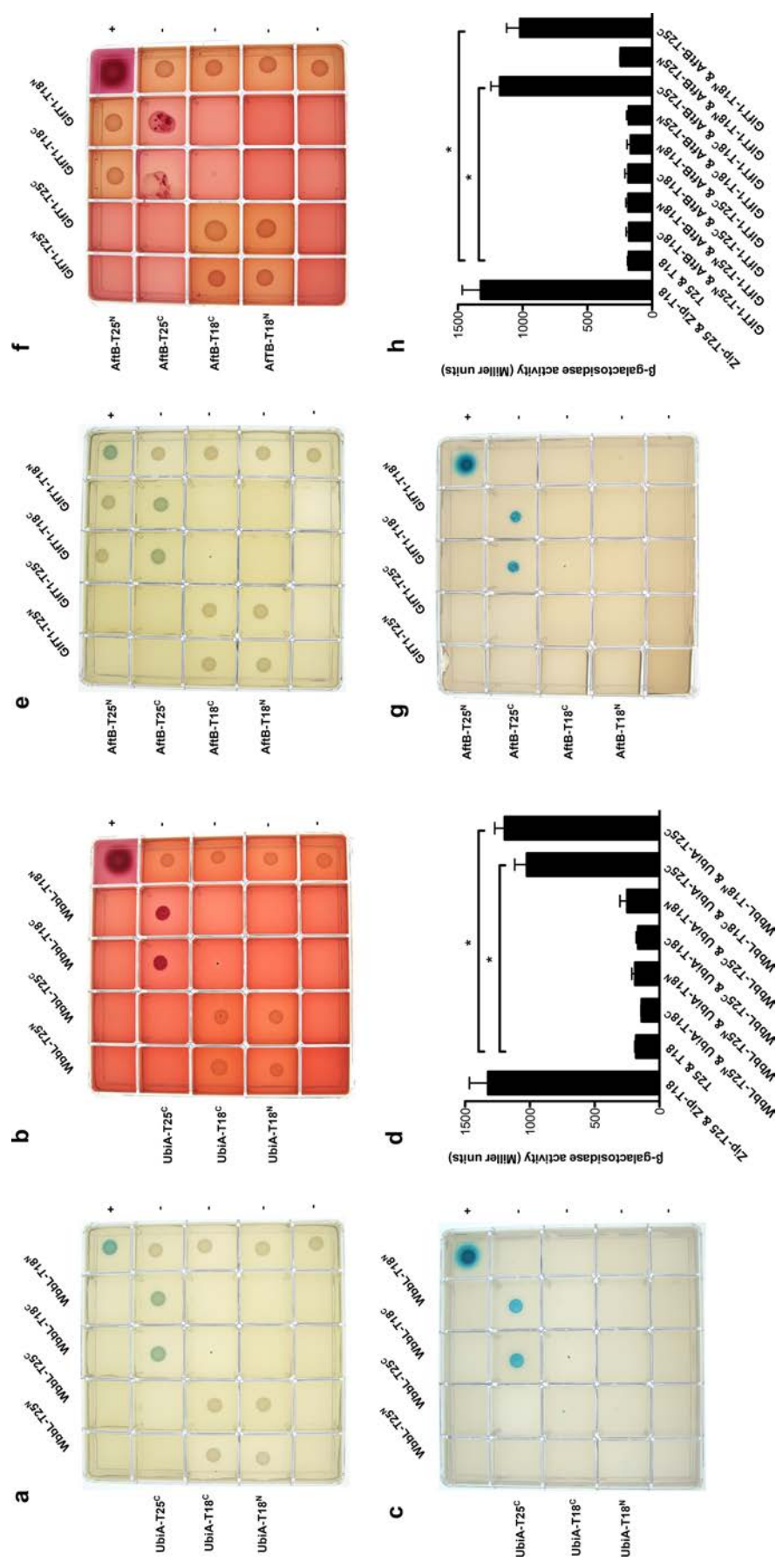


Figure 6 BACTH analysis of interactions between Wbbl-UbiA and Glt1-AftB from *C. glutamicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya* BTH101. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18, pKT25-pUT18c, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). **d, h** Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).

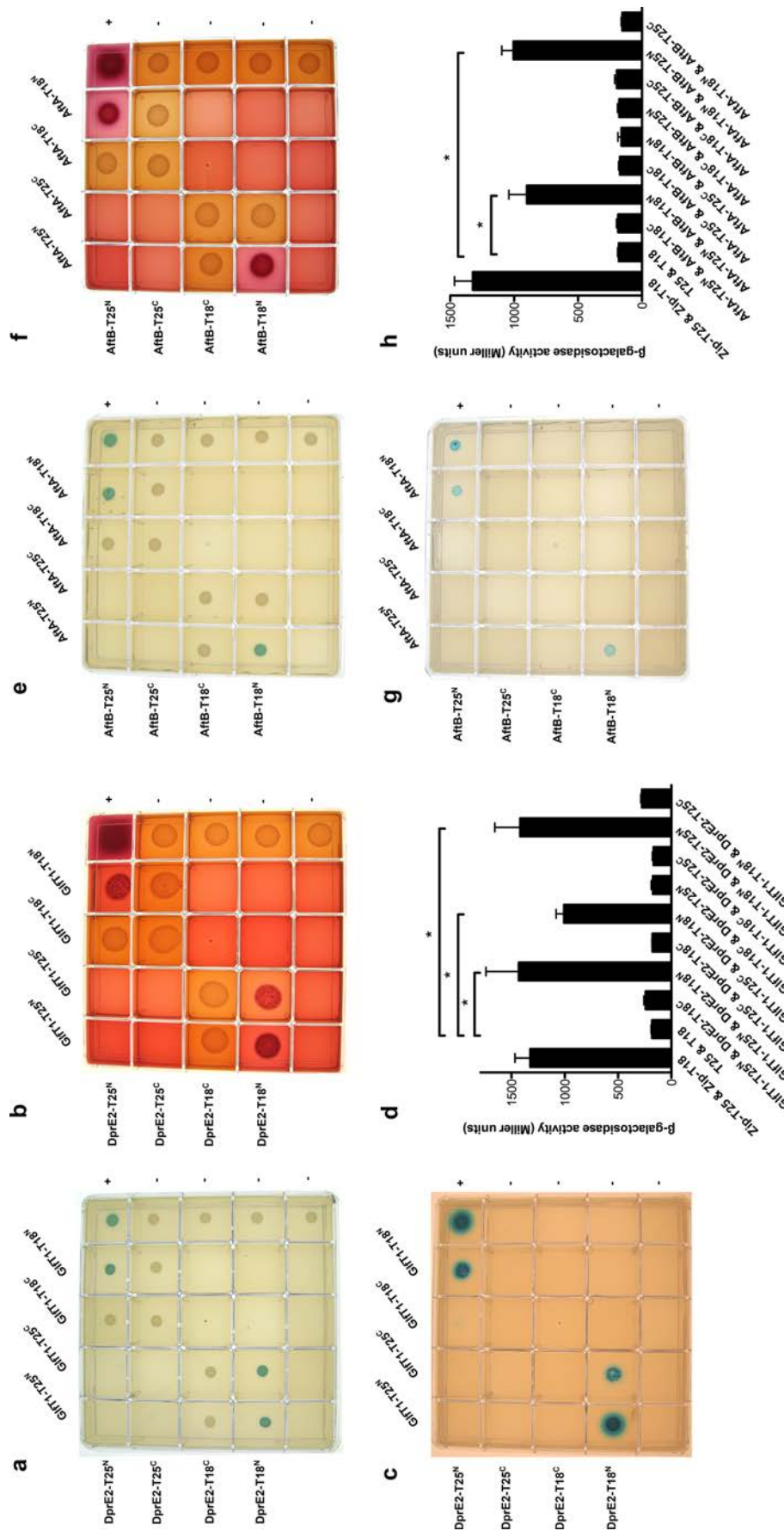


Figure 7 BACTH analysis of interactions between GIFT1-DprE2 and AftA-AftB from *C. glutamicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya⁻ BTH101*. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18, pKT25-pUT18c, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). **d, h** Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).

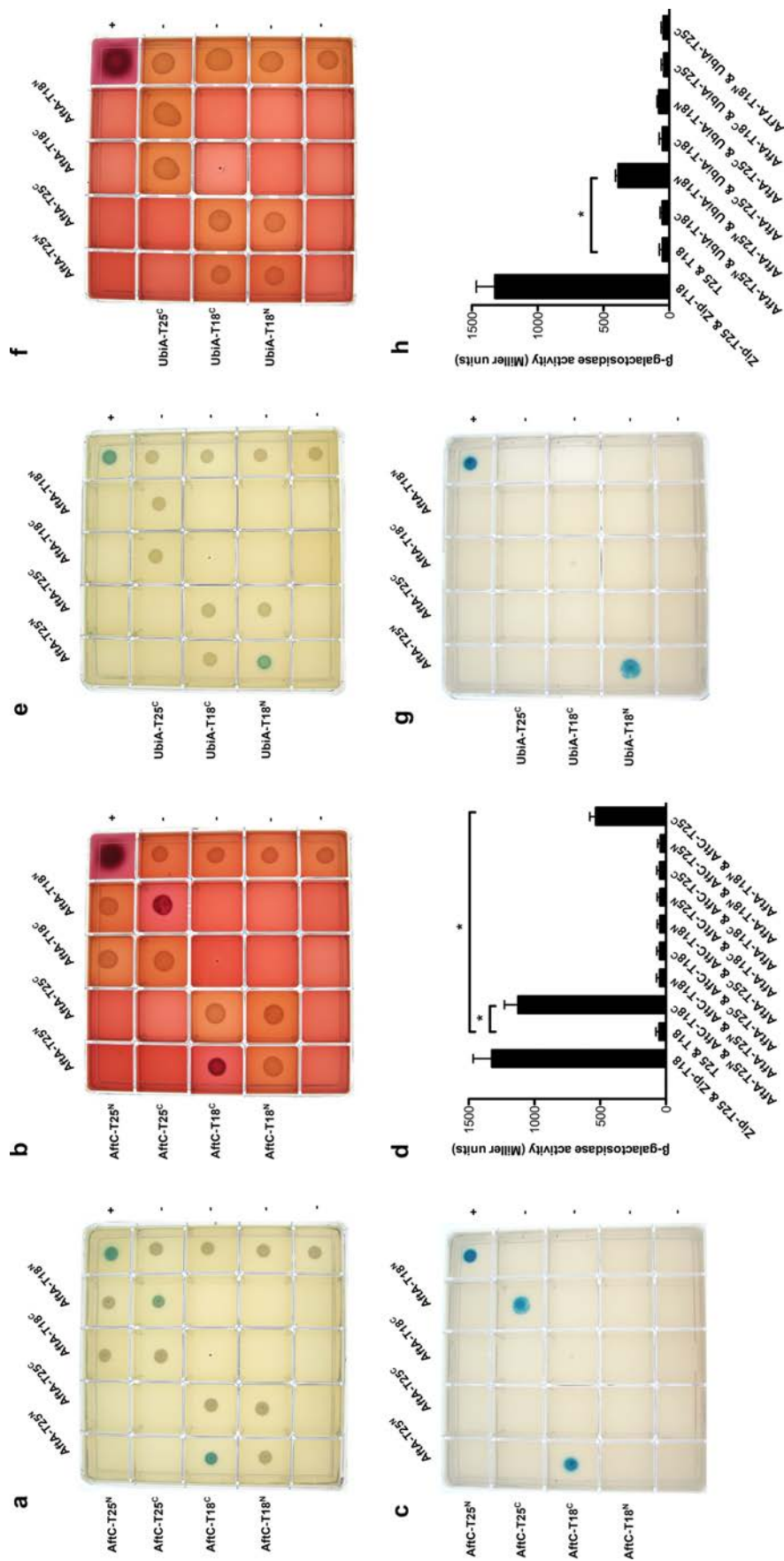


Figure 8 BACTH analysis of interactions between AftA-AftC and AftA-UbiA from *C. glutamicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya⁻ BTH101*. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18, pKT25-pUT18c, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). d, h Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).

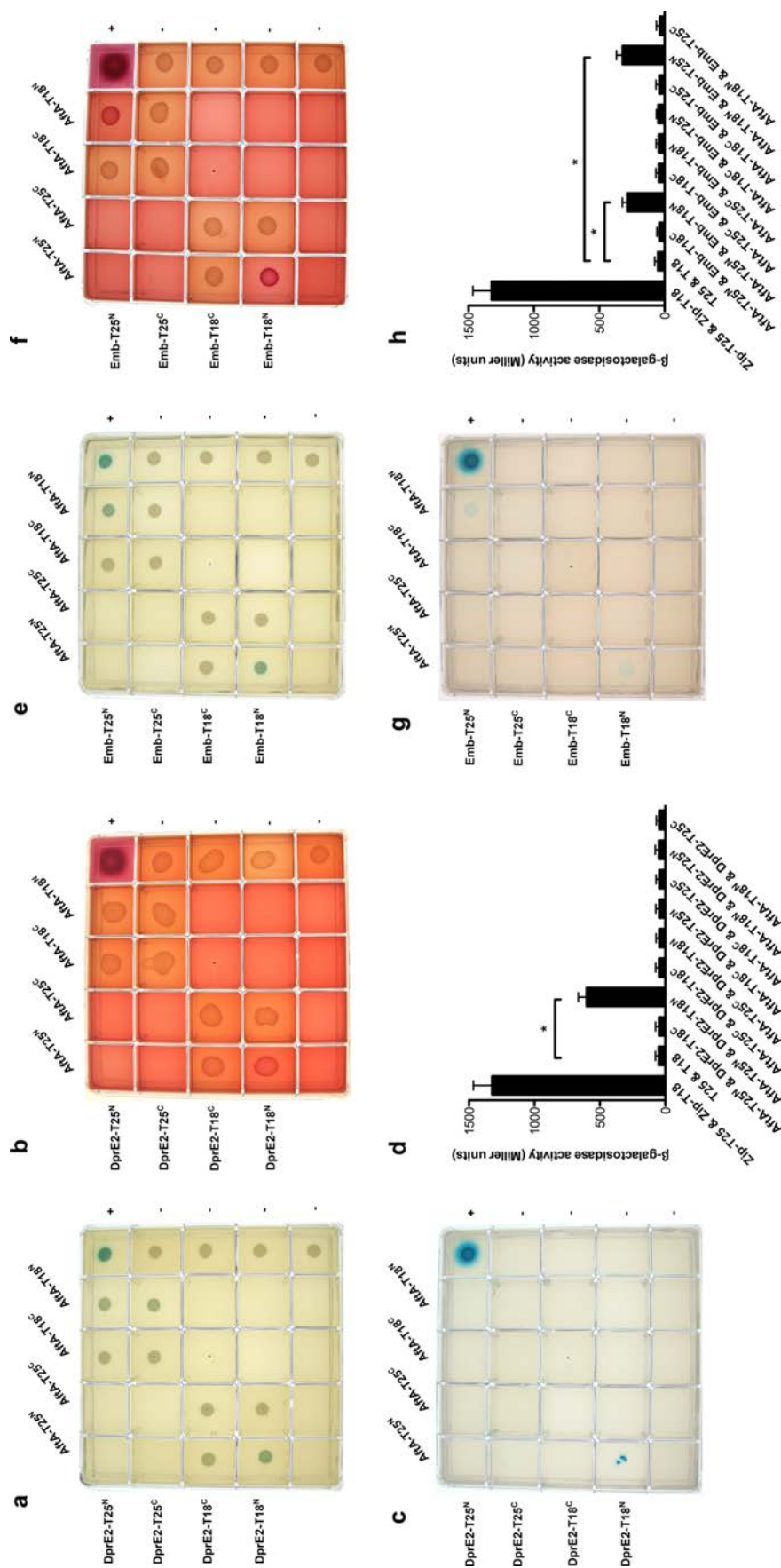


Figure 9 BACTH analysis of interactions between AftA-DprE2 and AftA-Emb from *C. glutamicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya⁻ BTH101*. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18, pKT25-pUT18c, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). d, h Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).

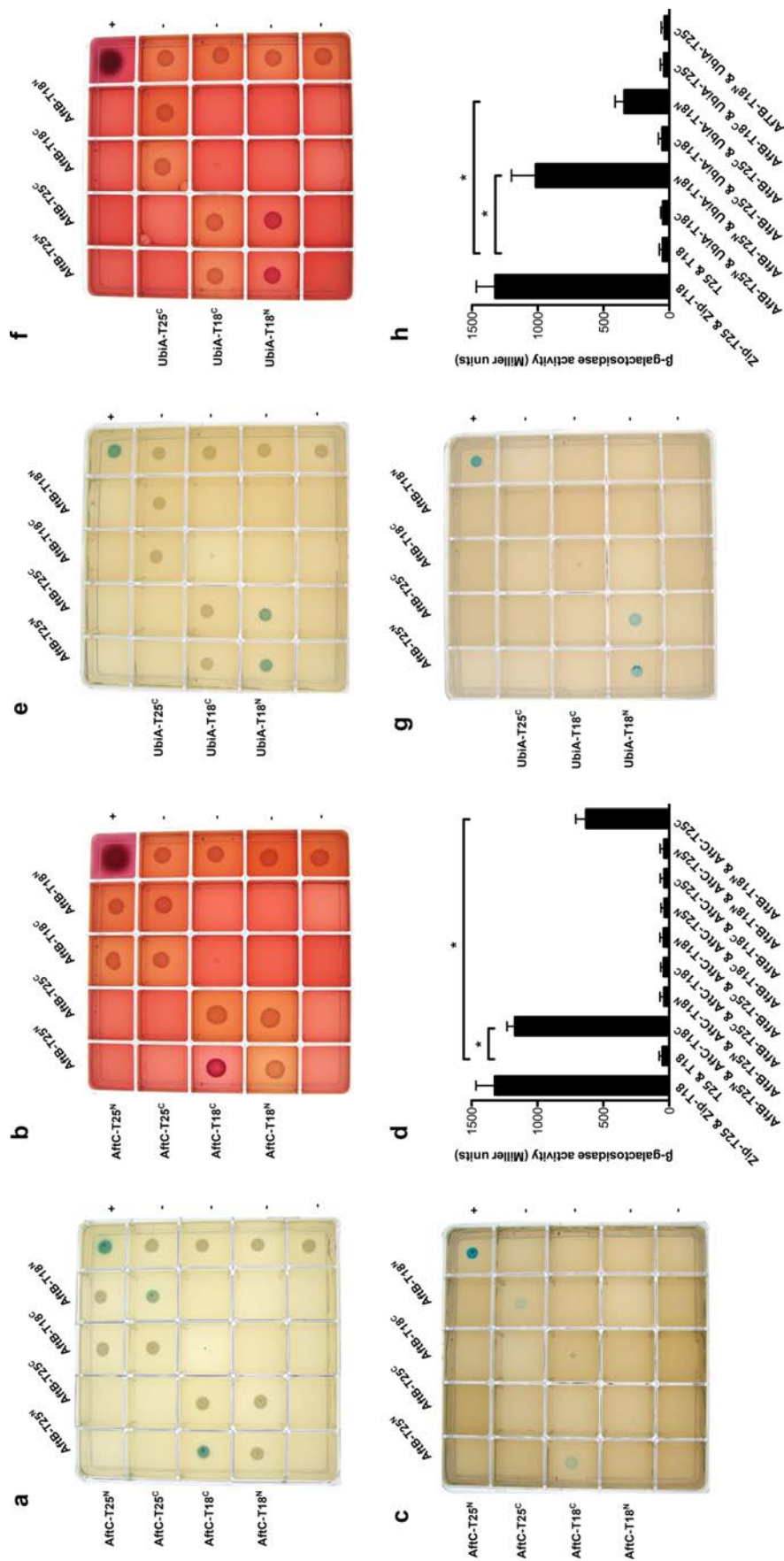


Figure 10 BACTH analysis of interactions between AftB-AftC and AftB-UbiA from *C. glutamicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya⁻ BTH101*. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18, pKT25-pUT18c, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). d, h Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).

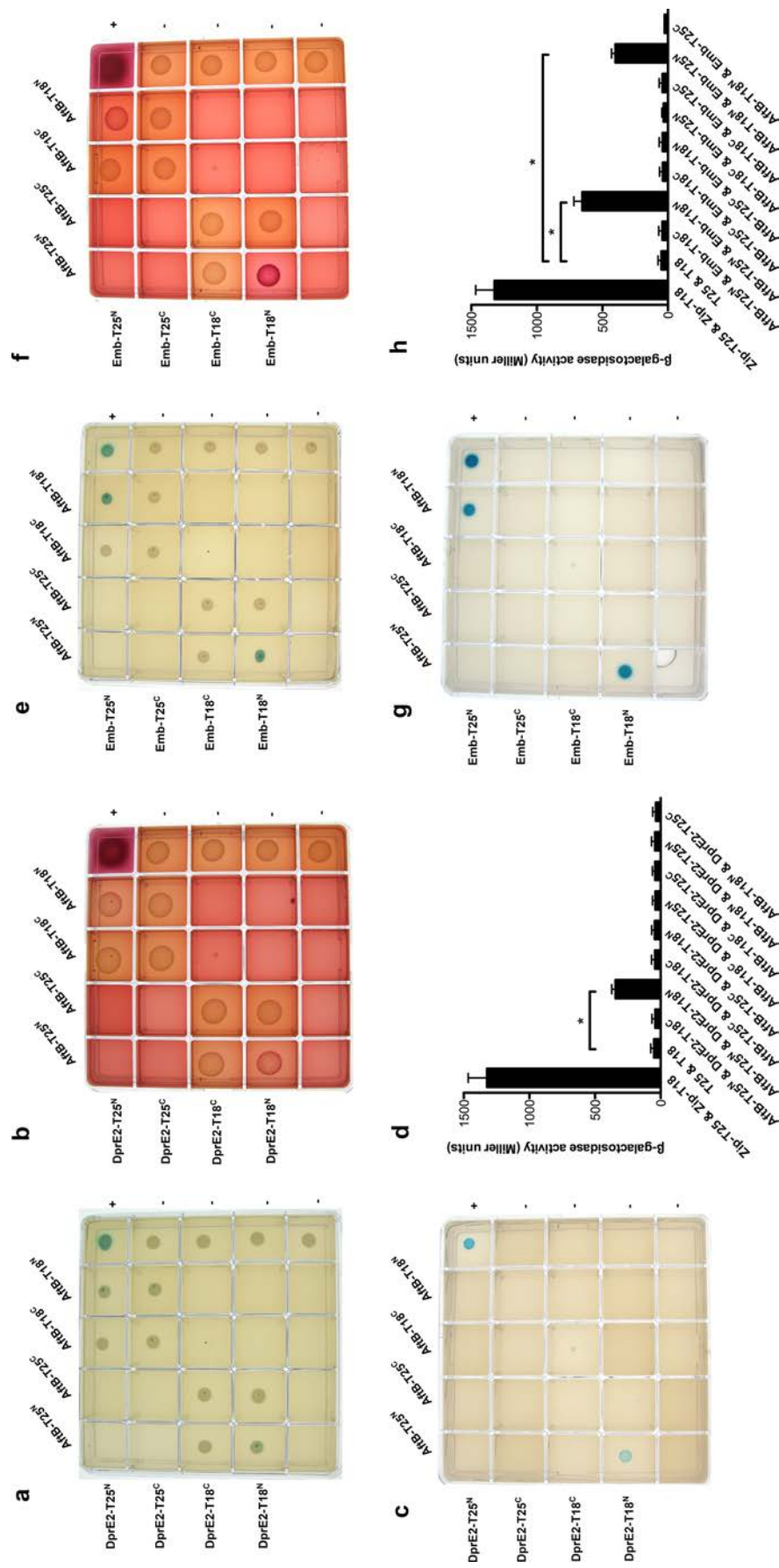


Figure 11 BACTH analysis of interactions between AftB-DprE2 and AftB-Emb from *C. glutamicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya⁻ BTH101*. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18, pKT25-pUT18c, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). **d, h** Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).

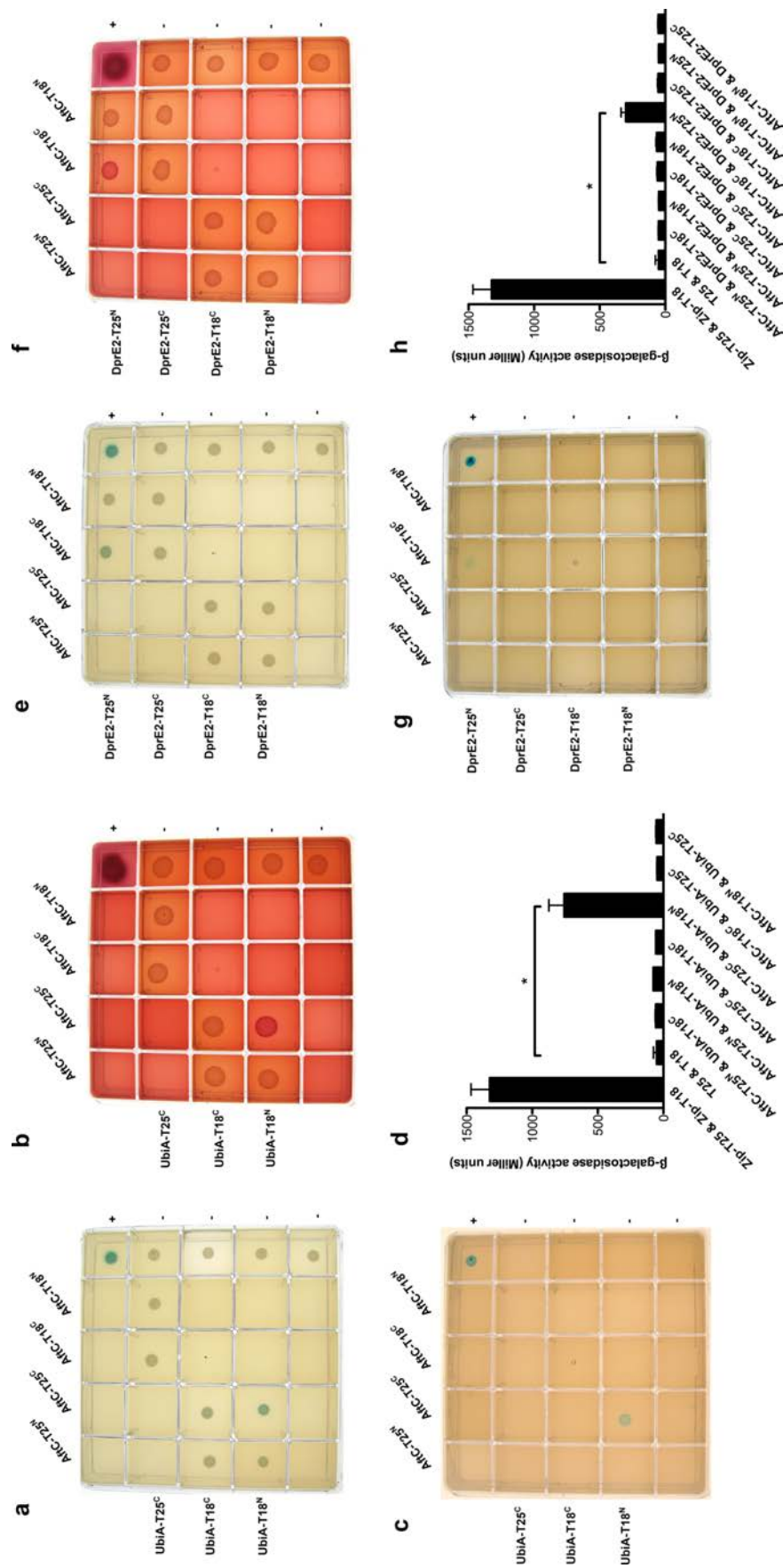


Figure 12 BACTH analysis of interactions between AftC-UbiA and AftC-DprE2 from *C. glutamicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya⁻ BTH101*. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18, pKT25-pUT18c, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). **d, h** Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).

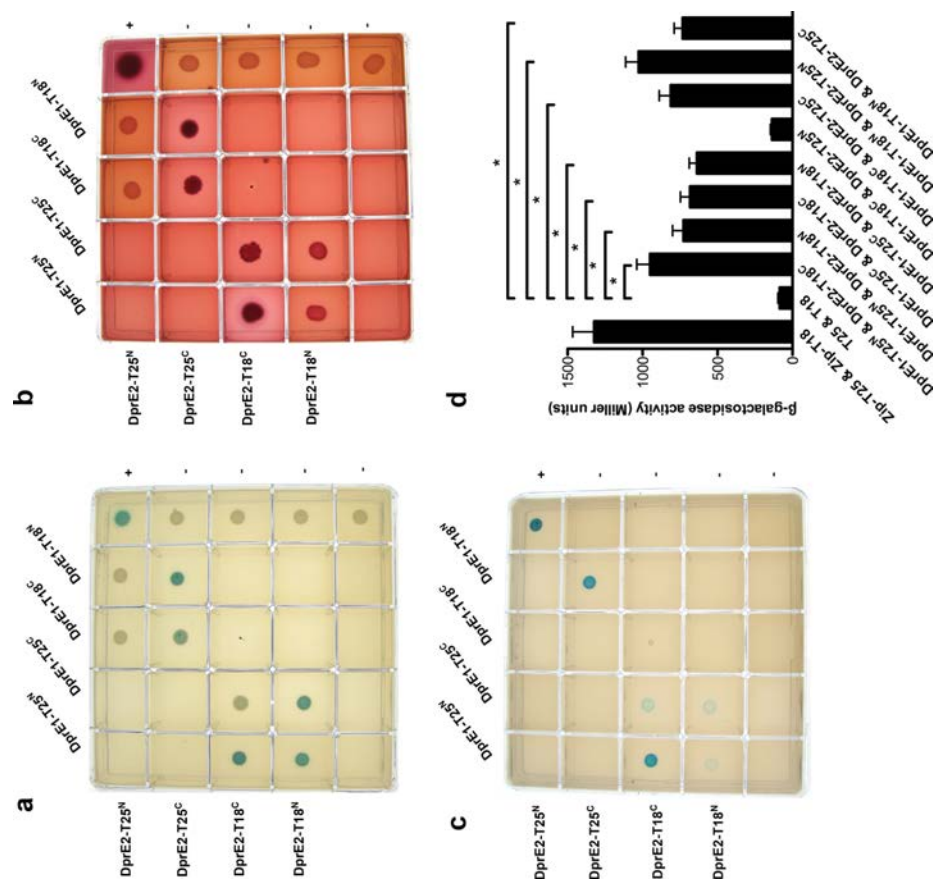


Figure 13 BACTH analysis of interactions between DprE1-DprE2 from *C. glutamicum*. The genes encoding full-length proteins were fused in frame with adenylate cyclase T25 or T18 fragments at N- or C-terminus and expressed in *E. coli cya* BTH101. Co-transformants containing two plasmids encoding putative interaction partners were spotted onto selective LB (a, e), MacConkey (b, f) and M63 (c, g) agar, as described in Materials and Methods. Protein-protein interactions are indicated by blue/red colonies through the reconstitution of adenylate cyclase catalytic domain. A strain co-expressing T25 and T18 fragments fused to leucine zipper domain was used as positive control (+), whereas empty pKT25-pUT18, pKT25-pUT18c, pKNT25-pUT18, and pKNT25-pUT18c were used as negative controls (-). **d, h** Relative β -galactosidase activities in *E. coli* BTH101 harboring the corresponding plasmids were measured to quantify the extent of protein-protein interactions. Results are expressed in Miller units and are the mean \pm standard deviation of three independent experiments. Statistical significance was determined by Student's t-test ($p < 0.01$).

| References

- ABE, S., TAKAYAMA, K., KINOSHITA, S. 1967. Taxonomical studies on glutamic acid producing bacteria. *J Gen Appl Microbiol*, 13, 279-301.
- ALCAIDE, F., PFYFFER, G. E. & TELENTI, A. 1997. Role of *embB* in natural and acquired resistance to ethambutol in mycobacteria. *Antimicrob Agents Chemother*, 41, 2270-3.
- ALDERWICK, L. J., DOVER, L. G., VEERAPEN, N., GURCHA, S. S., KREMER, L., ROPER, D. L., PATHAK, A. K., REYNOLDS, R. C. & BESRA, G. S. 2008. Expression, purification and characterisation of soluble GlfT and the identification of a novel galactofuranosyltransferase *Rv3782* involved in priming GlfT-mediated galactan polymerisation in *Mycobacterium tuberculosis*. *Protein Expr Purif*, 58, 332-41.
- ALDERWICK, L. J., LLOYD, G. S., GHADBANE, H., MAY, J. W., BHATT, A., EGGELING, L., FUTTERER, K. & BESRA, G. S. 2011. The C-terminal domain of the Arabinosyltransferase *Mycobacterium tuberculosis* EmbC is a lectin-like carbohydrate binding module. *PLoS Pathog*, 7, e1001299.
- ALDERWICK, L. J., LLOYD, G. S., LLOYD, A. J., LOVERING, A. L., EGGELING, L. & BESRA, G. S. 2011b. Biochemical characterisation of the *Mycobacterium tuberculosis* phosphoribosyl-1-pyrophosphate synthetase. *Glycobiology*, 21, 410-25.
- ALDERWICK, L. J., RADMACHER, E., SEIDEL, M., GANDE, R., HITCHEN, P. G., MORRIS, H. R., DELL, A., SAHM, H., EGGELING, L. & BESRA, G. S. 2005. Deletion of *Cg-emb* in *corynebacterianae* leads to a novel truncated cell wall arabinogalactan, whereas inactivation of *Cg-ubiA* results in an arabinan-deficient mutant with a cell wall galactan core. *J Biol Chem*, 280, 32362-71.
- ALDERWICK, L. J., SEIDEL, M., SAHM, H., BESRA, G. S. & EGGELING, L. 2006. Identification of a novel arabinofuranosyltransferase (AftA) involved in cell wall arabinan biosynthesis in *Mycobacterium tuberculosis*. *J Biol Chem*, 281, 15653-61.
- ALTAFA, M., MILLER, C. H., BELLOWS, D. S. & O'TOOLE, R. 2010. Evaluation of the *Mycobacterium smegmatis* and BCG models for the discovery of *Mycobacterium tuberculosis* inhibitors. *Tuberculosis (Edinb)*, 90, 333-7.
- AMER, A. O. & VALVANO, M. A. 2001. Conserved amino acid residues found in a predicted cytosolic domain of the lipopolysaccharide biosynthetic protein WecA are implicated in the recognition of UDP-N-acetylglucosamine. *Microbiology*, 147, 3015-25.
- AMIN, A. G., GOUDE, R., SHI, L., ZHANG, J., CHATTERJEE, D. & PARISH, T. 2008. EmbA is an essential arabinosyltransferase in *Mycobacterium tuberculosis*. *Microbiology*, 154, 240-8.
- APPELMELK, B. J., DEN DUNNEN, J., DRIESSEN, N. N., UMMELS, R., PAK, M., NIGOU, J., LARROUY-MAUMUS, G., GURCHA, S. S., MOVAHEDZADEH, F., GEURTSSEN, J., BROWN, E. J., EYSINK SMEETS, M. M., BESRA, G. S., WILLEMSSEN, P. T., LOWARY, T. L., VAN KOOYK, Y., MAASKANT, J. J., STOKER, N. G., VAN DER LEY, P., PUZO, G., VANDENBROUCKE-GRAULS, C.

- M., WIELAND, C. W., VAN DER POLL, T., GEIJTENBEEK, T. B., VAN DER SAR, A. M. & BITTER, W. 2008. The mannose cap of mycobacterial lipoarabinomannan does not dominate the *Mycobacterium*-host interaction. *Cell Microbiol*, 10, 930-44.
- APPELMELK, B. J., DEN DUNNEN, J., DRIESSEN, N. N., UMMELS, R., PAK, M., NIGOU, J., LARROUY-MAUMUS, G., GURCHA, S. S., MOVAHEDZADEH, F., GEURTSSEN, J., BROWN, E. J., EYSINK SMEETS, M. M., BESRA, G. S., WILLEMSSEN, P. T., LOWARY, T. L., VAN KOOYK, Y., MAASKANT, J. J., STOKER, N. G., VAN DER LEY, P., PUZO, G., VANDENBROUCKE-GRAULS, C. M., WIELAND, C. W., VAN DER POLL, T., GEIJTENBEEK, T. B., VAN DER SAR, A. M. & BITTER, W. 2008. The mannose cap of mycobacterial lipoarabinomannan does not dominate the *Mycobacterium*-host interaction. *Cell Microbiol*, 10, 930-44.
- ASSELINÉAU, J. & LEDERER, E. 1950. Structure of the mycolic acids of *Mycobacteria*. *Nature*, 166, 782-3.
- AZUMA, I., THOMAS, D. W., ADAM, A., GHUYSEN, J. M., BONALY, R., PETIT, J. F. & LEDERER, E. 1970. Occurrence of *N*-glycolylmuramic acid in bacterial cell walls. A preliminary survey. *Biochim Biophys Acta*, 208, 444-51.
- BABAOGU, K., PAGE, M. A., JONES, V. C., MCNEIL, M. R., DONG, C., NAISMITH, J. H. & LEE, R. E. 2003. Novel inhibitors of an emerging target in *Mycobacterium tuberculosis*; substituted thiazolidinones as inhibitors of dTDP-rhamnose synthesis. *Bioorg Med Chem Lett*, 13, 3227-30.
- BACHHAWAT, N. & MANDE, S. C. 1999. Identification of the INO1 gene of *Mycobacterium tuberculosis* H37Rv reveals a novel class of inositol-1-phosphate synthase enzyme. *J Mol Biol*, 291, 531-6.
- BALLOU, C. E., VILKAS, E. & LEDERER, E. 1963. Structural studies on the myo-inositol phospholipids of *Mycobacterium tuberculosis* (var. *bovis*, strain BCG). *J Biol Chem*, 238, 69-76.
- BARDAROV, S., BARDAROV JR, S., JR., PAVELKA JR, M. S., JR., SAMBANDAMURTHY, V., LARSEN, M., TUFARIELLO, J., CHAN, J., HATFULL, G. & JACOBS JR, W. R., JR. 2002. Specialized transduction: an efficient method for generating marked and unmarked targeted gene disruptions in *Mycobacterium tuberculosis*, *M. bovis* BCG and *M. smegmatis*. *Microbiology*, 148, 3007-17.
- BATT, S. M., JABEEN, T., BHOWRUTH, V., QUILL, L., LUND, P. A., EGGELING, L., ALDERWICK, L. J., FUTTERER, K. & BESRA, G. S. 2012. Structural basis of inhibition of *Mycobacterium tuberculosis* DprE1 by benzothiazinone inhibitors. *Proc Natl Acad Sci U S A*, 109, 11354-9.
- BATT, S. M., JABEEN, T., MISHRA, A. K., VEERAPEN, N., KRUMBACH, K., EGGELING, L., BESRA, G. S. & FUTTERER, K. 2010. Acceptor substrate discrimination in phosphatidyl-myo-inositol mannoside synthesis: structural and mutational analysis of mannosyltransferase *Corynebacterium glutamicum* PimB'. *J Biol Chem*, 285, 37741-52.

- BATTESTI, A. & BOUVERET, E. 2012. The bacterial two-hybrid system based on adenylate cyclase reconstitution in *Escherichia coli*. *Methods*, 58, 325-34.
- BAULARD, A. R., GURCHA, S. S., ENGOHANG-NDONG, J., GOUFFI, K., LOCHT, C. & BESRA, G. S. 2003. *In vivo* interaction between the polyprenol phosphate mannose synthase Ppm1 and the integral membrane protein Ppm2 from *Mycobacterium smegmatis* revealed by a bacterial two-hybrid system. *J Biol Chem*, 278, 2242-8.
- BEIS, K., SRIKANNATHASAN, V., LIU, H., FULLERTON, S. W., BAMFORD, V. A., SANDERS, D. A., WHITFIELD, C., MCNEIL, M. R. & NAISMITH, J. H. 2005. Crystal structures of *Mycobacteria tuberculosis* and *Klebsiella pneumoniae* UDP-galactopyranose mutase in the oxidised state and *Klebsiella pneumoniae* UDP-galactopyranose mutase in the (active) reduced state. *J Mol Biol*, 348, 971-82.
- BELANGER, A. E., BESRA, G. S., FORD, M. E., MIKUSOVA, K., BELISLE, J. T., BRENNAN, P. J. & INAMINE, J. M. 1996. The *embAB* genes of *Mycobacterium avium* encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. *Proc Natl Acad Sci U S A*, 93, 11919-24.
- BELANOVA, M., DIANISKOVA, P., BRENNAN, P. J., COMPLETEO, G. C., ROSE, N. L., LOWARY, T. L. & MIKUSOVA, K. 2008. Galactosyl transferases in mycobacterial cell wall synthesis. *J Bacteriol*, 190, 1141-5.
- BELISLE, J. T., VISSA, V. D., SIEVERT, T., TAKAYAMA, K., BRENNAN, P. J. & BESRA, G. S. 1997. Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. *Science*, 276, 1420-2.
- BERG, S., KAUR, D., JACKSON, M. & BRENNAN, P. J. 2007. The glycosyltransferases of *Mycobacterium tuberculosis* - roles in the synthesis of arabinogalactan, lipoarabinomannan, and other glycoconjugates. *Glycobiology*, 17, 35-56R.
- BERG, S., STARBUCK, J., TORRELLES, J. B., VISSA, V. D., CRICK, D. C., CHATTERJEE, D. & BRENNAN, P. J. 2005. Roles of conserved proline and glycosyltransferase motifs of EmbC in biosynthesis of lipoarabinomannan. *J Biol Chem*, 280, 5651-63.
- BESRA, G. S. & BRENNAN, P. J. 1997. The mycobacterial cell wall: biosynthesis of arabinogalactan and lipoarabinomannan. *Biochem Soc Trans*, 25, 845-50.
- BESRA, G. S., KHOO, K. H., MCNEIL, M. R., DELL, A., MORRIS, H. R. & BRENNAN, P. J. 1995. A new interpretation of the structure of the mycolyl-arabinogalactan complex of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosylalditol fragments by fast-atom bombardment mass spectrometry and ¹H nuclear magnetic resonance spectroscopy. *Biochemistry*, 34, 4257-66.

- BHAMIDI, S., SCHERMAN, M. S., RITHNER, C. D., PRENNI, J. E., CHATTERJEE, D., KHOO, K. H. & MCNEIL, M. R. 2008. The identification and location of succinyl residues and the characterization of the interior arabinan region allow for a model of the complete primary structure of *Mycobacterium tuberculosis* mycolyl arabinogalactan. *J Biol Chem*, 283, 12992-3000.
- BHATT, A. & JACOBS, W. R., JR. 2009. Gene essentiality testing in *Mycobacterium smegmatis* using specialized transduction. *Methods Mol Biol*, 465, 325-36.
- BHATT, A., BROWN, A. K., SINGH, A., MINNIKIN, D. E. & BESRA, G. S. 2008. Loss of a mycobacterial gene encoding a reductase leads to an altered cell wall containing beta-oxo-mycolic acid analogs and accumulation of ketones. *Chem Biol*, 15, 930-9.
- BHATT, A., KREMER, L., DAI, A. Z., SACCHETTINI, J. C. & JACOBS, W. R., JR. 2005. Conditional depletion of KasA, a key enzyme of mycolic acid biosynthesis, leads to mycobacterial cell lysis. *J Bacteriol*, 187, 7596-606.
- BHATT, A., MOLLE, V., BESRA, G. S., JACOBS, W. R., JR. & KREMER, L. 2007. The *Mycobacterium tuberculosis* FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development. *Mol Microbiol*, 64, 1442-54.
- BIRCH, H. L., ALDERWICK, L. J., APPELMELK, B. J., MAASKANT, J., BHATT, A., SINGH, A., NIGOU, J., EGGELING, L., GEURTSSEN, J. & BESRA, G. S. 2010. A truncated lipoglycan from mycobacteria with altered immunological properties. *Proc Natl Acad Sci U S A*, 107, 2634-9.
- BIRCH, H. L., ALDERWICK, L. J., BHATT, A., RITTMANN, D., KRUMBACH, K., SINGH, A., BAI, Y., LOWARY, T. L., EGGELING, L. & BESRA, G. S. 2008. Biosynthesis of mycobacterial arabinogalactan: identification of a novel $\alpha(1\rightarrow3)$ arabinofuranosyltransferase. *Mol Microbiol*, 69, 1191-206.
- BOUHSS, A., CROUVOISIER, M., BLANOT, D. & MENGIN-LECREULX, D. 2004. Purification and characterization of the bacterial MraY translocase catalyzing the first membrane step of peptidoglycan biosynthesis. *J Biol Chem*, 279, 29974-80.
- BOUHSS, A., MENGIN-LECREULX, D., LE BELLER, D. & VAN HEIJENOORT, J. 1999. Topological analysis of the MraY protein catalysing the first membrane step of peptidoglycan synthesis. *Mol Microbiol*, 34, 576-85.
- BRENNAN, P. J. & NIKAIDO, H. 1995. The envelope of mycobacteria. *Annu Rev Biochem*, 64, 29-63.
- BROSCH, R., GORDON, S. V., MARMIESSE, M., BRODIN, P., BUCHRIESER, C., EIGLMEIER, K., GARNIER, T., GUTIERREZ, C., HEWINSON, G., KREMER, K., PARSONS, L. M., PYM, A. S., SAMPER, S., VAN SOOLINGEN, D. & COLE, S. T. 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci U S A*, 99, 3684-9.

- BROWN-ELLIOTT, B. A., GRIFFITH, D. E. & WALLACE, R. J., JR. 2002. Newly described or emerging human species of nontuberculous mycobacteria. *Infect Dis Clin North Am*, 16, 187-220.
- BROWN, A. K., MENG, G., GHADBANE, H., SCOTT, D. J., DOVER, L. G., NIGOU, J., BESRA, G. S. & FUTTERER, K. 2007. Dimerization of inositol monophosphatase *Mycobacterium tuberculosis* SuhB is not constitutive, but induced by binding of the activator Mg²⁺. *BMC Struct Biol*, 7, 55.
- BROWN, E. D., VIVAS, E. I., WALSH, C. T. & KOLTER, R. 1995. MurA (MurZ), the enzyme that catalyzes the first committed step in peptidoglycan biosynthesis, is essential in *Escherichia coli*. *J Bacteriol*, 177, 4194-7.
- BUDELMEIJER, N. & BECKWITH, J. 2004. A complex of the *Escherichia coli* cell division proteins FtsL, FtsB and FtsQ forms independently of its localization to the septal region. *Mol Microbiol*, 52, 1315-27.
- BURGESS, R. R. 2009. Refolding solubilized inclusion body proteins. *Methods Enzymol*, 463, 259-82.
- CABRITA, L. D. & BOTTOMLEY, S. P. 2004. Protein expression and refolding-a practical guide to getting the most out of inclusion bodies. *Biotechnol Annu Rev*, 10, 31-50.
- CALMETTE, A. & GUERIN, C. 1924. Vaccination des bovidés contre la tuberculose et méthode.
- CALMETTE, A. 1928. On preventive vaccination of the new-born against tuberculosis by B.C.G. *British Journal of Tuberculosis*, 22.
- CHAN, K., KNAAK, T., SATKAMP, L., HUMBERT, O., FALKOW, S. & RAMAKRISHNAN, L. 2002. Complex pattern of *Mycobacterium marinum* gene expression during long-term granulomatous infection. *Proc Natl Acad Sci U S A*, 99, 3920-5.
- CHANG, J. C., MINER, M. D., PANDEY, A. K., GILL, W. P., HARIK, N. S., SASSETTI, C. M. & SHERMAN, D. R. 2009. *igr* Genes and *Mycobacterium tuberculosis* cholesterol metabolism. *J Bacteriol*, 191, 5232-9.
- CHATTERJEE, D., BOZIC, C. M., MCNEIL, M. & BRENNAN, P. J. 1991. Structural features of the arabinan component of the lipoarabinomannan of *Mycobacterium tuberculosis*. *J Biol Chem*, 266, 9652-60.
- CHATTERJEE, D., HUNTER, S. W., MCNEIL, M. & BRENNAN, P. J. 1992a. Lipoarabinomannan. Multiglycosylated form of the mycobacterial mannosylphosphatidylinositols. *J Biol Chem*, 267, 6228-33.
- CHATTERJEE, D., KHOO, K. H., MCNEIL, M. R., DELL, A., MORRIS, H. R. & BRENNAN, P. J. 1993. Structural definition of the non-reducing termini of mannose-

capped LAM from *Mycobacterium tuberculosis* through selective enzymatic degradation and fast atom bombardment-mass spectrometry. *Glycobiology*, 3, 497-506.

CHATTERJEE, D., LOWELL, K., RIVOIRE, B., MCNEIL, M. R. & BRENNAN, P. J. 1992b. Lipoarabinomannan of *Mycobacterium tuberculosis*. Capping with mannosyl residues in some strains. *J Biol Chem*, 267, 6234-9.

CHRISTOPHE, T., JACKSON, M., JEON, H. K., FENISTEIN, D., CONTRERAS-DOMINGUEZ, M., KIM, J., GENOVESIO, A., CARRALOT, J. P., EWANN, F., KIM, E. H., LEE, S. Y., KANG, S., SEO, M. J., PARK, E. J., SKOVIEROVA, H., PHAM, H., RICCARDI, G., NAM, J. Y., MARSOLLIER, L., KEMPF, M., JOLY-GUILLOU, M. L., OH, T., SHIN, W. K., NO, Z., NEHRBASS, U., BROSCHE, R., COLE, S. T. & BRODIN, P. 2009. High content screening identifies decaprenyl-phosphoribose 2' epimerase as a target for intracellular antimycobacterial inhibitors. *PLoS Pathog*, 5, e1000645.

CLARK-CURTISS, J. E. 1990. Genome structure of mycobacteria. *Molecular Biology of the Mycobacteria*, Academic Press Ltd., 77-96.

CLARKE, B. R., GREENFIELD, L. K., BOUWMAN, C. & WHITFIELD, C. 2009. Coordination of polymerization, chain termination, and export in assembly of the *Escherichia coli* lipopolysaccharide O9a antigen in an ATP-binding cassette transporter-dependent pathway. *J Biol Chem*, 284, 30662-72.

COLE, S. T., BROSCHE, R., PARKHILL, J., GARNIER, T., CHURCHER, C., HARRIS, D., GORDON, S. V., EIGLMEIER, K., GAS, S., BARRY, C. E., 3RD, TEKAIA, F., BADCOCK, K., BASHAM, D., BROWN, D., CHILLINGWORTH, T., CONNOR, R., DAVIES, R., DEVLIN, K., FELTWELL, T., GENTLES, S., HAMLIN, N., HOLROYD, S., HORNSBY, T., JAGELS, K., KROGH, A., MCLEAN, J., MOULE, S., MURPHY, L., OLIVER, K., OSBORNE, J., QUAIL, M. A., RAJANDREAM, M. A., ROGERS, J., RUTTER, S., SEEGER, K., SKELTON, J., SQUARES, R., SQUARES, S., SULSTON, J. E., TAYLOR, K., WHITEHEAD, S. & BARRELL, B. G. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*, 393, 537-44.

COLLINS, M. D., GOODFELLOW, M. & MINNIKIN, D. E. 1982a. Fatty acid composition of some mycolic acid-containing coryneform bacteria. *J Gen Microbiol*, 128, 2503-9.

COLLINS, M. D., GOODFELLOW, M. & MINNIKIN, D. E. 1982b. A survey of the structures of mycolic acids in *Corynebacterium* and related taxa. *J Gen Microbiol*, 128, 129-49.

CRELLIN, P. K., BRAMMANANTH, R. & COPPEL, R. L. 2011. Decaprenylphosphoryl- β -D-ribose 2'-epimerase, the target of benzothiazinones and dinitrobenzamides, is an essential enzyme in *Mycobacterium smegmatis*. *PLoS One*, 6, e16869.

CRELLIN, P. K., KOVACEVIC, S., MARTIN, K. L., BRAMMANANTH, R., MORITA, Y. S., BILLMAN-JACOB, H., MCCONVILLE, M. J. & COPPEL, R. L.

2008. Mutations in *pimE* restore lipoarabinomannan synthesis and growth in a *Mycobacterium smegmatis* *lpqW* mutant. *J Bacteriol*, 190, 3690-9.
- CRUBEZY, E., LUDES, B., POVEDA, J. D., CLAYTON, J., CROUAU-ROY, B. & MONTAGNON, D. 1998. Identification of *Mycobacterium* DNA in an Egyptian Pott's disease of 5,400 years old. *C R Acad Sci III*, 321, 941-51.
- CUI, T., ZHANG, L., WANG, X. & HE, Z. G. 2009. Uncovering new signaling proteins and potential drug targets through the interactome analysis of *Mycobacterium tuberculosis*. *BMC Genomics*, 10, 118.
- DAFFE, M., BRENNAN, P. J. & MCNEIL, M. 1990. Predominant structural features of the cell wall arabinogalactan of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosyl alditol fragments by gas chromatography/mass spectrometry and by ^1H and ^{13}C NMR analyses. *J Biol Chem*, 265, 6734-43.
- DANIEL, T. M. 2006. The history of tuberculosis. *Respir Med*, 100, 1862-70.
- DAVIS, J. M., CLAY, H., LEWIS, J. L., GHORI, N., HERBOMEL, P. & RAMAKRISHNAN, L. 2002. Real-time visualization of mycobacterium-macrophage interactions leading to initiation of granuloma formation in zebrafish embryos. *Immunity*, 17, 693-702.
- DE SOUZA, M. V., FERREIRA MDE, L., PINHEIRO, A. C., SARAIVA, M. F., DE ALMEIDA, M. V. & VALLE, M. S. 2008. Synthesis and biological aspects of mycolic acids: an important target against *Mycobacterium tuberculosis*. *ScientificWorldJournal*, 8, 720-51.
- DENG, L., MIKUSOVA, K., ROBUCK, K. G., SCHERMAN, M., BRENNAN, P. J. & MCNEIL, M. R. 1995. Recognition of multiple effects of ethambutol on metabolism of mycobacterial cell envelope. *Antimicrob Agents Chemother*, 39, 694-701.
- DINADAYALA, P., KAUR, D., BERG, S., AMIN, A. G., VISSA, V. D., CHATTERJEE, D., BRENNAN, P. J. & CRICK, D. C. 2006. Genetic basis for the synthesis of the immunomodulatory mannose caps of lipoarabinomannan in *Mycobacterium tuberculosis*. *J Biol Chem*, 281, 20027-35.
- DOETSCH, R. N. 1978. Benjamin Marten and his "New Theory of Consumptions". *Microbiol Rev*, 42, 521-8.
- DOHERTY, T. M. & ANDERSEN, P. 2005. Vaccines for tuberculosis: novel concepts and recent progress. *Clin Microbiol Rev*, 18, 687-702.
- DONG, X., BHAMIDI, S., SCHERMAN, M., XIN, Y. & MCNEIL, M. R. 2006. Development of a quantitative assay for mycobacterial endogenous arabinase and ensuing studies of arabinase levels and arabinan metabolism in *Mycobacterium smegmatis*. *Appl Environ Microbiol*, 72, 2601-5.

- DORMAN, S. E. 2010. New diagnostic tests for tuberculosis: bench, bedside, and beyond. *Clin Infect Dis*, 50 Suppl 3, S173-7.
- DOVER, L. G., CERDENO-TARRAGA, A. M., PALLEN, M. J., PARKHILL, J. & BESRA, G. S. 2004. Comparative cell wall core biosynthesis in the mycolated pathogens, *Mycobacterium tuberculosis* and *Corynebacterium diphtheriae*. *FEMS Microbiol Rev*, 28, 225-50.
- DRAPER, P., KHOO, K. H., CHATTERJEE, D., DELL, A. & MORRIS, H. R. 1997. Galactosamine in walls of slow-growing mycobacteria. *Biochem J*, 327 (Pt 2), 519-25.
- DURAND-HEREDIA, J., RIVKIN, E., FAN, G., MORALES, J. & JANAKIRAMAN, A. 2012. Identification of ZapD as a cell division factor that promotes the assembly of FtsZ in *Escherichia coli*. *J Bacteriol*, 194, 3189-98.
- ELLEBY, B., SVENSSON, S., WU, X., STEFANSSON, K., NILSSON, J., HALLEN, D., OPPERMANN, U. & ABRAHIMSEN, L. 2004. High-level production and optimization of monodispersity of 11 β -hydroxysteroid dehydrogenase type 1. *Biochim Biophys Acta*, 1700, 199-207.
- ESCUYER, V. E., LETY, M. A., TORRELLES, J. B., KHOO, K. H., TANG, J. B., RITHNER, C. D., FREHEL, C., MCNEIL, M. R., BRENNAN, P. J. & CHATTERJEE, D. 2001. The role of the *embA* and *embB* gene products in the biosynthesis of the terminal hexaarabinofuranosyl motif of *Mycobacterium smegmatis* arabinogalactan. *J Biol Chem*, 276, 48854-62.
- FORBES, M., KUCK, N. A. & PEETS, E. A. 1965. Effect of ethambutol on nucleic acid metabolism in *Mycobacterium smegmatis* and its reversal by polyamines and divalent cations. *J Bacteriol*, 89, 1299-305.
- GALLI, E. & GERDES, K. 2010. Spatial resolution of two bacterial cell division proteins: ZapA recruits ZapB to the inner face of the Z-ring. *Mol Microbiol*, 76, 1514-26.
- GARISSON, F. H. 1913. An introduction of Medicine, Philadelphia, Saunders.
- GARNIER, T., EIGLMEIER, K., CAMUS, J. C., MEDINA, N., MANSOOR, H., PRYOR, M., DUTHOY, S., GRONDIN, S., LACROIX, C., MONSEMPE, C., SIMON, S., HARRIS, B., ATKIN, R., DOGGETT, J., MAYES, R., KEATING, L., WHEELER, P. R., PARKHILL, J., BARRELL, B. G., COLE, S. T., GORDON, S. V. & HEWINSON, R. G. 2003. The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci U S A*, 100, 7877-82.
- GEORGE, K. M., YUAN, Y., SHERMAN, D. R. & BARRY, C. E., 3RD 1995. The biosynthesis of cyclopropanated mycolic acids in *Mycobacterium tuberculosis*. Identification and functional analysis of CMAS-2. *J Biol Chem*, 270, 27292-8.
- GEORGIADOU, M., CASTAGNINI, M., KARIMOVA, G., LADANT, D. & PELICIC, V. 2012. Large-scale study of the interactions between proteins involved in type IV pilus

biology in *Neisseria meningitidis*: characterization of a subcomplex involved in pilus assembly. *Mol Microbiol*, 84, 857-73.

GILLERON, M., HIMOUDI, N., ADAM, O., CONSTANT, P., VENISSE, A., RIVIERE, M. & PUZO, G. 1997. *Mycobacterium smegmatis* phosphoinositols-glyceroarabinomannans. Structure and localization of alkali-labile and alkali-stable phosphoinositides. *J Biol Chem*, 272, 117-24.

GLICKMAN, M. S. 2003. The *mmaA2* gene of *Mycobacterium tuberculosis* encodes the distal cyclopropane synthase of the α -mycolic acid. *J Biol Chem*, 278, 7844-9.

GLICKMAN, M. S., COX, J. S. & JACOBS, W. R., JR. 2000. A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of *Mycobacterium tuberculosis*. *Mol Cell*, 5, 717-27.

GOODFELLOW, M. & MINNIKIN, D. E. 1977. Nocardioform bacteria. *Annu Rev Microbiol*, 31, 159-80.

GOUDE, R., AMIN, A. G., CHATTERJEE, D. & PARISH, T. 2008. The critical role of *embC* in *Mycobacterium tuberculosis*. *J Bacteriol*, 190, 4335-41.

GOUDE, R., AMIN, A. G., CHATTERJEE, D. & PARISH, T. 2009. The arabinosyltransferase EmbC is inhibited by ethambutol in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*, 53, 4138-46.

GROVER, S., ALDERWICK, L. J., MISHRA, A. K., KRUMBACH, K., MARIENHAGEN, J., EGGELING, L., BHATT, A. & BESRA, G. S. 2014. Benzothiazinones mediate killing of *Corynebacterineae* by blocking decaprenyl phosphate recycling involved in cell wall biosynthesis. *J Biol Chem*, 289, 6177-87.

GUERARDEL, Y., MAES, E., ELASS, E., LEROY, Y., TIMMERMAN, P., BESRA, G. S., LOCHT, C., STRECKER, G. & KREMER, L. 2002. Structural study of lipomannan and lipoarabinomannan from *Mycobacterium chelonae*. Presence of unusual components with α 1,3-mannopyranose side chains. *J Biol Chem*, 277, 30635-48.

GUERIN, M. E., KAUR, D., SOMASHEKAR, B. S., GIBBS, S., GEST, P., CHATTERJEE, D., BRENNAN, P. J. & JACKSON, M. 2009. New insights into the early steps of phosphatidylinositol mannoside biosynthesis in mycobacteria: PimB' is an essential enzyme of *Mycobacterium smegmatis*. *J Biol Chem*, 284, 25687-96.

GUERIN, M. E., KORDULAKOVA, J., ALZARI, P. M., BRENNAN, P. J. & JACKSON, M. 2010. Molecular basis of phosphatidyl-myo-inositol mannoside biosynthesis and regulation in mycobacteria. *J Biol Chem*, 285, 33577-83.

GUERIN, M. E., KORDULAKOVA, J., SCHAEFFER, F., SVETLIKOVA, Z., BUSCHIAZZO, A., GIGANTI, D., GICQUEL, B., MIKUSOVA, K., JACKSON, M. & ALZARI, P. M. 2007. Molecular recognition and interfacial catalysis by the essential phosphatidylinositol mannosyltransferase PimA from mycobacteria. *J Biol Chem*, 282, 20705-14.

- GULATI, S., JAMSHAD, M., KNOWLES, T. J., MORRISON, K. A., DOWNING, R., CANT, N., COLLINS, R., KOENDERINK, J. B., FORD, R. C., OVERDUIN, M., KERR, I. D., DAFFORN, T. R. & ROTHNIE, A. J. 2014. Detergent-free purification of ABC (ATP-binding-cassette) transporters. *Biochem J*, 461, 269-78.
- GURCHA, S. S., BAULARD, A. R., KREMER, L., LOCHT, C., MOODY, D. B., MUHLECKER, W., COSTELLO, C. E., CRICK, D. C., BRENNAN, P. J. & BESRA, G. S. 2002. Ppm1, a novel polyprenol monophosphomannose synthase from *Mycobacterium tuberculosis*. *Biochem J*, 365, 441-50.
- GUTIERREZ, M. C., BRISSE, S., BROSCHE, R., FABRE, M., OMAIS, B., MARMIESSE, M., SUPPLY, P. & VINCENT, V. 2005. Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. *PLoS Pathog*, 1, e5.
- HALE, C. A., SHIOMI, D., LIU, B., BERNHARDT, T. G., MARGOLIN, W., NIKI, H. & DE BOER, P. A. 2011. Identification of *Escherichia coli* ZapC (YcbW) as a component of the division apparatus that binds and bundles FtsZ polymers. *J Bacteriol*, 193, 1393-404.
- HARA, Y., SEKI, M., MATSUOKA, S., HARA, H., YAMASHITA, A. & MATSUMOTO, K. 2008. Involvement of PlsX and the acyl-phosphate dependent sn-glycerol-3-phosphate acyltransferase PlsY in the initial stage of glycerolipid synthesis in *Bacillus subtilis*. *Genes Genet Syst*, 83, 433-42.
- HAYMAN, J. 1984. *Mycobacterium ulcerans*: an infection from Jurassic time? *Lancet*, 2, 1015-6.
- HAZBON, M. H., BOBADILLA DEL VALLE, M., GUERRERO, M. I., VARMA-BASIL, M., FILLIOL, I., CAVATORE, M., COLANGELI, R., SAFI, H., BILLMAN-JACOB, H., LAVENDER, C., FYFE, J., GARCIA-GARCIA, L., DAVIDOW, A., BRIMACOMBE, M., LEON, C. I., PORRAS, T., BOSE, M., CHAVES, F., EISENACH, K. D., SIFUENTES-OSORNIO, J., PONCE DE LEON, A., CAVE, M. D. & ALLAND, D. 2005. Role of *embB* codon 306 mutations in *Mycobacterium tuberculosis* revisited: a novel association with broad drug resistance and IS6110 clustering rather than ethambutol resistance. *Antimicrob Agents Chemother*, 49, 3794-802.
- HERSHKOVITZ, I., DONOGHUE, H. D., MINNIKIN, D. E., BESRA, G. S., LEE, O. Y., GERNAEY, A. M., GALILI, E., ESHED, V., GREENBLATT, C. L., LEMMA, E., BAR-GAL, G. K. & SPIGELMAN, M. 2008. Detection and molecular characterization of 9,000-year-old *Mycobacterium tuberculosis* from a Neolithic settlement in the Eastern Mediterranean. *PLoS One*, 3, e3426.
- HERZOG, H. 1998. History of tuberculosis. *Respiration*, 65, 5-15.
- HOFFMANN, C., LEIS, A., NIEDERWEIS, M., PLITZKO, J. M. & ENGELHARDT, H. 2008. Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc Natl Acad Sci U S A*, 105, 3963-7.

- HUANG, H., BERG, S., SPENCER, J. S., VEREECKE, D., D'HAENZE, W., HOLSTERS, M. & MCNEIL, M. R. 2008. Identification of amino acids and domains required for catalytic activity of DPPR synthase, a cell wall biosynthetic enzyme of *Mycobacterium tuberculosis*. *Microbiology*, 154, 736-43.
- IKEDA, M., WACHI, M., JUNG, H. K., ISHINO, F. & MATSUHASHI, M. 1990. Nucleotide sequence involving *murG* and *murC* in the *mra* gene cluster region of *Escherichia coli*. *Nucleic Acids Res*, 18, 4014.
- JACKSON, M., CRICK, D. C. & BRENNAN, P. J. 2000. Phosphatidylinositol is an essential phospholipid of mycobacteria. *J Biol Chem*, 275, 30092-9.
- JACKSON, M., RAYNAUD, C., LANEELLE, M. A., GUILHOT, C., LAURENT-WINTER, C., ENSERGUEIX, D., GICQUEL, B. & DAFFE, M. 1999. Inactivation of the antigen 85C gene profoundly affects the mycolate content and alters the permeability of the *Mycobacterium tuberculosis* cell envelope. *Mol Microbiol*, 31, 1573-87.
- JAMSHAD, M., LIN, Y. P., KNOWLES, T. J., PARSLOW, R. A., HARRIS, C., WHEATLEY, M., POYNER, D. R., BILL, R. M., THOMAS, O. R., OVERDUIN, M. & DAFFORN, T. R. 2011. Surfactant-free purification of membrane proteins with intact native membrane environment. *Biochem Soc Trans*, 39, 813-8.
- JANKUTE, M., GROVER, S., RANA, A. K. & BESRA, G. S. 2012. Arabinogalactan and lipoarabinomannan biosynthesis: structure, biogenesis and their potential as drug targets. *Future Microbiol*, 7, 129-47.
- JARLIER, V. & NIKAIDO, H. 1994. Mycobacterial cell wall: structure and role in natural resistance to antibiotics. *FEMS Microbiol Lett*, 123, 11-8.
- JIANG, T., HE, L., ZHAN, Y., ZANG, S., MA, Y., ZHAO, X., ZHANG, C. & XIN, Y. 2011. The effect of *MSMEG_6402* gene disruption on the cell wall structure of *Mycobacterium smegmatis*. *Microb Pathog*, 51, 156-60.
- JIN, Y., XIN, Y., ZHANG, W. & MA, Y. 2010. *Mycobacterium tuberculosis* Rv1302 and *Mycobacterium smegmatis* MSMEG_4947 have WecA function and MSMEG_4947 is required for the growth of *M. smegmatis*. *FEMS Microbiol Lett*, 310, 54-61.
- KACEM, R., DE SOUSA-D'AURIA, C., TROPIS, M., CHAMI, M., GOUNON, P., LEBLON, G., HOUSSIN, C. & DAFFE, M. 2004. Importance of mycoloyltransferases on the physiology of *Corynebacterium glutamicum*. *Microbiology*, 150, 73-84.
- KANTARDJIEFF, K. A., KIM, C. Y., NARANJO, C., WALDO, G. S., LEKIN, T., SEGELKE, B. W., ZEMLA, A., PARK, M. S., TERWILLIGER, T. C. & RUPP, B. 2004. *Mycobacterium tuberculosis* RmlC epimerase (Rv3465): a promising drug-target structure in the rhamnose pathway. *Acta Crystallogr D Biol Crystallogr*, 60, 895-902.
- KAPLAN, A. 1947. [The vaccine against tuberculosis by BCG]. *J Med (Oporto)*, 9, 595.

- KARIMOVA, G., DAUTIN, N. & LADANT, D. 2005. Interaction network among *Escherichia coli* membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis. *J Bacteriol*, 187, 2233-43.
- KARIMOVA, G., PIDOUX, J., ULLMANN, A. & LADANT, D. 1998. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci U S A*, 95, 5752-6.
- KARIMOVA, G., ROBICHON, C. & LADANT, D. 2009. Characterization of YmgF, a 72-residue inner membrane protein that associates with the *Escherichia coli* cell division machinery. *J Bacteriol*, 191, 333-46.
- KAUFMANN, S. H. & MCMICHAEL, A. J. 2005. Annulling a dangerous liaison: vaccination strategies against AIDS and tuberculosis. *Nat Med*, 11, S33-44.
- KAUFMANN, S. H. 2001. How can immunology contribute to the control of tuberculosis? *Nat Rev Immunol*, 1, 20-30.
- KAUR, D., GUERIN, M. E., SKOVIEROVA, H., BRENNAN, P. J. & JACKSON, M. 2009. Chapter 2: Biogenesis of the cell wall and other glycoconjugates of *Mycobacterium tuberculosis*. *Adv Appl Microbiol*, 69, 23-78.
- KAUR, D., MCNEIL, M. R., KHOO, K. H., CHATTERJEE, D., CRICK, D. C., JACKSON, M. & BRENNAN, P. J. 2007. New insights into the biosynthesis of mycobacterial lipomannan arising from deletion of a conserved gene. *J Biol Chem*, 282, 27133-40.
- KEERS, R. Y. 1978. Pulmonary tuberculosis : a journey down the centuries, London, Baillière Tindall.
- KHOO, K. H., DELL, A., MORRIS, H. R., BRENNAN, P. J. & CHATTERJEE, D. 1995. Inositol phosphate capping of the nonreducing termini of lipoarabinomannan from rapidly growing strains of *Mycobacterium*. *J Biol Chem*, 270, 12380-9.
- KHOO, K. H., DELL, A., MORRIS, H. R., BRENNAN, P. J. & CHATTERJEE, D. 1995b. Structural definition of acylated phosphatidylinositol mannosides from *Mycobacterium tuberculosis*: definition of a common anchor for lipomannan and lipoarabinomannan. *Glycobiology*, 5, 117-27.
- KHOO, K. H., TANG, J. B. & CHATTERJEE, D. 2001. Variation in mannose-capped terminal arabinan motifs of lipoarabinomannans from clinical isolates of *Mycobacterium tuberculosis* and *Mycobacterium avium* complex. *J Biol Chem*, 276, 3863-71.
- KILBURN, J. O. & GREENBERG, J. 1977. Effect of ethambutol on the viable cell count in *Mycobacterium smegmatis*. *Antimicrob Agents Chemother*, 11(3):534-540.
- KILBURN, J. O. & TAKAYAMA, K. 1981. Effects of ethambutol on accumulation and secretion of trehalose mycolates and free mycolic acid in *Mycobacterium smegmatis*. *Antimicrob Agents Chemother*, 20, 401-4.

- KILBURN, J. O., TAKAYAMA, K., ARMSTRONG, E. L. & GREENBERG, J. 1981. Effects of ethambutol on phospholipid metabolism in *Mycobacterium smegmatis*. *Antimicrob Agents Chemother*, 19, 346-8.
- KIM, W. J., SON, W. S., AHN, D. H., IM, H., AHN, H. C. & LEE, B. J. 2014. Solution structure of *Rv0569*, potent hypoxic signal transduction protein, from *Mycobacterium tuberculosis*. *Tuberculosis* (Edinb), 94, 43-50.
- KING, M. J. & PARK, W. H. 1929. Effect of Calmette's BCG Vaccine on Experimental Animals. *Am J Public Health Nations Health*, 19, 179-92.
- KLEIN, D. J. & FERRE-D'AMARE, A. R. 2006. Structural basis of *glmS* ribozyme activation by glucosamine-6-phosphate. *Science*, 313, 1752-6.
- KLOPPER, M., WARREN, R. M., HAYES, C., GEY VAN PITTIUS, N. C., STREICHER, E. M., MULLER, B., SIRGEL, F. A., CHABULA-NXIWENI, M., HOOSAIN, E., COETZEE, G., DAVID VAN HELDEN, P., VICTOR, T. C. & TROLLIP, A. P. 2013. Emergence and spread of extensively and totally drug-resistant tuberculosis, South Africa. *Emerg Infect Dis*, 19, 449-55.
- KNOWLES, T. J., FINKA, R., SMITH, C., LIN, Y. P., DAFFORN, T. & OVERDUIN, M. 2009. Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J Am Chem Soc*, 131, 7484-5.
- KOCH, R. 1882. Die Aetiologie der Tuberculose. *Berl Klinische Wochenschr*, 19, 221-230.
- KOLK, A. H., HO, M. L., KLATSER, P. R., EGGELTE, T. A., KUIJPER, S., DE JONGE, S. & VAN LEEUWEN, J. 1984. Production and characterization of monoclonal antibodies to *Mycobacterium tuberculosis*, *M. bovis* (BCG) and *M. leprae*. *Clin Exp Immunol*, 58, 511-21.
- KORDULAKOVA, J., GILLERON, M., MIKUSOVA, K., PUZO, G., BRENNAN, P. J., GICQUEL, B. & JACKSON, M. 2002. Definition of the first mannosylation step in phosphatidylinositol mannoside synthesis. PimA is essential for growth of mycobacteria. *J Biol Chem*, 277, 31335-44.
- KORDULAKOVA, J., GILLERON, M., PUZO, G., BRENNAN, P. J., GICQUEL, B., MIKUSOVA, K. & JACKSON, M. 2003. Identification of the required acyltransferase step in the biosynthesis of the phosphatidylinositol mannosides of mycobacterium species. *J Biol Chem*, 278, 36285-95.
- KOVACEVIC, S., ANDERSON, D., MORITA, Y. S., PATTERSON, J., HAITES, R., MCMILLAN, B. N., COPPEL, R., MCCONVILLE, M. J. & BILLMAN-JACOB, H. 2006. Identification of a novel protein with a role in lipoarabinomannan biosynthesis in mycobacteria. *J Biol Chem*, 281, 9011-7.
- KREMER, L., DOVER, L. G., MOREHOUSE, C., HITCHIN, P., EVERETT, M., MORRIS, H. R., DELL, A., BRENNAN, P. J., MCNEIL, M. R., FLAHERTY, C.,

- DUNCAN, K. & BESRA, G. S. 2001. Galactan biosynthesis in *Mycobacterium tuberculosis*. Identification of a bifunctional UDP-galactofuranosyltransferase. *J Biol Chem*, 276, 26430-40.
- KREMER, L., GURCHA, S. S., BIFANI, P., HITCHEN, P. G., BAULARD, A., MORRIS, H. R., DELL, A., BRENNAN, P. J. & BESRA, G. S. 2002. Characterization of a putative α -mannosyltransferase involved in phosphatidylinositol trimannoside biosynthesis in *Mycobacterium tuberculosis*. *Biochem J*, 363, 437-47.
- KREMER, L., MAUGHAN, W. N., WILSON, R. A., DOVER, L. G. & BESRA, G. S. 2002. The *M. tuberculosis* antigen 85 complex and mycolyltransferase activity. *Lett Appl Microbiol*, 34, 233-7.
- KUMAR, V., SARAVANAN, P., ARVIND, A. & MOHAN, C. G. 2011. Identification of hotspot regions of MurB oxidoreductase enzyme using homology modeling, molecular dynamics and molecular docking techniques. *J Mol Model*, 17, 939-53.
- LAI, X., WU, J., CHEN, S., ZHANG, X. & WANG, H. 2008. Expression, purification, and characterization of a functionally active *Mycobacterium tuberculosis* UDP-glucose pyrophosphorylase. *Protein Expr Purif*, 61, 50-6.
- LARROUY-MAUMUS, G., SKOVIEROVA, H., DHOUIB, R., ANGALA, S. K., ZUBEROGOITIA, S., PHAM, H., VILLELA, A. D., MIKUSOVA, K., NOGUERA, A., GILLERON, M., VALENTINOVA, L., KORDULAKOVA, J., BRENNAN, P. J., PUZO, G., NIGOU, J. & JACKSON, M. 2012. A small multidrug resistance-like transporter involved in the arabinosylation of arabinogalactan and lipoarabinomannan in mycobacteria. *J Biol Chem*, 287, 39933-41.
- LEA-SMITH, D. J., MARTIN, K. L., PYKE, J. S., TULL, D., MCCONVILLE, M. J., COPPEL, R. L. & CRELLIN, P. K. 2008. Analysis of a new mannosyltransferase required for the synthesis of phosphatidylinositol mannosides and lipoarabinomannan reveals two lipomannan pools in corynebacterineae. *J Biol Chem*, 283, 6773-82.
- LEDERER, E., ADAM, A., CIORBARU, R., PETIT, J. F. & WIETZERBIN, J. 1975. Cell walls of *Mycobacteria* and related organisms; chemistry and immunostimulant properties. *Mol Cell Biochem*, 7, 87-104.
- LEE, A. S., OTHMAN, S. N., HO, Y. M. & WONG, S. Y. 2004. Novel mutations within the *embB* gene in ethambutol-susceptible clinical isolates of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*, 48, 4447-9.
- LEE, A., WU, S. W., SCHERMAN, M. S., TORRELLES, J. B., CHATTERJEE, D., MCNEIL, M. R. & KHOO, K. H. 2006. Sequencing of oligoarabinosyl units released from mycobacterial arabinogalactan by endogenous arabinanase: identification of distinctive and novel structural motifs. *Biochemistry*, 45, 15817-28.
- LEE, H. N., JUNG, K. E., KO, I. J., BAIK, H. S. & OH, J. I. 2012. Protein-protein interactions between histidine kinases and response regulators of *Mycobacterium tuberculosis* H37Rv. *J Microbiol*, 50, 270-7.

- LEHMANN, K. B. & NEUMANN, R. 1896. Lehmann's Med Handatlanten. Atlas und grundriss der bakteriologie und lehrbuch der speciellen bakteriologischen diagnostik, 1, 1-448.
- LEMAIRE, H. G. & MULLER-HILL, B. 1986. Nucleotide sequences of the *galE* gene and the *galT* gene of *E. coli*. *Nucleic Acids Res*, 14, 7705-11.
- LETY, M. A., NAIR, S., BERCHE, P. & ESCUYER, V. 1997. A single point mutation in the *embB* gene is responsible for resistance to ethambutol in *Mycobacterium smegmatis*. *Antimicrob Agents Chemother*, 41, 2629-33.
- LEVY-FREBAULT, V. V. & PORTAELS, F. 1992. Proposed minimal standards for the genus *Mycobacterium* and for description of new slowly growing *Mycobacterium* species. *Int J Syst Bacteriol*, 42, 315-23.
- LI, S., KANG, J., YU, W., ZHOU, Y., ZHANG, W., XIN, Y. & MA, Y. 2012. Identification of *M. tuberculosis* Rv3441c and *M. smegmatis* MSMEG_1556 and essentiality of *M. smegmatis* MSMEG_1556. *PLoS One*, 7, e42769.
- LI, W., XIN, Y., MCNEIL, M. R. & MA, Y. 2006. *rmlB* and *rmlC* genes are essential for growth of mycobacteria. *Biochem Biophys Res Commun*, 342, 170-8.
- LIGER, D., MASSON, A., BLANOT, D., VAN HEIJENOORT, J. & PARQUET, C. 1995. Over-production, purification and properties of the uridine-diphosphate-N-acetylmuramate:L-alanine ligase from *Escherichia coli*. *Eur J Biochem*, 230, 80-7.
- LIU, J. & MUSHEGIAN, A. 2003. Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci*, 12, 1418-31.
- LIU, J., ROSENBERG, E. Y. & NIKAIDO, H. 1995. Fluidity of the lipid domain of cell wall from *Mycobacterium chelonae*. *Proc Natl Acad Sci U S A*, 92, 11254-8.
- LUO, H., PANG, L. & XIE, J. 2012. Biosynthesis and regulation of mycolic acids in *Mycobacterium tuberculosis*-a review. *Wei Sheng Wu Xue Bao*, 52, 146-51.
- LUSTGARTEN, S. 1884. Ueber spezifische bazillen in syphilitischen
- LUTKENHAUS, J., PICHOFF, S. & DU, S. 2012. Bacterial cytokinesis: From Z ring to divisome. *Cytoskeleton (Hoboken)*, 69, 778-90.
- MA, Y., MILLS, J. A., BELISLE, J. T., VISSA, V., HOWELL, M., BOWLIN, K., SCHERMAN, M. S. & MCNEIL, M. 1997. Determination of the pathway for rhamnose biosynthesis in mycobacteria: cloning, sequencing and expression of the *Mycobacterium tuberculosis* gene encoding α -D-glucose-1-phosphate thymidyltransferase. *Microbiology*, 143 (Pt 3), 937-45.
- MA, Y., PAN, F. & MCNEIL, M. 2002. Formation of dTDP-rhamnose is essential for growth of mycobacteria. *J Bacteriol*, 184, 3392-5.

- MA, Y., STERN, R. J., SCHERMAN, M. S., VISSA, V. D., YAN, W., JONES, V. C., ZHANG, F., FRANZBLAU, S. G., LEWIS, W. H. & MCNEIL, M. R. 2001. Drug targeting *Mycobacterium tuberculosis* cell wall synthesis: genetics of dTDP-rhamnose synthetic enzymes and development of a microtiter plate-based screen for inhibitors of conversion of dTDP-glucose to dTDP-rhamnose. *Antimicrob Agents Chemother*, 45, 1407-16.
- MAHAPATRA, S., CRICK, D. C. & BRENNAN, P. J. 2000. Comparison of the UDP-N-acetylmuramate:L-alanine ligase enzymes from *Mycobacterium tuberculosis* and *Mycobacterium leprae*. *J Bacteriol*, 182, 6827-30.
- MAI, D., JONES, J., RODGERS, J. W., HARTMAN, J. L. T., KUTSCH, O. & STEYN, A. J. 2011. A screen to identify small molecule inhibitors of protein-protein interactions in mycobacteria. *Assay Drug Dev Technol*, 9, 299-310.
- MAKAROV, V., MANINA, G., MIKUSOVA, K., MOLLMANN, 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., MCKINNEY, J. D., BRODIN, P., CHRISTOPHE, T., WADDELL, S., BUTCHER, P., ALBRETHSEN, J., ROSENKRANDS, I., BROSCHE, R., NANDI, V., BHARATH, S., GAONKAR, S., SHANDIL, R. K., BALASUBRAMANIAN, V., BALGANESH, T., TYAGI, S., GROSSET, J., RICCARDI, G. & COLE, S. T. 2009. Benzothiazinones kill *Mycobacterium tuberculosis* by blocking arabinan synthesis. *Science*, 324, 801-4.
- MARTEYN, B. S., KARIMOVA, G., FENTON, A. K., GAZI, A. D., WEST, N., TOUQUI, L., PREVOST, M. C., BETTON, J. M., POYRAZ, O., LADANT, D., GERDES, K., SANSONETTI, P. J. & TANG, C. M. 2014. ZapE is a novel cell division protein interacting with FtsZ and modulating the Z-ring dynamics. *MBio*, 5, e00022-14.
- MARUYAMA, I. N., YAMAMOTO, A. H. & HIROTA, Y. 1988. Determination of gene products and coding regions from the *murE-murF* region of *Escherichia coli*. *J Bacteriol*, 170, 3786-8.
- MAXSON, M. E. & DARWIN, A. J. 2006. PspB and PspC of *Yersinia enterocolitica* are dual function proteins: regulators and effectors of the phage-shock-protein response. *Mol Microbiol*, 59, 1610-23.
- MCNEIL, M. R., ROBUCK, K. G., HARTER, M. & BRENNAN, P. J. 1994. Enzymatic evidence for the presence of a critical terminal hexa-arabinoside in the cell walls of *Mycobacterium tuberculosis*. *Glycobiology*, 4, 165-73.
- MCNEIL, M., DAFFE, M. & BRENNAN, P. J. 1990. Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. *J Biol Chem*, 265, 18200-6.
- MCNEIL, M., DAFFE, M. & BRENNAN, P. J. 1991. Location of the mycolyl ester substituents in the cell walls of mycobacteria. *J Biol Chem*, 266, 13217-23.

- MCNEIL, M., DAFPE, M. & BRENNAN, P. J. 1991. Location of the mycolyl ester substituents in the cell walls of mycobacteria. *J Biol Chem*, 266, 13217-23.
- MCNEIL, M., WALLNER, S. J., HUNTER, S. W. & BRENNAN, P. J. 1987. Demonstration that the galactosyl and arabinosyl residues in the cell-wall arabinogalactan of *Mycobacterium leprae* and *Mycobacterium tuberculosis* are furanoid. *Carbohydr Res*, 166, 299-308.
- MENGIN-LECREULX, D. & VAN HEIJENOORT, J. 1993. Identification of the *glmU* gene encoding *N*-acetylglucosamine-1-phosphate uridyltransferase in *Escherichia coli*. *J Bacteriol*, 175, 6150-7.
- MENGIN-LECREULX, D. & VAN HEIJENOORT, J. 1994. Copurification of glucosamine-1-phosphate acetyltransferase and *N*-acetylglucosamine-1-phosphate uridyltransferase activities of *Escherichia coli*: characterization of the *glmU* gene product as a bifunctional enzyme catalyzing two subsequent steps in the pathway for UDP-*N*-acetylglucosamine synthesis. *J Bacteriol*, 176, 5788-95.
- MENGIN-LECREULX, D. & VAN HEIJENOORT, J. 1996. Characterization of the essential gene *glmM* encoding phosphoglucosamine mutase in *Escherichia coli*. *J Biol Chem*, 271, 32-9.
- MENGIN-LECREULX, D., PARQUET, C., DESVIAT, L. R., PLA, J., FLOURET, B., AYALA, J. A. & VAN HEIJENOORT, J. 1989. Organization of the *murE-murG* region of *Escherichia coli*: identification of the *murD* gene encoding the D-glutamic-acid-adding enzyme. *J Bacteriol*, 171, 6126-34.
- MENICHE, X., DE SOUSA-D'AURIA, C., VAN-DER-REST, B., BHAMIDI, S., HUC, E., HUANG, H., DE PAEPE, D., TROPIS, M., MCNEIL, M., DAFPE, M. & HOUSSIN, C. 2008. Partial redundancy in the synthesis of the D-arabinose incorporated in the cell wall arabinan of *Corynebacterineae*. *Microbiology*, 154, 2315-26.
- MIGLIORI, G. B., DE IACO, G., BESOZZI, G., CENTIS, R. & CIRILLO, D. M. 2007. First tuberculosis cases in Italy resistant to all tested drugs. *Euro Surveill*, 12, E070517 1.
- MIKUSOVA, K., BELANOVA, M., KORDULAKOVA, J., HONDA, K., MCNEIL, M. R., MAHAPATRA, S., CRICK, D. C. & BRENNAN, P. J. 2006. Identification of a novel galactosyl transferase involved in biosynthesis of the mycobacterial cell wall. *J Bacteriol*, 188, 6592-8.
- MIKUSOVA, K., HUANG, H., YAGI, T., HOLSTERS, M., VEREECKE, D., D'HAENZE, W., SCHERMAN, M. S., BRENNAN, P. J., MCNEIL, M. R. & CRICK, D. C. 2005. Decaprenylphosphoryl arabinofuranose, the donor of the D-arabinofuranosyl residues of mycobacterial arabinan, is formed via a two-step epimerization of decaprenylphosphoryl ribose. *J Bacteriol*, 187, 8020-5.
- MIKUSOVA, K., MIKUS, M., BESRA, G. S., HANCOCK, I. & BRENNAN, P. J. 1996. Biosynthesis of the linkage region of the mycobacterial cell wall. *J Biol Chem*, 271, 7820-8.

MILLS, J. A., MOTICHKA, K., JUCKER, M., WU, H. P., UHLIK, B. C., STERN, R. J., SCHERMAN, M. S., VISSA, V. D., PAN, F., KUNDU, M., MA, Y. F. & MCNEIL, M. 2004. Inactivation of the mycobacterial rhamnosyltransferase, which is needed for the formation of the arabinogalactan-peptidoglycan linker, leads to irreversible loss of viability. *J Biol Chem*, 279, 43540-6.

MINNIKIN, D. E. 1982. Lipids: complex lipids, their chemistry, biosynthesis and roles. In *The Biology of the Mycobacteria*. pp. 95-184. Edited by C. Ratledge & J. Stanford. London:Academic Press Inc. Ltd.

MINNIKIN, D. E., KREMER, L., DOVER, L. G. & BESRA, G. S. 2002. The methyl-branched fortifications of *Mycobacterium tuberculosis*. *Chem Biol*, 9, 545-53.

MISHRA, A. K., ALDERWICK, L. J., RITTMANN, D., TATITURI, R. V., NIGOU, J., GILLERON, M., EGGELING, L. & BESRA, G. S. 2007. Identification of an $\alpha(1\rightarrow6)$ mannosyltransferase (MptA), involved in *Corynebacterium glutamicum* lipomanann biosynthesis, and identification of its orthologue in *Mycobacterium tuberculosis*. *Mol Microbiol*, 65, 1503-17.

MISHRA, A. K., ALDERWICK, L. J., RITTMANN, D., WANG, C., BHATT, A., JACOBS, W. R., JR., TAKAYAMA, K., EGGELING, L. & BESRA, G. S. 2008a. Identification of a novel $\alpha(1\rightarrow6)$ mannosyltransferase MptB from *Corynebacterium glutamicum* by deletion of a conserved gene, *NCgl1505*, affords a lipomannan- and lipoarabinomannan-deficient mutant. *Mol Microbiol*, 68, 1595-613.

MISHRA, A. K., ALVES, J. E., KRUMBACH, K., NIGOU, J., CASTRO, A. G., GEURTSSEN, J., EGGELING, L., SARAIVA, M. & BESRA, G. S. 2012. Differential arabinan capping of lipoarabinomannan modulates innate immune responses and impacts T helper cell differentiation. *J Biol Chem*, 287, 44173-83.

MISHRA, A. K., BATT, S., KRUMBACH, K., EGGELING, L. & BESRA, G. S. 2009. Characterization of the *Corynebacterium glutamicum* $\Delta pimB'$ $\Delta mgtA$ double deletion mutant and the role of *Mycobacterium tuberculosis* orthologues *Rv2188c* and *Rv0557* in glycolipid biosynthesis. *J Bacteriol*, 191, 4465-72.

MISHRA, A. K., BATT, S.M., ALDERWICK, L.J., FUTTERER, K., BESRA, G.S. 2011. Bacterial lipoarabinomannan: structure to biogenesis. *Bacterial Glycomics: Current Research, Technology and Applications*, 97-109.

MISHRA, A. K., DRIESSEN, N. N., APPELMELK, B. J. & BESRA, G. S. 2011a. Lipoarabinomannan and related glycoconjugates: structure, biogenesis and role in *Mycobacterium tuberculosis* physiology and host-pathogen interaction. *FEMS Microbiol Rev*, 35, 1126-57.

MISHRA, A. K., KLEIN, C., GURCHA, S. S., ALDERWICK, L. J., BABU, P., HITCHEN, P. G., MORRIS, H. R., DELL, A., BESRA, G. S. & EGGELING, L. 2008b. Structural characterization and functional properties of a novel lipomannan variant isolated from a *Corynebacterium glutamicum* *pimB'* mutant. *Antonie Van Leeuwenhoek*, 94, 277-87.

- MISHRA, A. K., KRUMBACH, K., RITTMANN, D., APPELMELK, B., PATHAK, V., PATHAK, A. K., NIGOU, J., GEURTSSEN, J., EGGELING, L. & BESRA, G. S. 2011b. Lipoarabinomannan biosynthesis in *Corynebacterineae*: the interplay of two $\alpha(1 \rightarrow 2)$ -mannopyranosyltransferases MptC and MptD in mannan branching. *Mol Microbiol*, 80, 1241-59.
- MORITA, Y. S., SENA, C. B., WALLER, R. F., KUROKAWA, K., SERNEE, M. F., NAKATANI, F., HAITES, R. E., BILLMAN-JACOB, H., MCCONVILLE, M. J., MAEDA, Y. & KINOSHITA, T. 2006. PimE is a polyprenol-phosphate-mannose-dependent mannosyltransferase that transfers the fifth mannose of phosphatidylinositol mannoside in mycobacteria. *J Biol Chem*, 281, 25143-55.
- MOVAHEDZADEH, F., SMITH, D. A., NORMAN, R. A., DINADAYALA, P., MURRAY-RUST, J., RUSSELL, D. G., KENDALL, S. L., RISON, S. C., MCALISTER, M. S., BANCROFT, G. J., MCDONALD, N. Q., DAFFE, M., AV-GAY, Y. & STOKER, N. G. 2004. The *Mycobacterium tuberculosis ino1* gene is essential for growth and virulence. *Mol Microbiol*, 51, 1003-14.
- MOVAHEDZADEH, F., WHEELER, P. R., DINADAYALA, P., AV-GAY, Y., PARISH, T., DAFFE, M. & STOKER, N. G. 2010. Inositol monophosphate phosphatase genes of *Mycobacterium tuberculosis*. *BMC Microbiol*, 10, 50.
- MUNSHI, T., GUPTA, A., EVANGELOPOULOS, D., GUZMAN, J. D., GIBBONS, S., KEEP, N. H. & BHAKTA, S. 2013. Characterisation of ATP-dependent Mur ligases involved in the biogenesis of cell wall peptidoglycan in *Mycobacterium tuberculosis*. *PLoS One*, 8, e60143.
- MURPHY, K. P. 2001. Stabilization of protein structure. *Methods Mol Biol*, 168, 1-16.
- NASSAU, P. M., MARTIN, S. L., BROWN, R. E., WESTON, A., MONSEY, D., MCNEIL, M. R. & DUNCAN, K. 1996. Galactofuranose biosynthesis in *Escherichia coli* K-12: identification and cloning of UDP-galactopyranose mutase. *J Bacteriol*, 178, 1047-52.
- NERLICH, A. G., HAAS, C. J., ZINK, A., SZEIMIES, U. & HAGEDORN, H. G. 1997. Molecular evidence for tuberculosis in an ancient Egyptian mummy. *Lancet*, 350, 1404.
- NGUYEN, L. & THOMPSON, C. J. 2006. Foundations of antibiotic resistance in bacterial physiology: the mycobacterial paradigm. *Trends Microbiol*, 14, 304-12.
- NIESEN, F. H., BERGLUND, H. & VEDADI, M. 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc*, 2, 2212-21.
- NIGOU, J. & BESRA, G. S. 2002. Characterization and regulation of inositol monophosphatase activity in *Mycobacterium smegmatis*. *Biochem J*, 361, 385-90.

- NIGOU, J., GILLERON, M. & PUZO, G. 1999. Lipoarabinomannans: characterization of the multiacylated forms of the phosphatidyl-myo-inositol anchor by NMR spectroscopy. *Biochem J*, 337 (Pt 3), 453-60.
- NIGOU, J., GILLERON, M., CAHUZAC, B., BOUNERY, J. D., HEROLD, M., THURNHER, M. & PUZO, G. 1997. The phosphatidyl-*myo*-inositol anchor of the lipoarabinomannans from *Mycobacterium bovis* bacillus Calmette Guerin. Heterogeneity, structure, and role in the regulation of cytokine secretion. *J Biol Chem*, 272, 23094-103.
- ORWICK-RYDMARK, M., LOVETT, J. E., GRAZIADEI, A., LINDHOLM, L., HICKS, M. R. & WATTS, A. 2012. Detergent-free incorporation of a seven-transmembrane receptor protein into nanosized bilayer Lipodisq particles for functional and biophysical studies. *Nano Lett*, 12, 4687-92.
- PAI, M. 2013. Diagnosis of pulmonary tuberculosis: recent advances. *J Indian Med Assoc*, 111, 332-6.
- PAN, F., JACKSON, M., MA, Y. & MCNEIL, M. 2001. Cell wall core galactofuran synthesis is essential for growth of mycobacteria. *J Bacteriol*, 183, 3991-8.
- PAULIN, S., JAMSHAD, M., DAFFORN, T. R., GARCIA-LARA, J., FOSTER, S. J., GALLEY, N. F., ROPER, D. I., ROSADO, H. & TAYLOR, P. W. 2014. Surfactant-free purification of membrane protein complexes from bacteria: application to the staphylococcal penicillin-binding protein complex PBP2/PBP2a. *Nanotechnology*, 25, 285101.
- PEASE, A. 1940. Some remarks on the diagnosis and treatment of tuberculosis in antiquity. *Isis*, 31, 380-393.
- PENG, W., ZOU, L., BHAMIDI, S., MCNEIL, M. R. & LOWARY, T. L. 2012. The galactosamine residue in mycobacterial arabinogalactan is α -linked. *J Org Chem*, 77, 9826-32.
- PETIT, J. F., ADAM, A., WIETZERBIN-FALSZPAN, J., LEDERER, E. & GHUYSEN, J. M. 1969. Chemical structure of the cell wall of *Mycobacterium smegmatis*. I. Isolation and partial characterization of the peptidoglycan. *Biochem Biophys Res Commun*, 35, 478-85.
- PITARQUE, S., LARROUY-MAUMUS, G., PAYRE, B., JACKSON, M., PUZO, G. & NIGOU, J. 2008. The immunomodulatory lipoglycans, lipoarabinomannan and lipomannan, are exposed at the mycobacterial cell surface. *Tuberculosis (Edinb)*, 88, 560-5.
- POPOT, J. L. 2010. Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annu Rev Biochem*, 79, 737-75.
- PORTEVIN, D., DE SOUSA-D'AURIA, C., HOUSSIN, C., GRIMALDI, C., CHAMI, M., DAFPE, M. & GUILHOT, C. 2004. A polyketide synthase catalyzes the last

condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. *Proc Natl Acad Sci U S A*, 101, 314-9.

POSO, H., PAULIN, L. & BRANDER, E. 1983. Specific inhibition of spermidine synthase from mycobacteria by ethambutol. *Lancet*, 2, 1418.

PRIYA, R., BIUKOVIC, G., MANIMEKALAI, M. S., LIM, J., RAO, S. P. & GRUBER, G. 2013. Solution structure of subunit γ (γ (1-204)) of the *Mycobacterium tuberculosis* F-ATP synthase and the unique loop of γ (165-178), representing a novel TB drug target. *J Bioenerg Biomembr*, 45, 121-9.

PUCCI, M. J., DISCOTTO, L. F. & DOUGHERTY, T. J. 1992. Cloning and identification of the *Escherichia coli murB* DNA sequence, which encodes UDP-N-acetylenolpyruvoylglucosamine reductase. *J Bacteriol*, 174, 1690-3.

QU, H., XIN, Y., DONG, X. & MA, Y. 2007. An *rmlA* gene encoding D-glucose-1-phosphate thymidyltransferase is essential for mycobacterial growth. *FEMS Microbiol Lett*, 275, 237-43.

QURESHI, N., TAKAYAMA, K., JORDI, H. C. & SCHNOES, H. K. 1978. Characterization of the purified components of a new homologous series of α -mycolic acids from *Mycobacterium tuberculosis* H37Ra. *J Biol Chem*, 253, 5411-7.

RAMAN, K. & CHANDRA, N. 2008. *Mycobacterium tuberculosis* interactome analysis unravels potential pathways to drug resistance. *BMC Microbiol*, 8, 234.

RAMASWAMY, S. V., AMIN, A. G., GOKSEL, S., STAGER, C. E., DOU, S. J., EL SAHLY, H., MOGHAZEH, S. L., KREISWIRTH, B. N. & MUSSER, J. M. 2000. Molecular genetic analysis of nucleotide polymorphisms associated with ethambutol resistance in human isolates of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*, 44, 326-36.

RANA, A. K., SINGH, A., GURCHA, S. S., COX, L. R., BHATT, A. & BESRA, G. S. 2012. Ppm1-encoded polyprenyl monophosphomannose synthase activity is essential for lipoglycan synthesis and survival in mycobacteria. *PLoS One*, 7, e48211.

RAY, S., TALUKDAR, A., KUNDU, S., KHANRA, D. & SONTHALIA, N. 2013. Diagnosis and management of miliary tuberculosis: current state and future perspectives. *Ther Clin Risk Manag*, 9, 9-26.

RAYMOND, J. B., MAHAPATRA, S., CRICK, D. C. & PAVELKA, M. S., JR. 2005. Identification of the *namH* gene, encoding the hydroxylase responsible for the N-glycolylation of the mycobacterial peptidoglycan. *J Biol Chem*, 280, 326-33.

RITCHIE, T. K., GRINKOVA, Y. V., BAYBURT, T. H., DENISOV, I. G., ZOLNERCIKS, J. K., ATKINS, W. M. & SLIGAR, S. G. 2009. Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol*, 464, 211-31.

- ROBICHON, C., KARIMOVA, G., BECKWITH, J. & LADANT, D. 2011. Role of leucine zipper motifs in association of the *Escherichia coli* cell division proteins FtsL and FtsB. *J Bacteriol*, 193, 4988-92.
- ROSE, N. L., COMPLETO, G. C., LIN, S. J., MCNEIL, M., PALCIC, M. M. & LOWARY, T. L. 2006. Expression, purification, and characterization of a galactofuranosyltransferase involved in *Mycobacterium tuberculosis* arabinogalactan biosynthesis. *J Am Chem Soc*, 128, 6721-9.
- RUSSELL, D. G. 2001. *Mycobacterium tuberculosis*: here today, and here tomorrow. *Nat Rev Mol Cell Biol*, 2, 569-77.
- RUSSELL, D. G. 2007. Who puts the tubercle in tuberculosis? *Nat Rev Microbiol*, 5, 39-47.
- SAFI, H., SAYERS, B., HAZBON, M. H. & ALLAND, D. 2008. Transfer of *embB* codon 306 mutations into clinical *Mycobacterium tuberculosis* strains alters susceptibility to ethambutol, isoniazid, and rifampin. *Antimicrob Agents Chemother*, 52, 2027-34.
- SALO, W. L., AUFDERHEIDE, A. C., BUIKSTRA, J. & HOLCOMB, T. A. 1994. Identification of *Mycobacterium tuberculosis* DNA in a pre-Columbian Peruvian mummy. *Proc Natl Acad Sci U S A*, 91, 2091-4.
- SANDERS, D. A., STAINES, A. G., MCMAHON, S. A., MCNEIL, M. R., WHITFIELD, C. & NAISMITH, J. H. 2001. UDP-galactopyranose mutase has a novel structure and mechanism. *Nat Struct Biol*, 8, 858-63.
- SCHAEFFER, M. L., KHOO, K. H., BESRA, G. S., CHATTERJEE, D., BRENNAN, P. J., BELISLE, J. T. & INAMINE, J. M. 1999. The *pimB* gene of *Mycobacterium tuberculosis* encodes a mannosyltransferase involved in lipoarabinomannan biosynthesis. *J Biol Chem*, 274, 31625-31.
- SCHAFER, A., TAUCH, A., JAGER, W., KALINOWSKI, J., THIERBACH, G. & PUHLER, A. 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene*, 145, 69-73.
- SCHLEIFER, K. H. & KANDLER, O. 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev*, 36, 407-77.
- SEIDEL, M., ALDERWICK, L. J., BIRCH, H. L., SAHM, H., EGGELING, L. & BESRA, G. S. 2007. Identification of a novel arabinofuranosyltransferase AftB involved in a terminal step of cell wall arabinan biosynthesis in *Corynebacterianae*, such as *Corynebacterium glutamicum* and *Mycobacterium tuberculosis*. *J Biol Chem*, 282, 14729-40.
- SEIDEL, M., ALDERWICK, L. J., SAHM, H., BESRA, G. S. & EGGELING, L. 2007b. Topology and mutational analysis of the single Emb arabinofuranosyltransferase of

Corynebacterium glutamicum as a model of Emb proteins of *Mycobacterium tuberculosis*. *Glycobiology*, 17, 210-9.

SEVERN, W. B., FURNEAUX, R. H., FALSHAW, R. & ATKINSON, P. H. 1998. Chemical and spectroscopic characterisation of the phosphatidylinositol manno-oligosaccharides from *Mycobacterium bovis* AN5 and WAg201 and *Mycobacterium smegmatis* mc²155. *Carbohydr Res*, 308, 397-408.

SHA, S., ZHOU, Y., XIN, Y. & MA, Y. 2012. Development of a colorimetric assay and kinetic analysis for *Mycobacterium tuberculosis* D-glucose-1-phosphate thymidyltransferase. *J Biomol Screen*, 17, 252-7.

SHARMA, S. K. & MOHAN, A. 2013. Tuberculosis: From an incurable scourge to a curable disease - journey over a millennium. *Indian J Med Res*, 137, 455-93.

SHARMA, S. K., MOHAN, A., SHARMA, A. & MITRA, D. K. 2005. Miliary tuberculosis: new insights into an old disease. *Lancet Infect Dis*, 5, 415-30.

SHEN, X., SHEN, G. M., WU, J., GUI, X. H., LI, X., MEI, J., DERIEMER, K. & GAO, Q. 2007. Association between *embB* codon 306 mutations and drug resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*, 51, 2618-20.

SHI, L., BERG, S., LEE, A., SPENCER, J. S., ZHANG, J., VISSA, V., MCNEIL, M. R., KHOO, K. H. & CHATTERJEE, D. 2006. The carboxy terminus of EmbC from *Mycobacterium smegmatis* mediates chain length extension of the arabinan in lipoarabinomannan. *J Biol Chem*, 281, 19512-26.

SHI, L., ZHOU, R., LIU, Z., LOWARY, T. L., SEEBERGER, P. H., STOCKER, B. L., CRICK, D. C., KHOO, K. H. & CHATTERJEE, D. 2008. Transfer of the first arabinofuranose residue to galactan is essential for *Mycobacterium smegmatis* viability. *J Bacteriol*, 190, 5248-55.

SILVE, G., VALERO-GUILLEN, P., QUEMARD, A., DUPONT, M. A., DAFPE, M. & LANEELLE, G. 1993. Ethambutol inhibition of glucose metabolism in mycobacteria: a possible target of the drug. *Antimicrob Agents Chemother*, 37, 1536-8.

SINGH, S. M. & PANDA, A. K. 2005. Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng*, 99, 303-10.

SKOVIEROVA, H., LARROUY-MAUMUS, G., PHAM, H., BELANOVA, M., BARILONE, N., DASGUPTA, A., MIKUSOVA, K., GICQUEL, B., GILLERON, M., BRENNAN, P. J., PUZO, G., NIGOU, J. & JACKSON, M. 2010. Biosynthetic origin of the galactosamine substituent of arabinogalactan in *Mycobacterium tuberculosis*. *J Biol Chem*, 285, 41348-55.

SKOVIEROVA, H., LARROUY-MAUMUS, G., ZHANG, J., KAUR, D., BARILONE, N., KORDULAKOVA, J., GILLERON, M., GUADAGNINI, S., BELANOVA, M., PREVOST, M. C., GICQUEL, B., PUZO, G., CHATTERJEE, D., BRENNAN, P. J.,

- NIGOU, J. & JACKSON, M. 2009. AftD, a novel essential arabinofuranosyltransferase from mycobacteria. *Glycobiology*, 19, 1235-47.
- SMITH, N. H., HEWINSON, R. G., KREMER, K., BROSCHE, R. & GORDON, S. V. 2009. Myths and misconceptions: the origin and evolution of *Mycobacterium tuberculosis*. *Nat Rev Microbiol*, 7, 537-44.
- SOTOMAYOR, H., BURGOS, J. & ARANGO, M. 2004. Demonstration of tuberculosis by DNA ribotyping of *Mycobacterium tuberculosis* in a Colombian prehispanic mummy. *Biomedica*, 24 Supp 1, 18-26.
- STAMM, L. M. & BROWN, E. J. 2004. *Mycobacterium marinum*: the generalization and specialization of a pathogenic mycobacterium. *Microbes Infect*, 6, 1418-28.
- STAPLETON, M. R., SMITH, L. J., HUNT, D. M., BUXTON, R. S. & GREEN, J. 2012. *Mycobacterium tuberculosis* WhiB1 represses transcription of the essential chaperonin GroEL2. *Tuberculosis (Edinb)*, 92, 328-32.
- STARKS, A. M., GUMUSBOGA, A., PLIKAYTIS, B. B., SHINNICK, T. M. & POSEY, J. E. 2009. Mutations at *embB* codon 306 are an important molecular indicator of ethambutol resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*, 53, 1061-6.
- STEHLE, T., SREERAMULU, S., LOHR, F., RICHTER, C., SAXENA, K., JONKER, H. R. & SCHWALBE, H. 2012. The apo-structure of the low molecular weight protein-tyrosine phosphatase A (MtpA) from *Mycobacterium tuberculosis* allows for better target-specific drug development. *J Biol Chem*, 287, 34569-82.
- STINEAR, T. P., SEEMANN, T., HARRISON, P. F., JENKIN, G. A., DAVIES, J. K., JOHNSON, P. D., ABDELLAH, Z., ARROWSMITH, C., CHILLINGWORTH, T., CHURCHER, C., CLARKE, K., CRONIN, A., DAVIS, P., GOODHEAD, I., HOLROYD, N., JAGELS, K., LORD, A., MOULE, S., MUNGALL, K., NORBERTCZAK, H., QUAIL, M. A., RABBINOWITSCH, E., WALKER, D., WHITE, B., WHITEHEAD, S., SMALL, P. L., BROSCHE, R., RAMAKRISHNAN, L., FISCHBACH, M. A., PARKHILL, J. & COLE, S. T. 2008. Insights from the complete genome sequence of *Mycobacterium marinum* on the evolution of *Mycobacterium tuberculosis*. *Genome Res*, 18, 729-41.
- SZKLARCZYK, D., FRANCESCHINI, A., KUHN, M., SIMONOVIC, M., ROTH, A., MINGUEZ, P., DOERKS, T., STARK, M., MULLER, J., BORK, P., JENSEN, L. J. & VON MERING, C. 2011. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res*, 39, D561-8.
- TAKAYAMA, K. & KILBURN, J. O. 1989. Inhibition of synthesis of arabinogalactan by ethambutol in *Mycobacterium smegmatis*. *Antimicrob Agents Chemother*, 33, 1493-9.
- TAKAYAMA, K., WANG, C. & BESRA, G. S. 2005. Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin Microbiol Rev*, 18, 81-101.

- TARBOURIECH, N., CHARNOCK, S. J. & DAVIES, G. J. 2001. Three-dimensional structures of the Mn and Mg dTDP complexes of the family GT-2 glycosyltransferase SpsA: a comparison with related NDP-sugar glycosyltransferases. *J Mol Biol*, 314, 655-61.
- TATITURI, R. V., ALDERWICK, L. J., MISHRA, A. K., NIGOU, J., GILLERON, M., KRUMBACH, K., HITCHEN, P., GIORDANO, A., MORRIS, H. R., DELL, A., EGGELING, L. & BESRA, G. S. 2007. Structural characterization of a partially arabinosylated lipoarabinomannan variant isolated from a *Corynebacterium glutamicum ubiA* mutant. *Microbiology*, 153, 2621-9.
- TATITURI, R. V., ILLARIONOV, P. A., DOVER, L. G., NIGOU, J., GILLERON, M., HITCHEN, P., KRUMBACH, K., MORRIS, H. R., SPENCER, N., DELL, A., EGGELING, L. & BESRA, G. S. 2007b. Inactivation of *Corynebacterium glutamicum* NCgl0452 and the role of MgtA in the biosynthesis of a novel mannosylated glycolipid involved in lipomannan biosynthesis. *J Biol Chem*, 282, 4561-72.
- TELENTI, A., PHILIPP, W. J., SREEVATSAN, S., BERNASCONI, C., STOCKBAUER, K. E., WIELES, B., MUSSER, J. M. & JACOBS, W. R., JR. 1997. The emb operon, a gene cluster of *Mycobacterium tuberculosis* involved in resistance to ethambutol. *Nat Med*, 3, 567-70.
- TORRELLES, J. B., DESJARDIN, L. E., MACNEIL, J., KAUFMAN, T. M., KUTZBACH, B., KNAUP, R., MCCARTHY, T. R., GURCHA, S. S., BESRA, G. S., CLEGG, S. & SCHLESINGER, L. S. 2009. Inactivation of *Mycobacterium tuberculosis* mannosyltransferase *pimB* reduces the cell wall lipoarabinomannan and lipomannan content and increases the rate of bacterial-induced human macrophage cell death. *Glycobiology*, 19, 743-55.
- UDWADIA, Z. & VENDOTI, D. 2013. Totally drug-resistant tuberculosis (TDR-TB) in India: every dark cloud has a silver lining. *J Epidemiol Community Health*, 67, 471-2.
- UDWADIA, Z. F., AMALE, R. A., AJBANI, K. K. & RODRIGUES, C. 2012. Totally drug-resistant tuberculosis in India. *Clin Infect Dis*, 54, 579-81.
- VAN HEIJENOORT, J. 2001a. Formation of the glycan chains in the synthesis of bacterial peptidoglycan. *Glycobiology*, 11, 25R-36R.
- VAN HEIJENOORT, J. 2001b. Recent advances in the formation of the bacterial peptidoglycan monomer unit. *Nat Prod Rep*, 18, 503-19.
- VAN HEIJENOORT, J. 2007. Lipid intermediates in the biosynthesis of bacterial peptidoglycan. *Microbiol Mol Biol Rev*, 71, 620-35.
- VAN SOOLINGEN, D., HOOGENBOEZEM, T., DE HAAS, P. E., HERMANS, P. W., KOEDAM, M. A., TEPPEMA, K. S., BRENNAN, P. J., BESRA, G. S., PORTAELS, F., TOP, J., SCHOULS, L. M. & VAN EMBDEN, J. D. 1997. A novel pathogenic taxon of the *Mycobacterium tuberculosis* complex, *Canetti*: characterization of an exceptional isolate from Africa. *Int J Syst Bacteriol*, 47, 1236-45.

- VARELA, C., RITTMANN, D., SINGH, A., KRUMBACH, K., BHATT, K., EGGELING, L., BESRA, G. S. & BHATT, A. 2012. *MmpL* genes are associated with mycolic acid metabolism in mycobacteria and corynebacteria. *Chem Biol*, 19, 498-506.
- VEDADI, M., NIESEN, F. H., ALLALI-HASSANI, A., FEDOROV, O. Y., FINERTY, P. J., JR., WASNEY, G. A., YEUNG, R., ARROWSMITH, C., BALL, L. J., BERGLUND, H., HUI, R., MARSDEN, B. D., NORDLUND, P., SUNDSTROM, M., WEIGELT, J. & EDWARDS, A. M. 2006. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A*, 103, 15835-40.
- VELAYATI, A. A., FARNIA, P. & MASJEDI, M. R. 2013. The totally drug resistant tuberculosis (TDR-TB). *Int J Clin Exp Med*, 6, 307-9.
- VELAYATI, A. A., MASJEDI, M. R., FARNIA, P., TABARSI, P., GHANAVI, J., ZIAZARIFI, A. H. & HOFFNER, S. E. 2009. Emergence of new forms of totally drug-resistant tuberculosis bacilli: super extensively drug-resistant tuberculosis or totally drug-resistant strains in iran. *Chest*, 136, 420-5.
- VELEZ-RUIZ, G. A. & SUNAHARA, R. K. 2011. Reconstitution of G protein-coupled receptors into a model bilayer system: reconstituted high-density lipoprotein particles. *Methods Mol Biol*, 756, 167-82.
- VERSCHOOR, J. A., BAIRD, M. S. & GROOTEN, J. 2012. Towards understanding the functional diversity of cell wall mycolic acids of *Mycobacterium tuberculosis*. *Prog Lipid Res*, 51, 325-39.
- VEYRON-CHURLET, R., BIGOT, S., GUERRINI, O., VERDOUX, S., MALAGA, W., DAFPE, M. & ZERBIB, D. 2005. The biosynthesis of mycolic acids in *Mycobacterium tuberculosis* relies on multiple specialized elongation complexes interconnected by specific protein-protein interactions. *J Mol Biol*, 353, 847-58.
- VEYRON-CHURLET, R., GUERRINI, O., MOUREY, L., DAFPE, M. & ZERBIB, D. 2004. Protein-protein interactions within the Fatty Acid Synthase-II system of *Mycobacterium tuberculosis* are essential for mycobacterial viability. *Mol Microbiol*, 54, 1161-72.
- VOLLMER, W. & HOLTJE, J. V. 2004. The architecture of the murein (peptidoglycan) in gram-negative bacteria: vertical scaffold or horizontal layer(s)? *J Bacteriol*, 186, 5978-87.
- WADENPOHL, I. & BRAMKAMP, M. 2010. DivIC stabilizes FtsL against RasP cleavage. *J Bacteriol*, 192, 5260-3.
- WANG, Y., CUI, T., ZHANG, C., YANG, M., HUANG, Y., LI, W., ZHANG, L., GAO, C., HE, Y., LI, Y., HUANG, F., ZENG, J., HUANG, C., YANG, Q., TIAN, Y., ZHAO, C., CHEN, H., ZHANG, H. & HE, Z. G. 2010. Global protein-protein interaction network in the human pathogen *Mycobacterium tuberculosis* H37Rv. *J Proteome Res*, 9, 6665-77.

- WESTON, A., STERN, R. J., LEE, R. E., NASSAU, P. M., MONSEY, D., MARTIN, S. L., SCHERMAN, M. S., BESRA, G. S., DUNCAN, K. & MCNEIL, M. R. 1997. Biosynthetic origin of mycobacterial cell wall galactofuranosyl residues. *Tuber Lung Dis*, 78, 123-31.
- WHEATLEY, R. W., ZHENG, R. B., RICHARDS, M. R., LOWARY, T. L. & NG, K. K. 2012. Tetrameric structure of the GlfT2 galactofuranosyltransferase reveals a scaffold for the assembly of mycobacterial Arabinogalactan. *J Biol Chem*, 287, 28132-43.
- WHO 2004. Procurement manual for the DOTS-plus projects approved by the Green Light Committee World Health Organization.
- WHO 2012. WHO report 2012: Global tuberculosis control. World Health Organization.
- WIETZERBIN, J., DAS, B. C., PETIT, J. F., LEDERER, E., LEYH-BOUILLE, M. & GHUYSEN, J. M. 1974. Occurrence of D-alanyl-(D)-meso-diaminopimelic acid and meso-diaminopimelyl-meso-diaminopimelic acid interpeptide linkages in the peptidoglycan of *Mycobacteria*. *Biochemistry*, 13, 3471-6.
- WOLUCKA, B. A. 2008. Biosynthesis of D-arabinose in mycobacteria - a novel bacterial pathway with implications for antimycobacterial therapy. *FEBS J*, 275, 2691-711.
- WOLUCKA, B. A., MCNEIL, M. R., DE HOFFMANN, E., CHOJNACKI, T. & BRENNAN, P. J. 1994. Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on mycobacteria. *J Biol Chem*, 269, 23328-35.
- WU, Q., ZHOU, P., QIAN, S., QIN, X., FAN, Z., FU, Q., ZHAN, Z. & PEI, H. 2011. Cloning, expression, identification and bioinformatics analysis of *Rv3265c* gene from *Mycobacterium tuberculosis* in *Escherichia coli*. *Asian Pac J Trop Med*, 4, 266-70.
- XU, L., WU, D., LIU, L., ZHENG, Q., SONG, Y., YE, L., SHA, S., KANG, J., XIN, Y. & MA, Y. 2014. Characterization of mycobacterial UDP-N-acetylglucosamine enolpyruvyle transferase (MurA). *Res Microbiol*, 165, 91-101.
- YAGI, T., MAHAPATRA, S., MIKUSOVA, K., CRICK, D. C. & BRENNAN, P. J. 2003. Polymerization of mycobacterial arabinogalactan and ligation to peptidoglycan. *J Biol Chem*, 278, 26497-504.
- YANG, Z., ZHANG, L., ZHANG, Y., ZHANG, T., FENG, Y., LU, X., LAN, W., WANG, J., WU, H., CAO, C. & WANG, X. 2011. Highly efficient production of soluble proteins from insoluble inclusion bodies by a two-step-denaturing and refolding method. *PLoS One*, 6, e22981.
- YEW, W. W. & LEUNG, C. C. 2006. Update in tuberculosis 2005. *Am J Respir Crit Care Med*, 173, 491-8.

- YUAN, Y., LEE, R. E., BESRA, G. S., BELISLE, J. T. & BARRY, C. E., 3RD 1995. Identification of a gene involved in the biosynthesis of cyclopropanated mycolic acids in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A*, 92, 6630-4.
- ZHANG, J., KHOO, K. H., WU, S. W. & CHATTERJEE, D. 2007. Characterization of a distinct arabinofuranosyltransferase in *Mycobacterium smegmatis*. *J Am Chem Soc*, 129, 9650-62.
- ZHANG, N., TORRELLES, J. B., MCNEIL, M. R., ESCUYER, V. E., KHOO, K. H., BRENNAN, P. J. & CHATTERJEE, D. 2003. The Emb proteins of mycobacteria direct arabinosylation of lipoarabinomannan and arabinogalactan via an N-terminal recognition region and a C-terminal synthetic region. *Mol Microbiol*, 50, 69-76.
- ZHANG, W., JONES, V. C., SCHERMAN, M. S., MAHAPATRA, S., CRICK, D., BHAMIDI, S., XIN, Y., MCNEIL, M. R. & MA, Y. 2008. Expression, essentiality, and a microtiter plate assay for mycobacterial GlmU, the bifunctional glucosamine-1-phosphate acetyltransferase and *N*-acetylglucosamine-1-phosphate uridylyltransferase. *Int J Biochem Cell Biol*, 40, 2560-71.
- ZHANG, Y. 2005. The magic bullets and tuberculosis drug targets. *Annu Rev Pharmacol Toxicol*, 45, 529-64.
- ZHANG, Z., BULLOCH, E. M., BUNKER, R. D., BAKER, E. N. & SQUIRE, C. J. 2009. Structure and function of GlmU from *Mycobacterium tuberculosis*. *Acta Crystallogr D Biol Crystallogr*, 65, 275-83.
- ZHENG, J., WEI, C., ZHAO, L., LIU, L., LENG, W., LI, W. & JIN, Q. 2011. Combining blue native polyacrylamide gel electrophoresis with liquid chromatography tandem mass spectrometry as an effective strategy for analyzing potential membrane protein complexes of *Mycobacterium bovis* bacillus Calmette-Guerin. *BMC Genomics*, 12, 40.
- ZHOU, Y., XIN, Y., SHA, S. & MA, Y. 2011. Kinetic properties of *Mycobacterium tuberculosis* bifunctional GlmU. *Arch Microbiol*, 193, 751-7.
- ZHOU, Y., YU, W., ZHENG, Q., XIN, Y. & MA, Y. 2012. Identification of amino acids involved in catalytic process of *M. tuberculosis* GlmU acetyltransferase. *Glycoconj J*, 29, 297-303.
- ZINK, A. R., GRABNER, W., REISCHL, U., WOLF, H. & NERLICH, A. G. 2003. Molecular study on human tuberculosis in three geographically distinct and time delineated populations from ancient Egypt. *Epidemiol Infect*, 130, 239-49.

