Incorporation of Trehalose Analogues into *Mycobacterium Tuberculosis*: Antigen 85 and Probes of Bacterial Infection

Thesis submitted in partial fulfilment for the degree of

Doctor of Philosophy at the University of Oxford



By

Keriann Marie Backus

New College, Oxford

Trinity Term 2011

Incorporation of Trehalose Analogues into *Mycobacterium Tuberculosis*: Antigen 85 and Probes of Bacterial Infection

By Keriann M. Backus, New College

Submitted towards the degree of D.Phil in Trinity Term 2011

Diagnoses of tuberculosis, 'TB,' currently rely upon non-specific techniques such as X-ray exams and acid-fast microscopy. Improved diagnostics would preferably consider specific bacterial processes to provide real-time readouts of disease burden and response to chemotherapy. This dissertation presents the cell-wall incorporation of trehalose analogues (fluorescent and radioactive) by the mycobacterial antigen 85 enzymes as a novel method to label the causative bacteria of TB, *Mycobacterium tuberculosis (Mtb).*

The trehalose mycolyltransesterase enzymes (antigens 85A, B, and C (Ag85)) serve as essential mediators of cell envelope function and biogenesis in *Mtb.* We show that the Ag85 enzymes display activities so broad that they allow added non-natural carbohydrate probes to be incorporated into *Mtb* growing *in vitro* and within macrophages. Design and synthesis of a library of structurally-diverse analogs of the sugar trehalose (Tre) revealed that Ag85-enzymes catalyze esterification of a wide variety of non-natural Tre structures, even stereoisomers and those appended with charged or bulky groups (Chapter 2). A novel mass-spectrometry based Ag85 enzyme assay was developed and employed to screen the library of compounds against all three isoforms of Ag85 (Chapter 3). This screen revealed that the Ag85 enzymes exhibit preference for dissacharides over monosaccharides and show a broad tolerance for most modified trehalose compounds. This activity assay also afforded full kinetic analysis and the discovery of a novel, covalent inhibitor of the Ag85 enzymes.

The Ag85 activity assay informed the design of a fluorescent trehalose-based compound (FITC-Tre), which is the first, non-toxic, selective, small molecule probe for mycobacterial infection. FITC-Tre was acylated with mycolyl esters by growing mycobacteria, anchoring the probe in the cell envelope resulting in fluorescent bacteria (Chapter 4). Adding FITC-Tre to *Mtb*-infected macrophages allowed selective, fluorescent tagging of *Mtb in vivo* (Chapter 5). Colocalization studies with antibodies against a variety of phagosomal associated components have hinted at the possibility of FITC-Tre as readout of cellular trafficking of bacteria. ^{18}F-trehalose, biotin-trehalose and rhodamine-trehalose are also substrates of Ag85. ^{18}F-trehalose shows promise as *Mtb* selective PET probe in an infected rabbit model of

tuberculosis. Future work with these probes may allow for fluorescent tracking of the *Mtb* during the macrophage infection process, as well as the ability to label Mtb in infected tissue.

The functional differences between the three isoforms of Ag85, A, B and C, are not well understood and may have implications for the survival and persistence of mycobacteria within humans. The differences in substrate specificity and catalytic activity between the Ag85 isoforms (discussed in Chapter 3) has been further investigated (Chapter 6). Mutation of three secondary site amino acids from Ag85C into Ag85B afforded nearly a twenty-fold gain in enzyme activity. Mutation of the equivalent Ag85B residues into Ag85C triggered nearly a twenty-fold loss in activity. Dissection of the roles of these three amino acids helps to explain the previously reported large differences in catalytic activity between Ag85A, B and C. Overexpression of Ag85A, B and C under tetracycline regulation revealed that these enzymes differentially modulate incorporation of mycolates into the cell wall. The Ag85 enzymes are not functionally redundant, and instead serve unique purposes in cell wall biosynthesis.

In summary, this research has demonstrated that the broad substrate tolerance of Ag85 enzymes, coupled with their extracellular location, opens the door to probes of mycobacterial infection using many imaging modalities.

Acknowledgements

This research was made possible by the support, guidance and freedom bestowed by my supervisors, Dr Clifton Barry and Professor Benjamin Davis. This thesis is the product of an interdisciplinary, joint research venture between the Barry and Davis labs. As such, I have had the opportunity to spend two years researching in the Oxford Chemistry Research Laboratory and a subsequent two years at the NIH Tuberculosis Research Section and have been fortunate to work with outstanding scientists at both institutions. I thank Professor Davis and Dr. Barry for supporting me in this combined doctoral program.

I could not have completed this D.Phil without the generous support of the Rhodes Trust (2007-2009) and the NIH Oxford Cambridge Scholars program (2009-2011). This funding allowed me to pursue a challenging body of research and share my findings at several international conferences. I thank both funding bodies for investing in me and in this research project. As part of the NIH Oxford Cambridge program, I would also like to thank Dr. Michael Lenardo and Dr. John Hanover for their mentorship, guidance and support throughout my D.Phil.

Much of the research presented here would not have been possible without extensive collaborations with other members of the Davis and Barry groups and other labs at Oxford and the NIH. In addition to recognizing them here, throughout this thesis I will also acknowledge the individual(s) who assisted with portions of this project. This research was also greatly facilitated by the excellent support staff of the Oxford Chemistry Research Lab and core facility staff at the NIH. In particular, I would like to thank the Biological Imaging Facility, particularly, Dr. Owen Schwartz, Dr. Lily Koo, Dr. Juraj Kabat and Dr. Stephen Becker, for being an invaluable resource for the art of confocal microscopy. Dr. Michael Dolan also deserves acknowledgement for his assistance with the molecular dynamics simulations.

My coworkers in Oxford and at the NIH provided invaluable scientific assistance and, more importantly, are wonderful, caring people and great friends and colleagues. In particular I would like to thank, ‘team trehalose,’ Dr. Omar Boutureira, Dr. Seung Lee and Dr. François D’Hooge and particularly Dr. Conor Barry for all their work compiling the trehalose project into a cohesive, exciting story. Mitul Patel, as my fellow 2011 D.Phil and trehalose partner in crime, deserves mention for his synthetic assistance and refreshingly sardonic outlook on life.

I would also like to thank Dr. Alister French, Dr. Joanna McGouran, Dr. Justin Chalker, Dr. James Errey, Dr. Julia Koeppe and Dr. Filip Wyszynski for their guidance and support and companionship, particularly during the beginning of my D.Phil.

At the NIH I would like to thank my coworkers and colleagues for their support throughout my time at the Tuberculosis Research Section. Dr. Tathagata Mukherjee, Dr. Amit Nayyar, Dr. Kriti Arora and Dr. Todd Gruber have all been very supportive both in scientific and life choices and deserve acknowledgment for their help proofreading this thesis. The animal experiments detailed here were carried out with the excellent assistance of Jacquie Gonzales, Dan Schimel and Danielle Weiner. Mandie Samuels and Ben Winterroth were tremendous cubicle and coffee companions, who inspired me to return to running and triathlons. Dale Kieswetter patiently taught me some of the nuisances of radiochemistry and guided me through the synthesis of ^{18}F-trehalose. Dr. Peter McPhie merits recognition for his help with circular dichroism experiments. Dr. Helena Boshoff has been supportive throughout my D.Phil, both in the design and execution of experiments and in the negotiation of a trans-Atlantic thesis project.

I am grateful for the love and support from my friends and family. I would particularly like to thank my parents, Ned and Carol Backus, for instilling me with a spirit of adventure, sense of curiosity and a resilience that has provided the foundation for my research career. Without their love and support and that of my wonderful boyfriend, Will Pisarello, and my brother, Spencer Backus, this thesis would not have been possible.

List of Publications

Backus, K.M.. Boshoff, H. I., Barry, C. S., Boutureira, O., Patel, M. K., D'Hooge, F., Lee, S. S., Via, L. E., Tahlan, K., Barry, C. E., 3rd, Davis, B. G. Uptake of unnatural trehalose analogs as a reporter for *Mycobacterium tuberculosis*. *Nat Chem Biol* **7**, 228-35 (2011).

Barry, C.S., Backus, K.M., Barry, C.E. & Davis, B.G. ESI-MS Assay of *M. tuberculosis* Cell Wall Antigen 85 Enzymes Permits Substrate Profiling and Design of a Mechanism-Based Inhibitor. *J. Am. Chem. Soc.*, **133**, 13232-13235 (2011).

Table of Contents

Abstract ..... **i**
Acknowledgments ..... **iii**
List of Publications ..... **v**
Table of Contents ..... **vi**
Statement of Authorship ..... **xii**
Abbreviations ..... **xiii**
Explanatory Note ..... **xvi**

Chapter 1: Introduction 1

1.1 Tuberculosis epidemic: the need for novel diagnostics **1**
1.2 Model of the mycobacterial cell wall **2**
1.3 Mycolic acids **4**
1.4 Trehalose mono and dimycolate (TMM and TDM) **5**
1.4.1 Structural features of TDM 6
1.4.2 Prior syntheses of TDM and TMM 7
1.4.3 Biophysical studies of TDM properties 8
1.4.4 Biosynthesis of TDM 9
1.5 Antigen 85 Enzymes **11**
1.5.1 Ag85 partial redundancy 13
1.5.2 Ag85 crystal structures 14
1.5.3 Ag85 activity assays 15
1.5.4 Trehalose as a drug target 17
1.6 Immunopathology and other functions of TDM and Ag85 ..... **18**
1.6.1 TDM and granuloma formation 19
1.6.2 Immunopathology and biological properties of Ag85 22
1.6.3 Ag85, TDM and a TB vaccine 22
1.7 Selection of synthetic target **23**
1.7.1 Synthesis of trehalose analogues 24
1.8 Uptake of trehalose probes into *Mtb in vitro* and *in vivo* **25**
1.8.1 Small molecule labeling of mycobacteria 26
1.8.2 Uptake of trehalose into animals 26
1.9 Positron emission tomography and computed tomography (PET/CT) **27**
1.10 Conclusions **28**
1.11 Thesis Objectives **29**
1.12 References **29**

Chapter 2: Synthesis of trehalose probes 36

2.1 Introduction **36**
2.2 Results **36**
2.2.1 Retrosynthetic analysis 36
2.2.2 Open active site of Antigen 85 (Ag85) suggests tolerance to modifications 37
2.2.3 Overview of the trehalose library 38
2.2.4 Synthetic strategies 39
2.2.5 Enzymatic synthesis 40
2.2.6 Radical bromination 42

2.2.7 Ketoside trehalose synthesis 44
2.2.7.1 Retrosynthetic analysis 44
2.2.7.2 Syntheses of donors and acceptors 45
2.2.7.3 Glycosylation of ketosides 46
2.2.7.4 Glycosylation of exoglycals 47
2.2.8 Intramolecular aglycone delivery (IAD) 51
2.2.9 Mannose directed glycosylations 53
2.2.10 Methyl ketoside glycosylations with acetylated acceptors 55
2.2.11 Methyl ketosides: Reaction scope and library synthesis 58
2.2.11.1 C18-trehalose synthesis 63
2.2.11.2 NOE assignment of stereochemistry 65
2.2.11.3 Isotopic ^{13}C and ^{14}C labeling of trehalose 66
2.2.12 Alternative uses for isotopically labeled trehalose: MRI 71
2.2.13 Modification of 2-aminotrehalose analogue 37 72
2.2.13.1 Selective removal of the Cbz protecting group 79
2.2.13.2 Peptide coupling 80
2.2.13.3 Reductive amination 81
2.2.13.4 Trehalose coated quantum dots 55 83
2.2.13.5 Synthesis of DOTA-trehalose analogue 57 85
2.2.13.6 Synthesis of biotin-trehalose analogue 54 86
2.2.13.7 Synthesis of rhodamine-trehalose analogue 56 87
2.3 Conclusions 89
2.4 References 90

Chapter 3: Ag85 enzyme assays 92

3.1 Introduction 92
3.2 Results 93
3.2.1 Antigen 85 Expression. 93
3.2.2 Radiometric assays 95
3.2.3 Isothermal titration calorimetry (ITC) 95
3.2.4 Coupled assays 96
3.2.5 Design of the Ag85 enzyme assay 97
3.2.5.1 Synthesis of trehalose mono and dihexanoate (TMH, 212, TDH, 215) 98
3.2.5.2 Mass Spectrometry' 101
3.2.6 Green Amber Red (GAR) screen 106
3.3 Resistance of modified trehalose to trehalase degradation 111
3.4 Design of an Ag85 inhibitor 114
3.4.1 The inhibitory properties of trehalose-6-azide 48 114
3.4.2 Protein modification- covalent inhibitors 116
3.4.3 Alternative inhibitor strategies 120
3.4.3.1 Synthesis of Ag85 inhibitor trehalose fluorophosphonate 60 121
3.5 Reaction of Ag85 with potential inhibitors 123
3.6 Conclusions 127
3.7 References 127

Chapter 4: Labeling mycobacteria *in vitro* 130

4.1 Introduction 130
4.2 Results 130
4.2.1 Trehalose uptake by whole cells of *Mtb* 130
4.2.2 ^{14}C-trehalose labels TDM and TMM of mycobacteria 132
4.2.3 FITC-trehalose specifically labels *M. tuberculosis* in batch culture 134
4.2.4 Polar labeling of Mtb microscopy 136
4.2.5 MIC data 138
4.3 Conclusions 139
4.4 References 140

Chapter 5: Labeling of mycobacteria *in vivo* 141

5.1 Introduction 141
5.2 Results 141
5.2.1 Labeling of *Mtb* in macrophages by ^{14}C-Tre 141
5.2.2 Labeling of *Mtb* in macrophages by FITC- Tre 142
5.2.3 Establishing the non-toxic, selective nature of the FITC-Tre 146
5.2.4 Determining the selectivity of the labeling 147
5.2.5 Differential labeling by FITC-Tre of *Mtb* macrophages 149
5.2.6 Statistical analysis of BCG vs. *Mtb* fluorescence 150
5.2.7 Trafficking studies: Labeling of endosomes and lysosomes 151
5.2.8 *Mtb* subpopulation that does not label with FITC-Tre (red-but-not-green) 153
5.2.9 FITC-Tre labeling may be linked to mycobacterial trafficking 153
5.2.10 Bone marrow macrophage labeling and colocalization studies 158
5.2.10.1 Colocalization results 160
5.2.11 IFN-γ vs. Non-INF-γ treated cells 165
5.3 Ag85 antibodies show non-specific labeling 168
5.3.1 Attempts at *in vivo* labeling 169
5.3.2 Attempts at tissue labeling 170
5.3.3 ^{18}F-Trehalose and PET imaging of infected rabbits 172
5.3.4 Pharmacokinetics and pharmacodynamics (PK/PD) 174
5.4 Conclusions 175
5.5 References 176

Chapter 6: Differential roles of Ag85s and related enzymes 178

6.1 Introduction 178
6.2 Results 179
6.2.1 Identification and analysis of key residues in Ag85 secondary site loop 179
6.2.2 Molecular Dynamics 183
6.2.3 Mutagenesis of Ag85C 186
6.2.3.1 Dimerization state of mutants 191
6.2.4 Overexpression of Rv1288 and Rv0519c 193
6.2.5 Transcriptional profiling of Ag85 196
6.2.6 Overexpression of His-tagged Ag85 constructs in *M. smegmatis* and *Mtb* 199
6.2.7 Overexpression of Ag85 as Tet inducible GFP fusion proteins 205
6.2.8 Microscopy of Tet inducible plasmids 210

6.2.9 Overexpression of Tet inducible constructs in *Mtb* 213
6.2.9.1 MIC analysis of overexpression strains 217
6.2.9.2 TDM/TMM profiling of Tet inducible GFP fusion proteins in H37RV 218
6.3 Conclusions 222
6.4 References 222

Chapter 7: Conclusions and further considerations 225

7.1. References 229

Appendix 1: Chemistry synthesis and characterization 231

8.1 General Experimental 231
8.2 Synthetic Experimental Details 233
8.2.1 Trehalose-6 phosphate synthase (OtsA) synthesis of compounds 11 and 51. 233
8.2.2 Radical bromination: attempted synthesis of compounds 73, 74 and 76. 238
8.2.3 Ketoside trehalose synthesis 240
8.2.3.1 Synthesis of methyl ketosides 96, 97 and 103 240
8.2.4 Intramolecular aglycone delivery glycosylation 247
8.2.5 Mannose glycosylations: Synthesis of compounds 126-129 251
8.2.6 Synthesis of trehalose methyl ketosides 253
8.2.6.1 Synthesis of compound 32. 254
8.2.6.2 Synthesis of compound 83 257
8.2.6.3 Synthesis of compound 33 260
8.2.6.4 Synthesis of compound 34. 263
8.2.6.5 Synthesis of compound 35. 265
8.2.6.6 Synthesis of compound 36. 268
8.2.6.7 Synthesis of compound 145 273
8.2.6.8 Synthesis of compounds 146 and 147 275
8.2.7 Isotopic labeling of trehalose: Synthesis of compounds 148 and 149 279
8.2.7.1 Attempted synthesis of ^{14}C-trehalose 149 281
8.2.8 Synthesis of amino-trehalose 37 285
8.2.8.1 Alternate synthesis of compound 138 288
8.2.8.2 Deprotection of compound compound 150 292
8.2.9 Other selective modifications of 37 295
8.2.9.1 Synthesis of ^{18}F-Trehalose 39 298
8.2.10 Synthesis of Quantum dot-trehalose, 55 301
8.2.11 Synthesis of DOTA-trehalose 57 303
8.2.12 Synthesis of biotin-trehalose, 54 305
8.2.13 Synthesis of rhodamine-trehalose, 56 and 201 307
8.2.14 Synthesis of trehalose-6-Azide 48 310
8.2.15 Synthesis of nitrophenol esters, 203-206 313
8.2.16 Synthesis of compounds TMH and TDH 212 and 215 315
8.2.16.1 Alternate synthesis of compounds 210 and 4 323
8.2.17 Synthesis of arabinofuranosides 221 and 222 326
8.2.18 Synthesis of compound 60 327
8.2.19 Synthesis of octanesulfonyl-trehalose 61 332
8.2.20 Synthesis of FITC-glucose 254 334
8.2.21 Protein modifications 336
8.2.21.1 Reaction of Ag85C with 60 336

8.2.21.2 Reaction of Subtilisin Bacillus lentus (SBL) with 60 337
8.2.21.3 Reaction of Arctica candida lipase with 60 338
8.2.21.4 Reaction of Ag85C with PMSF 230 339
8.2.21.5 Reaction of Ag85-124PMS with Selenide 231. 340
8.3 References 341

Appendix 2: Biological materials and methods 344

9.1 General Experimental 344
9.2 Protein Expression 345
9.2.1 Ag85 general protein expression methods 345
9.2.2 Expression of Ag85A 346
9.2.3 Expression of Ag85B 349
9.2.4 Expression of Ag85C 351
9.2.5 Expression of Trehalose-6-phosphate synthase 353
9.3 Kinetic Analysis 354
9.3.1 Ag85 ESI MS assays 356
9.4 Minimum Inhibitory Concentration (MIC) Determination 358
9.5 Trehalase assay 358
9.6 Mycobacterial cell culture 359
9.6.1 Middlebrook 7H9 359
9.6.2 GAST 359
9.7 ^{14}C-Trehalose, ^{14}C-glycerol and ^{14}C-glucose uptake into H37Rv 359
9.7.1 ^{14}C- and FITC-Tre lipid extractions 360
9.7.2 Isoniazid (InH) inhibition of mycolate production 360
9.7.3 Mycolate analysis 360
9.8 FITC-Tre and FITC-Glc uptake into *Mtb* 361
9.9 FITC-Tre uptake into other bacteria 361
9.10 Trehalose inhibition of labeling *in vitro* 362
9.11 Polar labeling *in vitro* 362
9.12 Test of growth inhibition due to FITC-Tre in macrophages 362
9.13 *Mtb* antibody labeling 362
9.14 Tre-Quantum dots labeling of *Mtb* infected macrophages 363
9.15 FITC-Glc labeling of Mtb infected macrophages 363
9.16 Trehalose inhibition of FITC-Tre labeling in vivo 363
9.17 Primary bone marrow macrophages (BMM) isolation[6] 363
9.18 Immunocytochemistry (ICC) procedure 364
9.19 Image collection and processing 365
9.20 Correction for 63X blue offset 365
9.21 Data compilation 366
9.22 *In vivo* labeling experiments in rabbits and mice 369
9.22.1 Animals and ethics assurance 369
9.22.2 Labeling of mouse and rabbit tissue 369
9.22.3 *In vivo* labeling with FITC-Tre in infected mice 369
9.22.4 *In vitro* labeling: FITC-Tre labeling of tissue slices with and without Ag85. . 370
9.22.5 Auramine stain 370
9.22.6 PET experiment with ^{18}F-fluorobenzyl trehalose 370
9.22.6.1 Anesthesia and parameters for PET/CT scanning 370
9.22.7 Pharmacokinetic Studies in Uninfected Animals 372
9.23 Molecular dynamics 372

9.24 Phylogenetic Analysis 373
9.25 Site Directed Mutagenesis 374
9.25.1 Mutagenesis of Ag85C 374
9.25.2 Expression of Ag85C mutant proteins 375
9.25.3 Circular Dichroism 376
9.25.4 Mutagenesis of Ag85B 377
9.25.5 Overexpression of Ag85B mutants 377
9.25.6 Native Page Gels 378
9.26 General Cloning: PCR and ligation and transformation 378
9.26.1 Restriction Digests 379
9.26.2 Ligation Reactions 379
9.27 Cloning and expression of Rv0519C and Rv1288 379
9.27.1 Alignment of RV1288 and Rv0519C with Ag85 proteins 381
9.27.2 Cloning of Rv0519C and Rv1288 382
9.28 Microarray analysis 385
9.29 Cloning of Ag85 into pVV16 His-tag vector 386
9.30 Western blots 387
9.31 Mycolyl transferase genes into tetracycline inducible GFP fusions 388
9.31.1 Preparation of electrocompetent *M. smegmatis* and *M. tuberculosis* 390
9.31.2 Electroporation into *M. smegmatis* 390
9.31.3 Electroporation into Mtb 391
9.31.4 Induction of strains overexpressing Tet inducible GFP fusion proteins 391
9.31.5 Minimum Inhibitory Concentration (MIC) with Tet inducible strains 391
9.31.6 Lipid extractions of ATc inducible constructs 394
9.31.6.1 TDM and TMM extraction 394
9.31.6.2 Mycolate analysis 394
9.32 Summary of all plasmids 395
9.32.1 Selected plasmid maps 397
9.33 References 410

Authorship Declaration

I declare that this thesis represents my own work, except where stated below.

It has never been submitted for any other degrees.

Compounds **41**-**45** were synthesized by Dr. Omar Boutureira in the laboratory of Prof. Benjamin Davis.

Compounds **49**-**51**, **53**, **58**, **59**, **254** were synthesized by Dr. Francois D'Hooge in the laboratory of Prof. Benjamin Davis.

Compounds **46**-**48** were synthesized by Mitul K Patel in the laboratory of Prof. Benjamin Davis.

Compounds **52** and **71** were synthesized by Dr. Seung S. Lee in the laboratory of Prof. Benjamin Davis.

Molecular Dynamics Simulations were conducted by Dr. Michael Dolan in the National institute of bioinformatics (NIBIB, NIAID).

Ag85 assays were conducted by Dr. Conor Barry in the laboratory of Prof. Benjamin Davis.

Trehalose-6-azide Ag85 inhibition assay was conducted by Dr. Kapil Tahlan in the laboratory of Dr. Clifton Barry.

The pTE-2M1L plasmid construct was obtained from Dr. Sabine Ehrt.

The RFP BCG strain was obtained from the laboratory of Dr. Alan Sher.

The RFP TB strain was made by Dr. Taegwon Oh in the laboratory of Dr. Clifton Barry.

RNA Microarrays were conducted by Dr. Kriti Arora and Dr. Helena Boshoff

Sections of this thesis were reproduced in part with permission from: Barry, C.S., Backus, K.M., Barry, C.E. & Davis, B.G. ESI-MS Assay of M. tuberculosis Cell Wall Antigen 85 Enzymes Permits Substrate Profiling and Design of a Mechanism-Based Inhibitor. *J. Am. Chem. Soc*., **133**, 13232-13235 (2011). Copyright 2011, American Chemical Society.

Sections of this thesis were reproduced in part with permission from: Backus, K.M. et al. Uptake of unnatural trehalose analogs as a reporter for Mycobacterium tuberculosis. *Nat. Chem. Biol.* **7**, 228-35 (2011). Copyright 2011, Nature Chemical Biology.

Common Abbreviations

Ac	Acetyl
Ag85	Antigen 85
Ala	A or Alanine
AMP	ampicillin,
aq.	Aqueous
Ar	Aryl
at	Apparent triplet
ATc	Anhydrotetracycline
ATP	Adensosine triphosphate
BCG	Bacille Calmette-Guérin
Bn	Benzyl
br	Broad
BTFFH	Bis(tetramethylene)fluoroformamidinium hexafluorophosphate
Bz	Benzoyl
C	Celsius
Cam	chloramphenicol
Cbz	Carboxybenzyl
conc.	Concentrated
COSY	Correlation spectroscopy
CV	Column volume
d	Doublet
dd	Doublet of doublets
dt	Doublet of triplets
DAPI	4',6-diamidino-2-phenylindole
DAST	Diethylaminosulfur trifluoride
DEPT	Distortionless enhancement by polarisation transfer
DCM	CH_2Cl_2 or dichloromethane
DIC	Differential intereference contrast image
DIPEA	N,N-diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMB	3,4-Dimethoxybenzyl
DMSO	Dimethylsulfoxide
DMF	Dimethylformamide
EDTA	Ethylenediaminetetraacetic acid
Et	Ethyl
Ether	Et_2O = diethyl ether
Et_3N	Triethylamine
EtOH	Ethanol
EtOAc	Ethyl acetate
ESI MS	Electrospray ionization mass spectrometry
eq.	Equiv. or equivalents
FPLC	Fast protein liquid chromatography
FITC	Fluorescein isothiocyanate
Gal	Galactose
GAR	Green Amber Red
GFP	Green fluorescent protein
g	Grams
Glc	Glucose

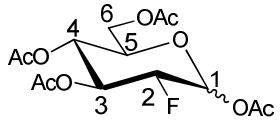
Gly	G or Glycine
h	Hours
H_2O	Water
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His-tag	Hexa histidine-tag
HMBC	Heteronuclear multiple bond correlation
HMQC	Heteronuclear multiple quantum coupling
HPLC	High-performace liquid chromatography
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum correlation
Hyg	Hygromycine
Hz	Hertz
INH	Isoniazid
*i*PrOH	Isopropanol
IPTG	Isopropyl-β-D-thiogalactopyranoside
IR	Infra-red
J	Coupling constant
k_{cat}	Overall catalytic rate of an enzyme
K_d	Enzyme dissociation constant
K_i	Absolute inhibition constant
K_M	Substrate concentration at which enzyme achieves 1/2 V_{max}
K_2CO_3	Potassium carbonate
Kan	Kanamycin
L	Litres
LB	Luria broth
LCMS	Liquid chromatography-mass spectrometry,
Lit.	Literature reference
M	Molar
m	Multiplet or milli
µ	Micro
M. tuberculosis	*Mycobacterium tuberculosis*
Me	methyl
MeCN	Acetonitrile
MeOH	Methanol
Met	Methionine
MIC	Minimum inhibitory concentration
min	Minutes
Mol	Moles
m.p.	melting point,
MS	Mass spectrometry
Mtb	*Mycobacterium tuberculosis*
m/z	Mass to charge ratio
n	Nano
$NaHCO_3$	Sodium bicarbonate
NaOH	Sodium hydroxide
NBS	*N*-bromosuccinamide
NH_4Cl	Ammonium chloride
NH_4OH	Ammonium hydroxide
NIS	*N*-iodosuccinamide
nm	Nanometers

NMR	Nuclear magnetic resonance
p	*para*
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline or Tuberculosis
PCR	Polymerase chain reaction
Ph	Phenyl
PMB	*para*-methoxybenzyl
PMSF	Phenylmethanesulfonylfluoride
Py	Pyridine
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
QD	Quantum dot
RBS	Ribosome Binding Site
R_f	Retention factor
Satd.	Saturated
SPE	Solid phase extraction
RFP	Red fluorescent protein
ROI	Region of interest
rpm	Revolutions per minute
RT	Room temperature
s	Singlet, seconds, strong
S/N	Signal to noise
SDS	Sodium dodecyl sulfate
T	Triplet
TB	*Tubercle bacillus* or Tuberculosis
TBAF	Tetra-*n*-butylammonium fluoride,
TBDPS	*tert*-butyldiphenylsilyl
TEA	Triethanolamine
TFA	Trifluoroacetic acid
Tet	Tetracycline
THF	Tetrahydrofuran,
TLC	Thin layer chromatography
TMSOTf	Trimethylsilyl trifluoromethanesulfonate
TFE	Trifluoro ethanol
TOF	Time of flight
Tre	Trehalose
Trp	W or Tryptophan
UDP	Uridine diphosphate
V_{max}	Maximal enzyme velocity

Explanatory Note

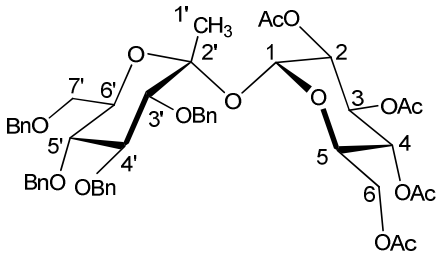
Carbohydrates and their derivatives have been named in accordance with the IUPAC recommendations and numbered according to the carbohydrate convention.

1,3,4,6-Tetra-*O*-acetyl-2-deoxy-2-fluoro-D-glucopyranose



Ketoside trehalose analogues are numbered in Appendix 1 in the following manner. This numbering follows the precedent of analogue number set by IUPAC nomenclature for ketosides. In the main text, for clarity, compounds are are referred to as methyl-trehalose and are numbered according to the convention for unmodified trehalose.

3,4,5,7-Tetra-*O*-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2,3,4,6-tetra-*O*-acetyl-2-α-D-glucopyranoside



1 Chapter 1: Introduction

1.1 Tuberculosis epidemic: the need for novel diagnostics

Tuberculosis (TB) is an infection that has plagued mankind for millennia. It remains a leading cause of death worldwide with an estimated 1.6 million fatalities in 2005[1]. A substantial obstacle to the development of new tuberculosis diagnostics, drugs, and vaccines is the lack of TB-specific probes that can be used to rapidly assess infection and monitor response to treatment[2]. Approximately one third of the world's population is infected with *Mycobacterium tuberculosis* (*Mtb*) and roughly 5-10% of these individuals will develop symptomatic tuberculosis[3]. Co-infection with immuno-compromising diseases, particularly human immunodeficiency virus (HIV), greatly increases the probability that symptoms will manifest[4]. With the current HIV epidemic and the emergence of drug resistant tuberculosis, including multi-drug resistant (MDR) and extensively-drug resistant (XDR) TB, TB has reemerged as one of the world's leading health threats, especially in regions of reduced socioeconomic status[5].

The causative agent of tuberculosis, *Mycobacterium tuberculosis (Mtb),* is a rod-shaped, extremely slow growing, acid-fast bacterium. *Mtb* is notorious for its high levels of drug resistance, a characteristic traditionally attributed to its greasy, impermeable, hydrophobic cell envelope. The thick cell envelope consists of an inner lipid bilayer, peptidoglycan linked to an arabinogalactan biopolymer, which is then sheathed in a second, outer bilayer composed of mycolic acids and extractable lipids. This thick envelope precludes entry of most substances (including potential drugs and probes) to the bacterial cytoplasm[6].

Even with this intrinsic drug resistance, many effective front-line antibiotics have been developed to successfully treat TB[7]. Prophylactic vaccines are also in development, but have been challenged by poor efficacy. The current vaccine, Bacille Calmette-Guérin (BCG), albeit widely employed, does not protect against adult pulmonary disease, but does provide some protection against extra pulmonary and pediatric disease[8]. One of the overarching problems within the fields of vaccine development and TB treatment is a lack of diagnostics for the assessment of the efficacy of TB chemotherapy. Sixth month relapse rates[9] are the accepted method for monitoring progress and these rates do not provide any short course, *in vivo*, methods for tracking the effectiveness of a particular agent on individual lesions. No methods currently exist to evaluate the metabolic state of the mycobacterium *in vivo*. Also, we lack tools to determine the specific microenvironments of particular lesions *in vivo*.

Besides sputum samples and biopsies, the medical community lacks quick, effective methods to monitor a patient throughout the course of therapy.

1.2 Model of the mycobacterial cell wall

An understanding of the structure, function and biosynthesis of the mycobacterial cell envelope is an important framework for the development of both novel mycobacterial therapeutics and diagnostics. While the chemical composition of this envelope has been known for decades[10-12], the 3-dimensional structure remains elusive. The mycobacterial cell wall does not conform to a "fluid mosaic" model of membrane structure and instead the structure of a single cell is composed of a single polymeric molecule with three layers: peptidoglycan, arabinogalactan and mycolic acids. Mycolic acids are long (C_{60}-C_{90}), often cyclopropanated, lipids found covalently attached to the arabinogalactan, and are important for bacterial outer membrane structure, virulence and persistence within the host[13]. These lipids, which give mycobacteria their name and distinct cording behavior, are one of the important components of the cell wall and will be the focus of this discussion of the mycobacterial cell envelope structure.

The physical organization of the cell wall polymer has been explored through biophysical analysis. X-ray diffraction experiments illustrated that the main chains of the mycolic acids are organized parallel to each other and freeze-fracture electron microscopy studies support, visually, the notion that a second hydrophobic barrier exists outside of the inner plasma membrane, peptidoglycan and arabinogalactan layers (**Figure 1.1**)[14]. Early electron microscopic evaluation of the coarse structure in thin-sections revealed an electron-transparent zone delimited by a darkly stained boundary[15,16]. These observations inspired a model of the covalently attached mycolic acids forming the inner leaflet[17,18] of the outer bilayer with the solvent extractable surface lipids intercalated between the asymmetric meromycolate chains[10,14].

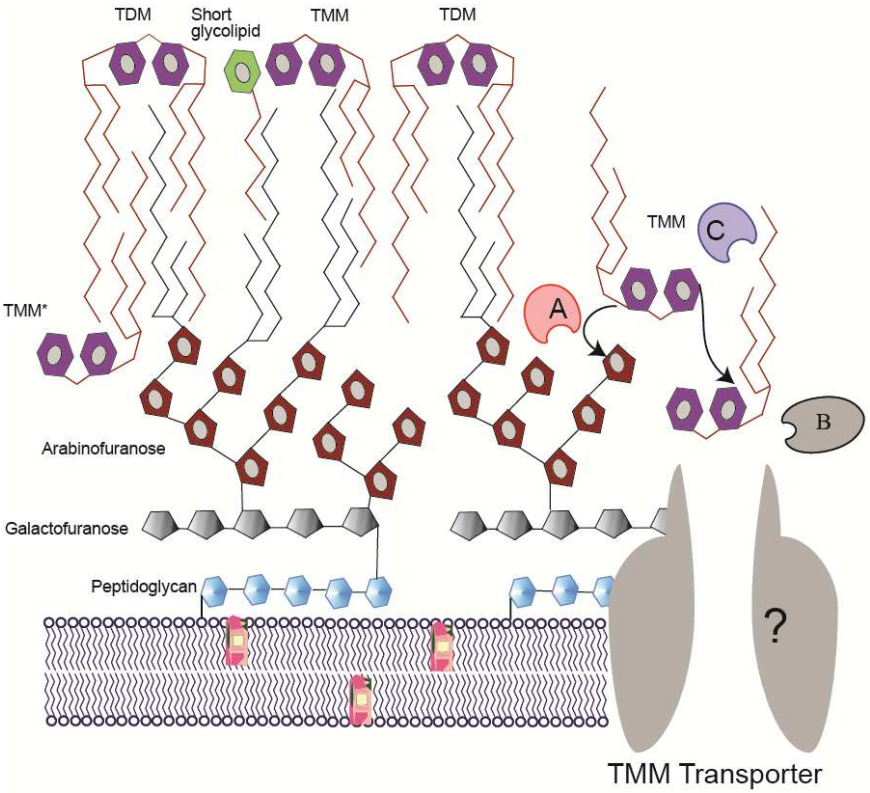


Figure 1.1. Schematic depiction of *Mtb* cellular envelope. Trehalose dimycolate (TDM) and trehalose monomycolate (TMM) associate to the outermost portion of membrane with their mycolates intercalated with the covalent mycolates. The Ag85 proteins transesterify the mycolates from TMM to either the arabinogalactan (AG) or to a second unit of TMM, producing TDM and free trehalose. The substrate specificity of Ag85A, B and C for the AG and/or TMM has not been fully characterized. TMM*- depicts opposite, putative orientation. A-Ag85A, B-Ag85B, C-Ag85C. Putative TMM transporter has not been identified. Diagram not depicted to scale. This figure was adapted from Tonge, 2000[19].

This model of the cell envelope dominated the field for more than twenty years but has been under recent scrutiny due to two observations. Distance measurements from crystal structure of the mycobacterial water-filled porin, MspA, revealed that for MspA to span the membrane and to function properly, the phospholipid bilayer must be narrower than was originally proposed[20-22]. Second, electron microscopy of cryopreserved cells revealed that the mycobacterial cellular envelope is no more than 8 nm in thickness (compared with 7 nm for the cell membrane alone)[23,24]. These recent observations provide evidence that the cell wall is no thicker than the plasma membrane and have fueled speculation about alternative physical arrangements of the envelope.

A consensus view has yet to emerge concerning the precise arrangement of these complex molecules. The hydrophobic mycolates are unlikely to directly contact with the hydrophilic environment. Outside of this layer, but still in intimate contact with the layer itself, reside the

numerous amphipathic molecules that are the direct interface for the pathogenic mycobacterium with the surrounding environment. This matrix of materials has been called the capsule and is the site of such structurally diverse families of glycolipids as trehalose mono- and dimycolate (TMM and TDM), the phthiocerol dimycocerosate waxes, sulfolipids, lipoarabinomannans, and glycopeptidolipids[10,17,25]. This survey will focus on the TDM and TMM glycolipids, as they serve as the mechanism presented in this thesis by which trehalose analogues can be incorporated into the *Mtb* cell wall.

1.3 Mycolic acids

The most significant part by mass and volume of the mycobacterial cellular envelope is the mycolic acids; hence taxonomists refer to these organisms as "mycolata." The presence or absence of mycolic acids, α-alkyl-β-hydroxy acids with length of $C_{>30}$, is a genus-specific trait and the fine details of their structure are used in chemotaxonomic approaches to species identification. The mycolic acids are found in two very important molecules in the cell: bound to arabinogalactan in the inner leaflet and esterified as trehalose-6,6'-dimycolate (TDM) in the non-covalently bound outer leaflet. There do not appear to be any structural differences between mycolic acids present in TDM when compared to those anchored in the cell wall (**Figure 1.1**). In several detailed studies of mycolic acid structure and function, the structures of mycolic acids throughout the mycolata were found to vary in three ways:

i) Chain length – longer chain lengths increase toxicity, granulomagenicity, pathogenicity, and the phase transition temperature of the cell wall.
ii) Oxygenated moieties – for example, the keto- and methoxy-mycolic acids characteristic of virulent strains and the epoxymycolates and wax esters that characterize some nontuberculous mycobacteria (NTM).
iii) *Cis*- or *trans*-cyclopropanations, and *cis*- or *trans*-alkenes some with adjacent methyl branches.

The analysis of cell wall-bound mycolates requires treatment of whole cells, or delipidated cells, achieved with *tert*-butyl ammonium hydroxide, which saponifies the bound mycolyl esters. The mixture of mycolates obtained can be converted to methyl esters and purified by silica gel column chromatography or silylated and analyzed by gas chromatography mass spectrometry (GC/MS)[26]. Together with matrix-assisted laser desorption/ionization (MALDI)

analysis[27-29], these methods have allowed full comparison of the lipid lengths, keto versus methoxy variability, and the degree of cyclopropanation from a variety of *Mtb*, *Mycobacterium bovis* and *Mycobacterium kansasii* strains. The mycolic acids of *Mtb* contain three general classes: α, methoxy, and keto, while the mycolates of other mycolata can contain additional functionalities and lengths as illustrated in **Figure 1.2a**.

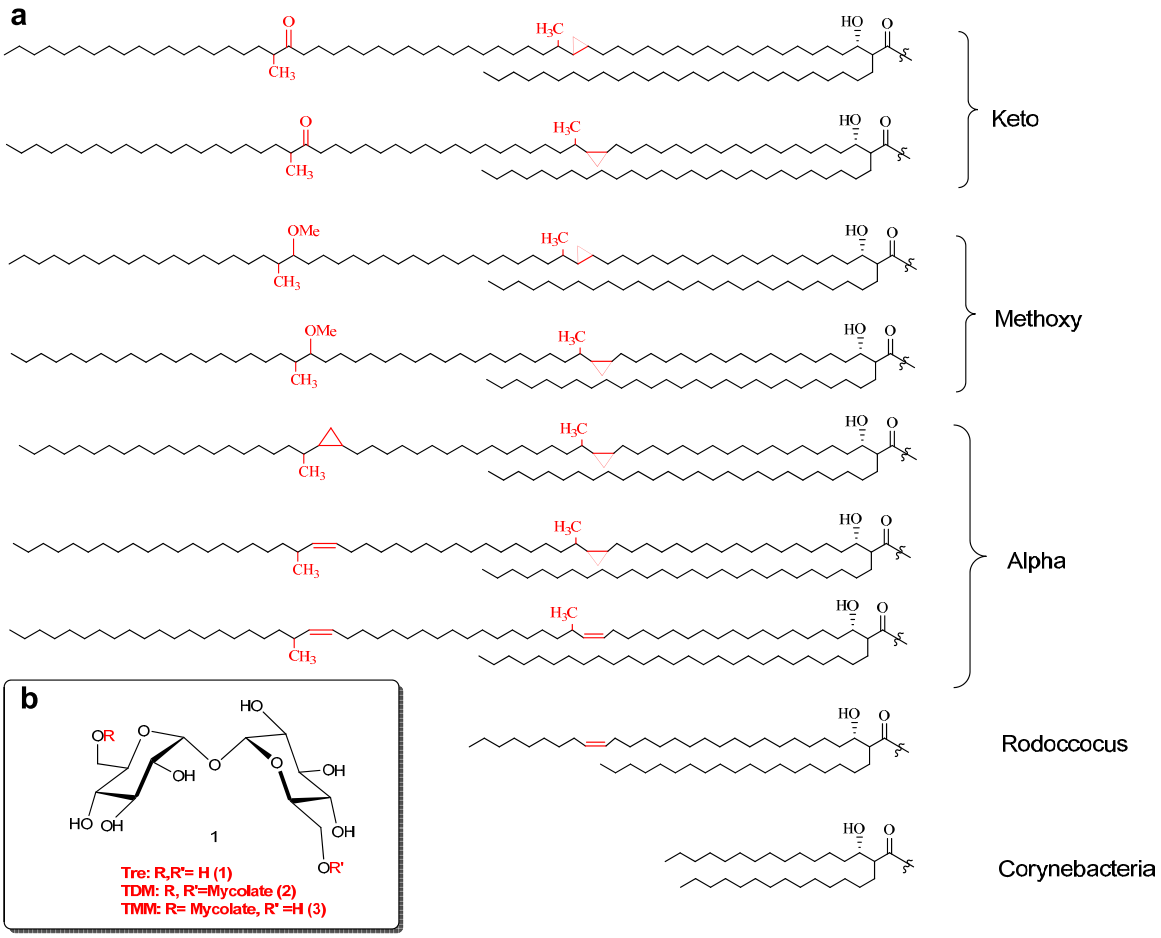


Figure 1.2. (a) Varieties of mycolic acids (b) Depiction of the structure of trehalose (1) trehalose dimycolate (2, TDM) and trehalose monomycolate (3, TMM), transesterified with the mycolates found in 'a.'

1.4 Trehalose mono and dimycolate (TMM and TDM)

Trehalose monomycolate and dimycolate (TMM and TDM, **3** and **2**), the mycolic acid esters of trehalose, are present in all Actinomycetes. They are widely distributed among both pathogenic and nonpathogenic strains of the mycolata, notably in the acid-fast mycobacteria. These glycolipids are the most abundant, non-covalently associated cell wall lipids, which

provide these organisms with a waxy, impermeable outer membrane and the characteristic of acid-fastness. TMM and TDM serve a variety of purposes during infection and are purportedly trafficked during initial infection and persistence of mycobacteria within macrophages[30]. TDM and TMM are thought to be oriented with their mycolates towards the cell cytoplasm, presenting the disaccharide moiety to the external environment of the bacterium as shown in **Figure 1.1**; although TMM has been found to be buried more deeply in the cell wall, suggesting that it and may also orientate itself with trehalose head group directed inwards[31]. Because of its importance for disease progression, we have chosen to highlight some of the characteristics of TDM, focusing on its discovery, structural characterization, biosynthesis, and immunomodulatory behavior as well as some of the implications of TDM research in tuberculosis drug and vaccine development.

1.4.1 Structural features of TDM

Interest in the trehalose mycolates began in the 1950s with their discovery and characterization by Noll and Bloch[32-34]. Before their structure was assigned, the dimycoloyl trehaloses were named "cord factors" by Bloch, who observed that the pathogenic strains of *Mtb* grew in tightly wound bundles of snake-like "cords." The crude, extractable, non-covalently cell wall-associated lipids that bestow this property were characterized by hydrolytic degradation, acetylation, reductive cleavage and combustion analysis, yielding the structure 6,6'-dimycolyl-α,α'-D-trehalose (**Figure 1.2b**)**.** The carbohydrate moiety of TDM and TMM is trehalose (Tre, **1**), a non-reducing disaccharide, consisting of two glucose molecules with an α,α',1,1'-linkage. TDM is symmetrically modified on the 6,6'-hydroxyl groups with mycolic acid esters.

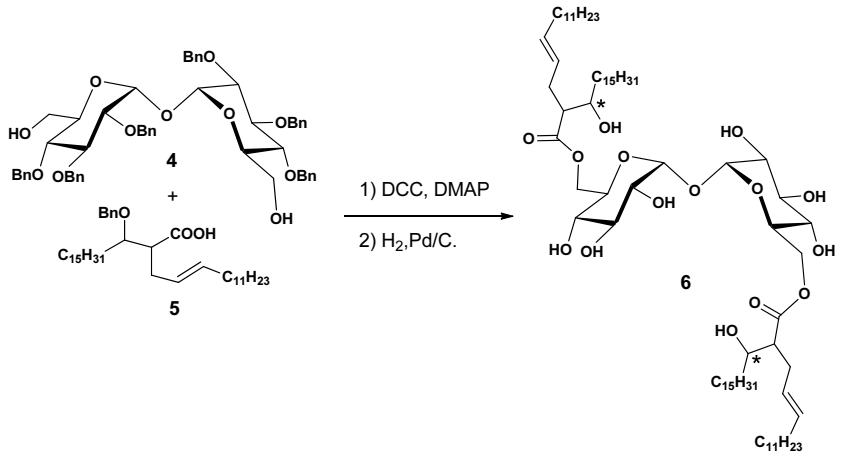
Trehalose is an important sugar for many of *Mtb*'s glycolipids and is also found in sulfolipid-1 (SL-1), diacyl trehalose (DAT), and triacyl trehalose (TAT). To analyze TDM and other cell wall-associated lipids, cells are harvested, dissolved in methanol and aqueous brine, and sequentially extracted with petroleum ether to remove apolar lipids (trehalose mycolates, phthiocerol dimycocerosate (PDIM), PGL, and acyl-trehalose) followed by chloroform:methanol (2:1) to remove the relatively polar lipids [phosphatidyl-*myo*-inositol mannoside (PIM), lipoarabinomannan (LAM), and glycopeptidolipid (GPL)]. The TDMs contain a diverse mixture of mycolic acid components that can be separated from other cell wall components by preparative TLC and then saponification to remove the trehalose. The remaining fatty acids are then methylated to provide the carboxylated methyl esters. As an

alternative to saponification, MALDI-TOF mass spectrometry has proven useful for characterizing heterogeneous mixtures of TDM[27,29] as whole glycolipids. Prior to MALDI, analysis by electron ionization mass spectrometry (EI/MS), fast atom bombardment mass spectrometry (FAB/MS)[35] and GC/MS proved helpful for characterization. However, these methods were insufficient for determining the mass of intact TDMs, which is typically greater than 2.6 kDa.

1.4.2 Prior syntheses of TDM and TMM

Natural TDMs contain an inseparable mixture of mycolic acids of varying lengths and degrees of cyclopropanation. Due to this heterogeneity, synthetic methods have been used to obtain pure samples of some non-mycobacterial TDMs. *Mtb* TDM has not been synthesized through fully chemical means because the chiral cyclopropanated and oxygenated mycolates are difficult to access. *Mtb* TDMs have not been reproduced through fully synthetic processes, but have been prepared from natural extracts by semisynthesis[36]. TDMs from related bacteria[37] and various analogues of TMM and TDM[38-40] have been created fully synthetically. Pure, synthetic TDMs are useful tools to probe structure-function relationships and have also been used as adjuvants for the development of vaccines against *Mtb*[38,41,42].

Semi-synthetic strategies capitalize on the ability to purify native mycolic acids from *Mtb* and then couple them to benzyl or silyl protected trehalose by common esterification techniques[41,43]. The mycolyl β-hydroxyl is an additional complication, and must be protected to prevent dehydration. The 2,2', 3,3', and 4,4' protected trehaloses can be obtained by utilizing either selective silyl protection[44] or selective mesylation[45]. Once the 6 and 6' hydroxyl groups are freed, a variety of esters can be coupled to form TDM analogs. A summary of this synthetic strategy can be found in **Scheme 1.1**.



Scheme 1.1. Example synthesis of TDM analogue by Nishizawa, 1996[41] demonstrates that TDM analogues can be synthesized from hexabenzyl-trehalose 4. Final DCC coupling and deprotection yielded the native corynebacterial TDM as well as 3 stereoisomers (*).

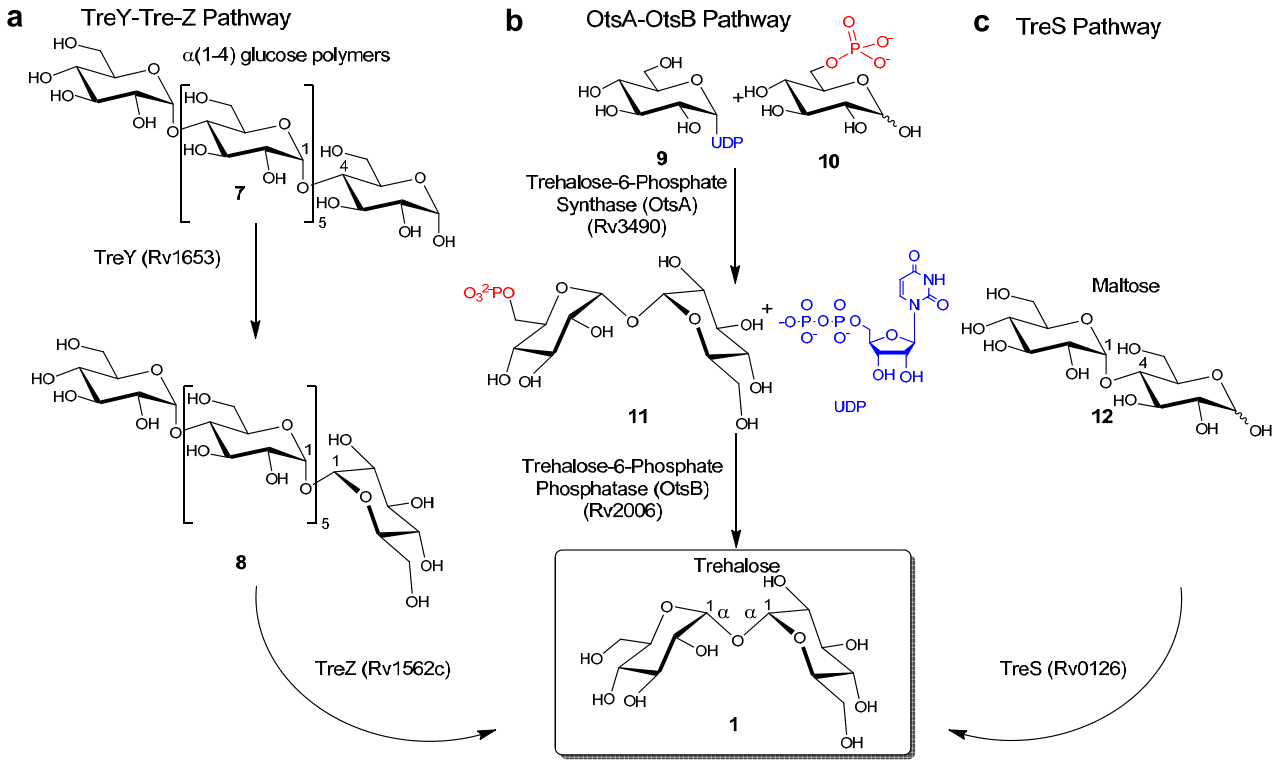
Synthesis has allowed researchers to probe the important features of TDM. The 6,6' trehalose di-ester of 2-eicosyl-3-hydroxy-tetracosanoic (behenylbehenic) acid, which has a similar structure to TDM, has been produced semi synthetically by Toubiana, et al.[36], as well as the straight chain TDM analogue. The 2*R*,3*R* stereochemistry of the mycolic acids in the trehalose dicorynomycolates (TDCMs) of corynebacteria can be manipulated chemically and all four stereoisomers of TDCM have immunoadjuvant activity. Natural R,R TDCM and the synthetic S,S TDCM showed the most significant immunomodulatory activities[41].

1.4.3 Biophysical studies of TDM properties

A peculiar feature of TDM is its dichotomous ability to form both monolayers and micelles. These structures have been confirmed by scanning tunneling microscopy (STM) and have been linked to the cord-like growth patterns of mycobacteria. Biologically, these two structures determine toxicity; micelles with trehalose on the surface, which are non-toxic even in high concentrations, versus surface adsorbed monolayers of TDM which expose the hydrophobic tail surface and exhibit extreme toxicity to cells[46]. TDM is thought to be a substantial reason for the low permeability of the mycobacterial membrane[47].

1.4.4 Biosynthesis of TDM

The biosynthesis of TDM and TMM begins with the separate synthesis of trehalose and mycolic acids. Trehalose is synthesized by *Mtb* through three independent pathways, and genetic knockouts of any of these three pathways results in either non-viable *Mtb* or organisms displaying growth defects[48,49]. Of these three pathways, it has been demonstrated that the OtsA/OtsB pathway (**Scheme 1.2**) is essential for trehalose biosynthesis[48]. Knockouts of the Otsa/OtsB pathway fail to grow, and addition of trehalose to the medium reverses the growth defect. Several of the enzymes responsible for trehalose synthesis, including trehalose-phosphate synthase (TPS or OtsA)[50,51] and trehalose-phosphate phosphatase (TPP or OtsB)[52] have been purified and analyzed. The crystal structure of OtsA from *Escherichia coli* has been solved[53], and from this structure Gibson and coworkers have deduced a possible mechanism. TPS (OtsA) catalyzes the transfer of glucose from UDP-glucose to glucose-6-phosphate to form trehalose 6-phosphate, a double glycoside in which both anomeric positions are set simultaneously. Trehalose phosphatase (OtsB) then cleaves off the phosphate group to generate trehalose (**Scheme 1.2**). Trehalose is metabolized by endogenous trehalase enzymes, which cleave the glycosidic linkage; the *Mtb* trehalase has been purified from *Mycobacterium smegmatis*[54]. A trehalose transporter has been identified[55] and is thought to salvage extracellular trehalose for mycobacterial use. One of the most important functions of trehalose in mycobacteria is as the sugar scaffold for mycolic acids in TDM.



Scheme 1.2. Three biosynthetic pathways for the synthesis of trehalose are present in *Mtb*. (a) In the TreY-TreZ pathways TreY isomerizes the α-1,4 linkage in glucose polymers (amylose) to an α,α-1,1 linkage and then TreZ hydrolyzes the trehalose from the polymer. (b) In the OtsA-OtsB pathway OtsA couples UDP-glucose and glucose-6-phosphate to form trehalose-6-phosphate, which is then dephosphorylated by OtsB. (c) TreS isomerizes the 1,4 linkage of maltose to the 1,1 trehalose linkage

Once synthesized, trehalose is transesterified with a mycolate, through a partially characterized pathway, to generate TDM and TMM. In the biosynthetic model proposed by Takayama, et al (**Figure 1.3)**, mycolic acids are synthesized by the well characterized FAS-II enzymes[56]; fully formed mycolic acids are transferred from the polyketide synthase (*pks13*) to a heptaprenyl group, by coupling D-mannosyl-1-phosphoheptaprenol to the mycolates, yielding 6-*O*-mycolyl-β-D-mannosyl-1-phosphoheptaprenol (Myc-PL)[57]. This model is supported by characterization of Myc-PL by Besra et al[58], but the full scope and properties of the Myc-PL carrier have yet to be researched. This transesterification of mycolate to the mannosyl carrier is thought to be catalyzed by the putative mycolyl transferase I. The prenyl tail of the Myc-PL causes it to migrate and dock with the proposed ATP-binding cassette (ABC) transporter, after which a second reaction transfers the mycolate from Myc-PL, *via* putative mycolyl transferase II to trehalose-6-phosphate (synthesized by OtsA) to form TMM-phosphate. A membrane associated phosphatase for trehalose-6-phosphate cleaves the

phosphate, and the TMM is then immediately exported through the putative transporter. The specific role of this transporter has not been biochemically validated, but the current data corroborates this model well[59,60]. This coupling constructs the TMM molecule and exports it to the cell wall. TDM is not produced by a second acylation from Myc-PL, as might be expected. Instead, three extracellular proteins, Antigens 85A, B, and C [Ag85A,B,C, also known as fibronectin-binding proteins (Fbp)], perform a transesterification reaction to generate one TDM molecule and one free trehalose out of two TMM units, or to transesterify mycolates to the arabinogalactan, generating the covalent outer bilayer[61-63].

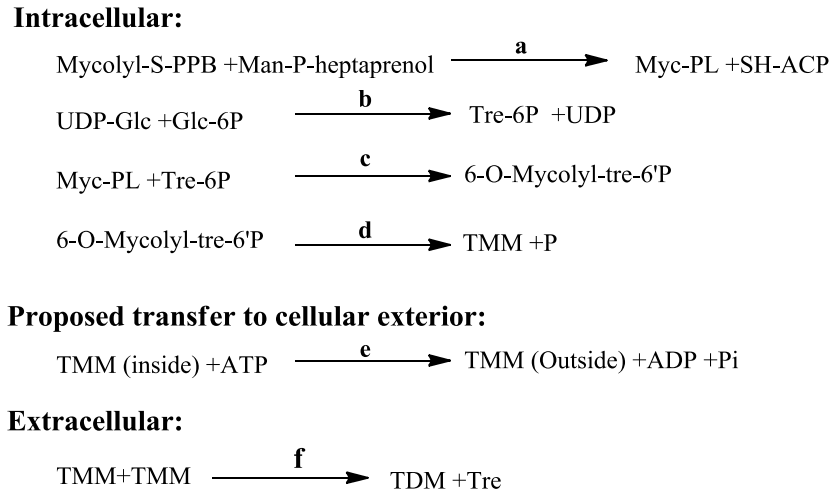


Figure 1.3. Proposed synthesis of TDM (a) Putative mycolyl transferase I (b) Trehalose-6-phosphate synthase (OtsA) (c) Putative Mycolyl Transferase II (d) Trehalose-6-phosphate phosphatase (OtsB) (e) Proposed ABC transporter (f) Antigen 85A, B and C. This figure was adapted from Takayama et al., 2005[57].

1.5 Antigen 85 Enzymes

Trehalose is anchored into the mycobacterial cell wall as mono- (TMM) or di- (TDM) mycolates by the action of the extracellular proteins Antigens 85A, 85B and 85C (Ag85A,B and C). Although these proteins were long known to immunologists by virtue of their immunogenicity, their enzymatic activity was first described in 1982 as the active fraction of a cell free extract from *M. smegmatis* that was capable of synthesizing TDM[39]. Ag85A, B, and C are the most abundant extracellular *Mtb* proteins, making up as much as 41% of the culture supernatant protein content[64]. They are expressed in a ratio of 3:2:1 (Ag85B, Ag85A and Ag85C, respectively)[65]. Genetic inactivation of 85C causes a 40% decrease in covalent mycolic acid content of the cell wall[63], suggesting that the Ag85 enzymes play a role in

mycolation of the cell wall. Triple knockouts of the *ag85ABC* genes have not been reported in *Mtb*, presumably because they are not viable. The three Ag85 *Mtb* proteins have high sequence homology to TMM transesterases from *M. leprae*, *M. bovis*, *M. avium* and C*orynebacterium glutamicum*. The corynemycolyl transferases, particularly PS1, are the best studied of this class of enzymes and have provided much insight into the *Mtb* antigen 85 complex[14,61].

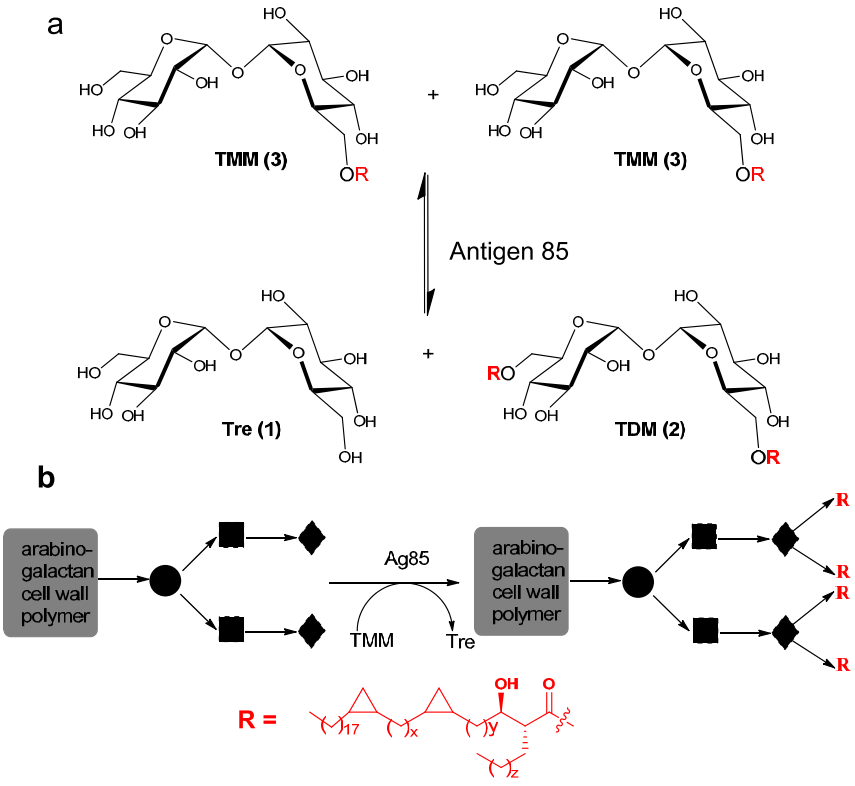


Figure 1.4. Antigen 85 catalyzes the mycolation of the mycobacterial cell wall. (a) Ag85A, B, and C catalyzed transesterification of trehalose (Tre) 1, trehalose dimycolate (TDM) and trehalose monomycolate (TMM). x (10-20), y (17-20), z (25-30). (b)Ag85 catalyzed mycolylation of arabinogalactan. Diamond: β-D-Araf, Square: 2-α-D-Araf, Circle: 3,5-α-D-Araf.

Ag85 isoforms all catalyse the reversible transesterification reaction between two units of TMM, generating TDM and free Tre; the reverse reaction also allows for direct esterification of Tre. Ag85 can also covalently introduce mycolates into arabinogalactan to form the base polymer of the cell wall (**Figure 1.4**)[61,63].

1.5.1 Ag85 partial redundancy

A brief survey of the antigen 85 and mycolate literature revealed that the Ag85 enzymes can partially but not fully rescue one another, suggesting partial but incomplete redundancy of these enzymes[61-63,66]. Deletion of the corynebacterial PS1 disrupts the ratio of the trehalose di and monocorynemycolates (TDCM and TMCM) and depletes the cell wall corynemycolates. PS1 knockouts can be complemented by over expression of PS1 and such complementation restores the TDCM/TMCM ratio and results in an over-incorporation of mycolates into the cell wall[61]. Complementation of PS1 with the *Mtb* Ag85A,B,C does not restore the TDCM/TMCM ration, but does increase the covalent mycolates. This result suggests that over expression of the Ag85 enzymes can partially complement for PS1 and can lead to over incorporation of mycolates into the cell wall. Originally it was believed that corynebacteria expressed only one mycolyl transferase, PS1. Subsequently, 5 additional mycolyl transferases were identified in corynebacteria. These fall into two classes, which can partially complement for one another, hinting at structural differences that might relate to complementation[66].

With the knowledge of the partial redundancy of the corynemycolyl transferases, the authors of the corynebacteria studies turned their attention towards similar study of the Ag85 enzymes in *Mtb*. Analysis of the Ag85A,B,C knockout mutants and their complements revealed that A and B can partially restore the *ag85C* deletion, but no better than C itself[62]. The authors also found that, unlike the *ag85C* knockout, deletion of *ag85A* and *ag85B* do not cause a cell wall mycolyl defect. This group, Daffé et al, also considered that the defect might not be in the gross level of mycolates but in the linkage of the mycolate to the arabinan. Contrary to this hypothesis, Daffé et al found that neither Ag85A nor Ag85B attaches mycolyl residues to specific sites but, like Ag85C, each enzyme transfers mycolyl residues onto all four sites present in the arabinogalactan non-reducing end hexaarabinosides. Degradation of mycolates from the TMM and TDM to triacyl glycerol (TAG) has been documented[67], but the reverse has not, hinting that one or more of the Ag85 enzymes may process TAGs as a substrate. The sum of these results suggests that there are subtle differences between the Ag85 enzymes that have yet to be fully characterized.

Other notable observations regarding the Ag85 enzymes and mycolates have implications for permeability of mycobacteria to drugs. It was found that the absence of mycolic acids does not affect cell wall structure, but does increase bacterial permeability and susceptibility to

antibiotics[68]. Generation of the double mutant of *cmtA* and *cmtB* in corynebacteria causes changes in freeze fracture and growth defect[69]. The results hint at the complex and important structural and functional roles of the mycolates and their corresponding mycolyl transferases responsible for cell wall synthesis.

1.5.2 Ag85 crystal structures

The crystal structures of the three Ag85 isoforms been solved[70-72] and these structures reveal that A, B and C share high sequence and structural homology, characterized by an α,β-hydrolase fold and a hydrophobic fibronectin-binding domain. The X-ray structures support a transesterification mechanism analogous to the serine hydrolases, in which the formation of a covalent Ser-mycolic acid enzyme intermediate is followed by attack from the 6' hydroxyl of the trehalose[70]. The active sites of 85A, B, and C are highly conserved, featuring a His, Glu, Ser catalytic triad, a hydrophobic tunnel for the mycolic acid, two trehalose binding sites, and in Ag85A and B, a key phenylalanine residue (Phe232) that is thought to regulate a hydrophobic channel between the two binding sites. This second channel may support a "scooting mechanism," whereby the second TMM substrate can flip from its binding site to the active site without exposing the mycolate to the solvent (Figure 1.5). Notably, in Ag85C Phe232 is replaced by Leu230, which may have implications in the relative catalytic efficiencies and substrate affinities of the three enzymes.

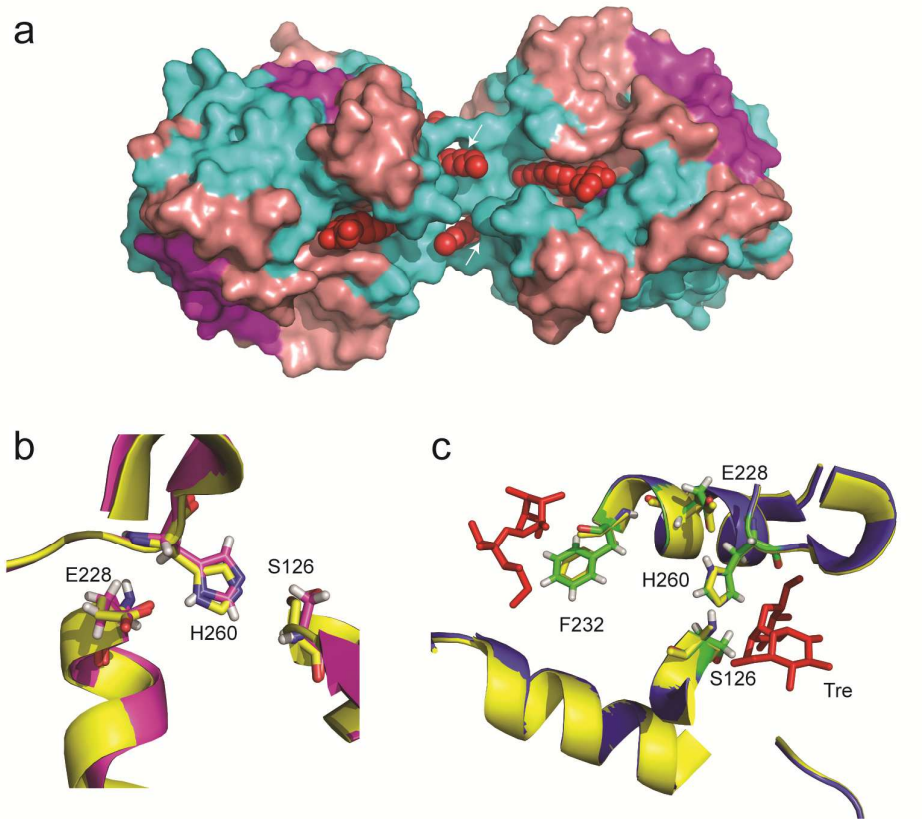


Figure 1.5. (a) Crystal structure of antigen 85C bound to S-octylglucose (PDB 1VAS). Structure shown is a dimer of two Ag85C proteins each bound to two octylglucose ligands. Proposed hydrophobic channels, through which a second TMM could traverse the active site are visible and indicated by white arrows. (b) active site overlay of Ag85B and Ag85C shows very strong structural homology (PDB 1f0n and 1DQZ) (c) Crystal structures of Ag85B and 85B bound to 2 trehalose molecules (Tre, red sticks) with one co-crystallized in the active site and one in the secondary site. Active sites overlaid show very little motion when bound to substrate. Phe232 moves slightly to make room for second trehalose to enter active site. In Ag85C Phe232 is replaces by Leu230 (PDB 1f0n and 1f0p).

1.5.3 Ag85 activity assays

The development of Ag85 enzymatic assays has been challenged by the heterogeneity and poor solubility of the TDM and TMM mycolate containing substrates. This enzyme activity was used as a basis for the original enzyme purification and demonstration that TDM was derived from TMM. Antigen 85 activity can be measured in a ^{14}C-trehalose exchange assay, where TMM is mixed with ^{14}C -trehalose and enzyme, as depicted in **Figure 1.6**[73,74]. Turnover is measured by scintillation counting of the lipid extract, monitoring for ^{14}C-trehalose incorporation into the lipids. Auto radiographic TLC analysis can be used to calculate approximate k_{cat} values. Using this assay, 85C was found to be much more active than either A or B[74]. Through this assay the full kinetic analysis, substrate specificity, and inhibitor screening of these enzymes could not be explored fully.

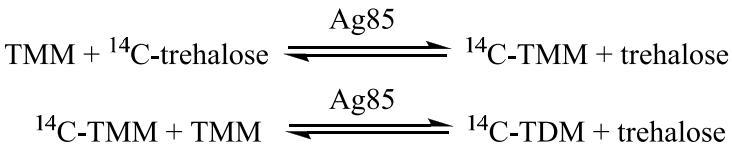
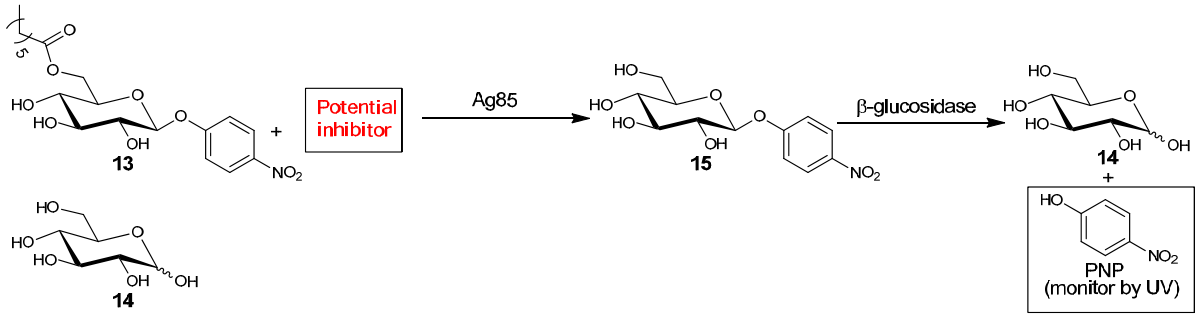


Figure 1.6. ^{14}C -trehalose Antigen 85 (Ag85) activity assay has afforded rough kinetic parameters for the AG85 enzymes[73,74]. Assay is not high throughput and is challenged by heterogeneity of the natural substrates.

High throughput screening of inhibitors is not possible with the cumbersome ^{14}C-trehalose assay. Ronning and coworkers[75] have devised a coupled assay, utilizing a *p*-nitrophenol-glucose substrate **13**, antigen 85 and β-glucosidase (**Scheme 1.3**). This assay has potential for high throughput screening applications and has helped to prove that antigen 85 transfers mycolates onto arabinogalactan[76]. It is challenged by both the non-natural substrates (glucose **14** in place of trehalose **1**) and by the secondary readout provided by coupled assays. A second coupled Ag85 assay has also recently been reported[77]. Elamin et al employed native TMM and Ag85A to generate free trehalose, which is then hydrolyzed by a trehalase included in the reaction mixture, and the glucose production is measure colorimetrically. This assay has the advantage of employing the native substrate, but as with the previous coupled assay, relies on a secondary readout and fails to differentiate between mycolyl transfer and substrate hydrolysis. Both of these assays can only access catalytic parameters for the first step in catalysis and therefore do not provide a full enzymatic analysis of the Ag85 enzymes. Even with these advances there remains a need for precise, high-throughput Antigen 85 enzyme assays.



Scheme 1.3. Novel antigen 85 coupled activity assay[75] based on hydrolysis of *p*-nitrophenol (PNP)-glucose (15) by β-glucosidase, yielding PNP, which can be quantified by UV.

1.5.4 Trehalose as a drug target

Because of Antigen 85 and TDM's roles in pathogenesis (see **Section 1.6**, Hunter et al.[78] and Saunders et al.[79]), inhibitors of the Ag85 complex have been proposed as leads for anti-tubercular antibiotics. Studies with antigen 85 mutants have suggested that inhibition of these proteins would increase susceptibility of the bacterium to other drugs. Antigen 85A mutants displayed increased sensitivity to front-line TB drugs and other broad-spectrum antibiotics[80]. One of the most substantial problems with antimycobacterial drug development is targeting both replicating and non-replicating bacteria[81]. Trehalose is involved both in cell-wall biosynthesis and protection from osmotic stress. Therefore, trehalose analogs have been proposed as possible drug targets for both replicating and latent bacteria and potentially provide a novel target for treatment of XDR strains[82].

The designed inhibitors of Antigen 85 (**Figure 1.7**) have large focused on trehalose scaffold[83-85] with modifications made to the reactive C-6 such as azides[74], amines[83], alkyl amides or sulfonamides[84,85], or phosphonates[86,87]. Phosphonates and sulfonamides, such as **21**, mimic the tetrahedral intermediate. The presence of larger lipophilic groups correlates with increased inhibition, but this causes problems with solubility in the assay. Recent work developing non trehalose-based inhibitors of Ag85 focused on diaryl phosphonate trehalose mimics **18**[88] and methyl 5-*S*-alkyl-5-thio-D-arabinofuranosides **19**[89] (**Figure 1.7a**), both of which suffer from poor solubility and low inhibitory activity against Antigen 85. Trehalose analogues with free 6' hydroxyl are complicated by the fact that they could function as either substrates or inhibitors of Antigen 85. Although desirable, to date, no potent, selective Antigen 85 inhibitors have been developed.

The Antigen 85 crystal structures reveal two trehalose binding sites separated by a hydrophobic channel. Anderson et al.[71] have proposed a novel series of inhibitors that would target both binding sites, consisting of two trehalose molecules linked with a hydrophobic linker. However, these compounds have yet to be synthesized and validated as inhibitors. The fundamental problem in design of all Ag85 inhibitors is the lack an effective assay to directly measure antigen 85 inhibition. MIC and other growth inhibition assays are useful, but run the risk of highlighting off-target effects. The assays detailed above are coupled to other enzymes, so could read off-target activity and fail to differentiate between mycolyl transferase and hydrolase activity of the Ag85 enzymes.

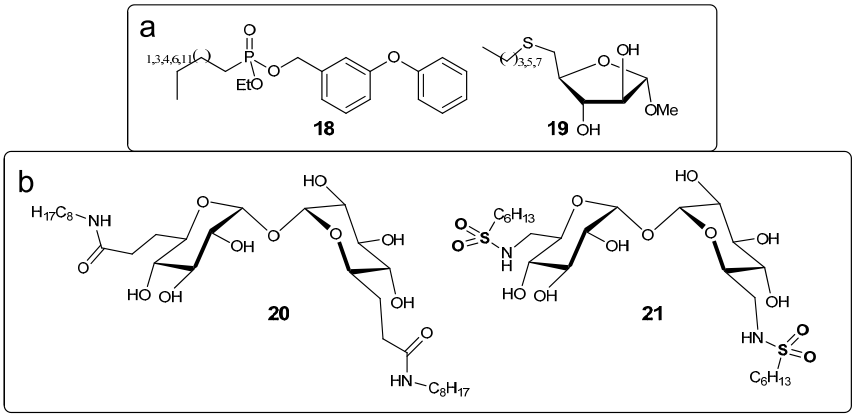


Figure 1.7. Representative work towards inhibitors of antigen 85. (a) Non-trehalose based scaffolds, (b) Trehalose based inhibitors.

Trehalose derivatives also offer the potential to inhibit the trehalose synthase (OtsA) and the trehalose phosphatase (OtsB). OtsA offers an unusual glycosyltransferase mechanism that might be inhibited by fluorinated trehalose. OtsA is known to be a rare glycosyltransferase that can use any of the nucleoside diphosphate glucose derivatives[51]. Trehalose synthesis proceeds with net retention, placing OtsA in the family of retaining glycosyl transferases (GTs). The crystal structure of OtsA suggests an oxocarbenium transition state and S_{Ni}-type mechanism[90]. Work with 5-F-glucose derivatives suggests that fluorination at the 5 position can destabilize the oxocarbenium transition state[91] and so 5-F-trehalose has been suggested as possible mechanistic inhibitors of OtsA and trehalase.

Few potent inhibitors of Ag85 and OtsA exist and inhibition of trehalose and TDM biosynthesis presents a novel mechanism to target and potentially kill *Mtb*. The evaluation of both trehalose analogues and non-trehalose based inhibitors of Ag85 is hindered by the lack of effective, high throughput enzyme assays.

1.6 Immunopathology and other functions of TDM and Ag85

In addition to its roles in the mycobacterial cell envelope, TDM (or cord factor) also serves as a toxic signaling and potently immunomodulatory molecule that is thought to be one of the most important players in mycobacterial survival and persistence within the host. Research has suggested that TDM acts as a signaling molecule[92] and that it prevents the membrane from oxidative stress[93] and that TDM diffuses into host cell membranes and damages their function. TDM also induces apoptosis[94] and can singlehandedly induce granulomas even in the absence of mycobacterial infection[95]. There is vast literature, dating back to the 1960s on

how mycobacterial cell wall lipids induce granulomas[95-102]. Their potency as adjuvants explains the presence of these lipids in the commonly utilized Freund's adjuvant.

Alongside TDM, the Antigen 85 proteins are potently immunogenic and much of the current Ag85 research focuses on the use of Ag85 in vaccines and TB diagnostics[103]. The Ag85 proteins are one of the major targets in the immune response and the IgG antibody response to the antigen 85 complex is correlated with disease outcome[104]. The Ag85 complex contain well documented T cell epitopes[105] and the synthesis of glucose monomycolate (GMM) by Ag85 has been reported as another mechanism by which *Mtb* modulates the host immune response[106].

Mycobacteria evade the host immune system in part by hiding within host cells, particularly macrophages. Although one of the functions of macrophages is to engulf and destroy foreign objects, *Mtb* inhibits and evades the initial macrophage destructive onslaught. *Mtb* are engulfed by macrophages through a process known as phagocytosis. Once inside the macrophage phagosome, the *Mtb* block the normal process of phagosomal acidification and fusion with the lysosome, arresting maturation, allowing the bacteria to survive at a mild pH within the host cell. The causes of phagosomal arrest are not fully understood, but TDM is documented to be involved in this process[107,108]. When macrophages fail to destroy the bacteria, the immune system invokes its next line of defense, forming granulomas around the infected macrophages and sealing off the area of infection with layers of host cells. The formation of the granuloma has the effect of containing the disease in a state that is characterized by low bacterial counts and regions of T cell, macrophage and neutrophil infiltration and more general inflammation.

1.6.1 TDM and granuloma formation

Granulomas induced by mycobacterial TDMs are caused by two responses: 1) foreign body derived (innate immunity) and 2) hypersensitivity derived (adaptive immunity, mediated by T cell responses for specific antigens)[109]. The innate response is characterized by the non-specific interception of foreign material by complement proteins and phagocytotic cells. Phagocytes and antigen-presenting cells capture foreign material, release the pro-inflammatory mediator tumor necrosis factor-α (TNF-α), and digest peptides and lipids. Upon human inhalation of *Mtb*, the bacterium reaches the alveoli and releases TDM and other antigens that interact with host lipids, macrophages, and proximal tissue. TDM is toxic to these host cells, in a manner thought to be not unlike that of silicon dioxide[110]. The

aggregate debris does not clear well, possesses adhesive properties, and triggers recruitment of more immune cells, which results in formation of a granuloma to enclose the debris. The structure of the granuloma formed from this immune response features an infected core, surrounded by necrotized cells, macrophages infiltration, neutrophils and additional damaged host tissue.

After an infection has occurred, the adaptive immune processes begin and subsequent exposure to TDM antigen results in hypersensitivity. As part of the adaptive immune response the TDM and other immune epitopes are digested and are presented to T cells. Presentation occurs *via* MHC II molecules for peptides, or CD1 molecules for most lipids, and via a unique C-Type lectin, Mincle, for TDM[111]. Notably, TDM activated macrophages produce inflammatory cytokines and nitric oxide, which are completely suppressed in Mincle-deficient macrophages. Not all TDMs are equally efficient at triggering an immune response; TDMs from pathogenic species, which tend to contain longer mero branch hydrocarbon chains with additional functionality, have larger effects than those from their avirulent cousins[112].

As part of both the innate and adaptive immune response, TDM plays a predominant role in stimulating pro-inflammatory cytokine production and in inhibition of phagosome maturation. Mycobacteria arrest phagosome maturation at an early stage, preventing acidification. Delipidated mycobacteria cannot do this and have decreased survival in macrophages[107], pointing to TDM as a potential culprit. After 24 hours inside macrophages, *Mtb* bacilli contain up to 50% more TDM than at initial infection. Delipidation of *Mtb* also leads to reduced cytokine production, particularly IL-6, IL-12 and TNF-α, and decreased bacterial viability in the host[107,108,113]. The inflammatory cascade initiated by TDM is known as the cytokine/chemokine storm. This storm of activity does not continue unabated during *Mtb* infection. Instead IFN-γ, which is also produced during cellular immune response, down regulates the pro-inflammatory response. At this point, a stable granuloma has been established. Characteristics of a stable granuloma include “foamy macrophages” and giant cells as well as regions of necrosis, the presence of T cells, B cells, vascularization, hypoxia, and extracellular mycobacteria trapped within the granuloma. [114,115] This cascade of events is illustrated in **Figure 1.8**.

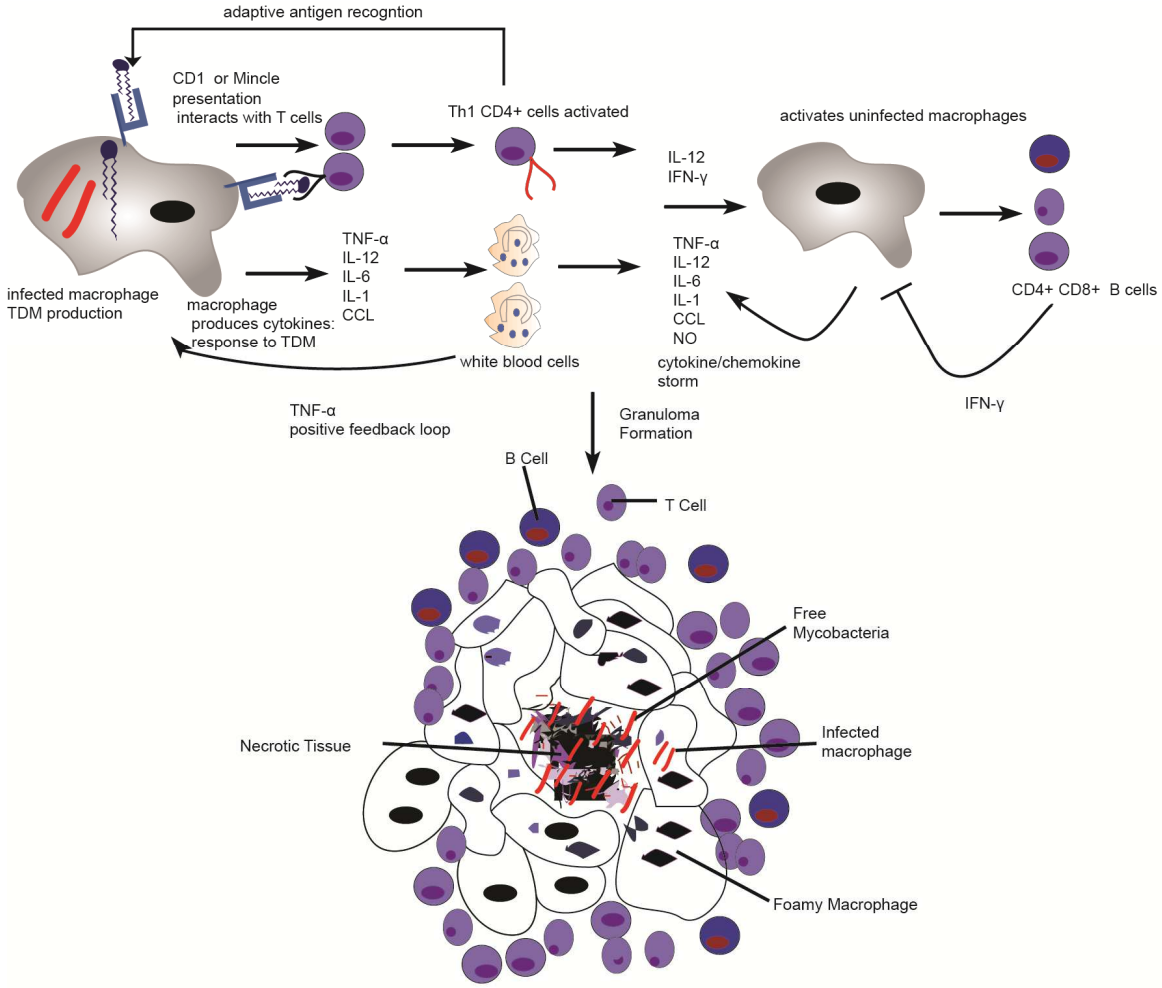


Figure 1.8. TDM plays an important role in the survival and persistence of *Mtb* within the host, including stimulating the immune response and granuloma formation. Arrows indicate cascade of events caused by TDM interaction with host immune system. This model was proposed by Ryll et al[116].

The formation of a granuloma, triggered in part by TDM, has ambiguous results. To the host's benefit, it helps to prevent the spread of infection. To the host's disadvantage, the granuloma creates a walled-off, locally isolated environment that suppresses diagnostic markers, presents a barrier to diffusion of drugs and prevents the immune system from clearing the entire infection. Granulomas provide a niche for the survival of latent, metabolically inactive bacteria. Rupture of these lesions provides a means for the infection to spread to a new host. This phenomenon of survival and containment underlies the stable "latent" infection present in one-third of the global population.

1.6.2 Immunopathology and biological properties of Ag85

Alongside TDM, the antigen 85 proteins are potent immunogens in their own right, invoking a diverse set of immunological responses, including immunoglobins (Ig) IgG and IgA[117]. Antigen 85-derived peptides are presented by the MHC II complex. The Ag85 peptide epitopes induce T cell differentiation, modulate the production of cytokine-producing T cells and the release of IFN-γ[118]. The apparent redundancy of three isoforms of the antigen 85 proteins has been proposed to help the bacteria evade the host's immune system, although the hyper conserved nature of mycobacterial immune epitopes contradicts this model[72,119].

The antigen 85 proteins also bind fibronectin. Crystallographic studies have suggested the presence of a fibronectin binding domain[70], although exact location of the domain remains unclear[120]. It has been demonstrated through binding studies that fibronectin binds to antigen 85 from BCG and that this binding is quite strong[121,122]. Thus, antigen 85 proteins could adhere to cell surfaces and be presented to T lymphocytes. Pasula et al., demonstrated that fibronectin facilitates attachment of *Mtb* to macrophages[123]. Release of antigen 85 could also prevent fibronectin from binding to the bacteria, by attaching the fibronectin to the antigen 85 before it reaches the bacteria. Binding to fibronectin could also encourage fibronectin-mediated phagocytosis.

1.6.3 Ag85, TDM and a TB vaccine

Both TDM and Ag85 have been considered as vaccine candidates. TDM has proved intriguing as a potential vaccine against tuberculosis itself, alone or as an additive to the BCG vaccine in order to improve its efficacy. TDM has been proven as a potent adjuvant, raising a strong response by both the adaptive and humoral immune systems. Just recently it was discovered that a synthetic analog of glycerol monomycolate MMG based on 32 carbons (C_{32}) was found to exhibit comparable levels of immunostimulatory activities to the native MMG[124]. This work casts into doubt the focus on TDM as the best adjuvant and also the focus on lipid length.

Like TDM, the antigen 85 proteins are also important candidates as additives to an *Mtb* vaccine. They are one of the major antigens in the humoral immune response to *Mtb*. As such, vaccination with these proteins either alone, or in conjunction with the BCG vaccine has been proposed as a means to raise immunity against *Mtb*. Protective immunity induced by vaccination with antigen 85 has even been demonstrated in the guinea pig model

system[125,126]. A fusion protein of antigen 85B and ESAT-6 conferred a high level of protection against *Mtb* in the guinea pig aerosol model of infection[127]. ESAT-6 is a 6 kDa early-secreted antigenic target that is a common additive in *Mtb* vaccine development. Preliminary results in mice suggest that vaccination with the Hsp65 protein[128] or with the new BCG-type ESAT vaccine[129] are also possible routes toward an effective TB vaccine. More recently Ag85B has been added to a variety of vaccine candidates, all of which have demonstrated efficacy to some degree[130,131].

These and numerous other results suggest that inclusion the antigen 85 proteins and potentially TDM analogues into a vaccine mixture of killed *Mtb* can raise a stronger immune response than killed bacteria alone. Vaccine trials are currently taking place in which the BCG vaccine has been boosted with a modified vaccinia Ankara that expresses antigen 85 and ESAT-6. It waits to be seen whether these trials will prove effective at preventing *Mtb* infection[103].

1.7 Selection of synthetic diagnostic target

Because of its essentiality, importance in cell wall biosynthesis and, particularly, due to the extracellular incorporation by the Ag85 enzymes, trehalose was selected as a synthetic target of interest. Several other, largely non-mammalian, biologically relevant sugars including arabinose, rhamnose, galactose and fucose[132] were also under consideration as alternative candidates for the development of probes that are specific for tuberculosis. Trehalose (Tre) plays an important role in mycobacterial membrane structure and function as the carbohydrate moiety in trehalose mono and dimycolate (TMM and TDM) glycolipids. The synthesis of TDM is understood to occur extracellularly, making the incorporation of trehalose- probes attractive and chemically tractable, as compounds would not have to transverse the hydrophobic cell wall. The unique immunmodulatory roles of the Ag85 enzymes and TDM, made trehalose of particular interest not only for its potential to label mycobacteria, but also for the potential to develop a better understanding of the roles of Ag85 system in the survival of mycobacteria within the host.

1.7.1 Synthesis of trehalose analogues

A survey the trehalose literature revealed that the synthesis of trehalose derivatives presents a synthetic challenges. Trehalose analogues are synthesized through two basic strategies: 1) modification of the trehalose scaffold[133-136] and 2) coupling of monosaccharides to produce trehalose[137-140]. Syntheses beginning with trehalose require difficult, low yielding desymmetrizing steps to asymmetrically modify the sugar scaffold. Most modifications, beginning from native trehalose have been at 4- and 6-positions because of ease of modification, using benzylidene protecting groups, affording the 6-sulfo and 6-phosphotrehalose[84] as well as 4- and 6-fluorotrehalose[133,134,141] and 6-azido and 6-aminotrehalose (**see Figure 1.9a**). Modifications at the 2- and 3- positions have been limited to the deoxy sugars. Bertozzi and coworkers have made the 3-monodeoxy- and 2-monodeoxy-trehalose analogues and the corresponding 3,3'- and 2,2'-dideoxy analogues utilizing selective protection and radical deoxygenation[135]. Although synthetically less tractable, fluorination at the 2- and 5-positions is potentially more desirable for PET probes and inhibitors because these analogs should mimic the oxocarbenium transition state[91].

Syntheses beginning with glucose or other monosaccharides require that two stereocenters must be set simultaneously, in the relatively unfavorable α configuration. The intramolecular aglycone delivery (IAD) tethered glycosylation strategy developed by Bertozzi et al.[140] offers the advantage of exclusively yielding the α,α linkage, but is challenged by difficult synthesis of the building blocks **24** and **25** and inconsistent reaction yields. The IAD glycosylation was employed to partially synthesize sulfolipid-1[137], capitalizing on the unique asymmetric protection of the 2-hydroxyl (**Figure 1.9b**). Glycosylation of ketosides **28**[138,142-144] capitalizes on the inherent tendency for the ketoside sugars to form the α,α-linkage, but incorporates an extra, bulky group at the anomeric position (**see Figure 1.9c**). Dehydrative glycosylations[139] afford alpha selectivity, but result in the synthesis of the 2',2'-dideoxy analogues, precluding further modification of the 2-position. Chemoenzymatic strategies utilizing the trehalose-6-phosphate synthase (OtsA) take advantage of the enzymes' ability to form exclusively the α,α stereochemistry[145], but as yet have not been demonstrated to tolerate unnatural donor and acceptor sugars. Trehalose and ^{14}C-trehalose can be purified from yeast, which produce vast quantities of trehalose under stressful, non-proliferating growth conditions in the presence of glucose or ^{14}C-glucose[146]. Incorporation of fluorosugars by yeast is not well tolerated[147]. Fluorination of trehalose with ^{18}F has also never been achieved;

however ^{18}FDG is readily available[148]. In summary, the synthesis, of trehalose analogues, particularly ^{18}F-trehalose presents a novel synthetic challenge.

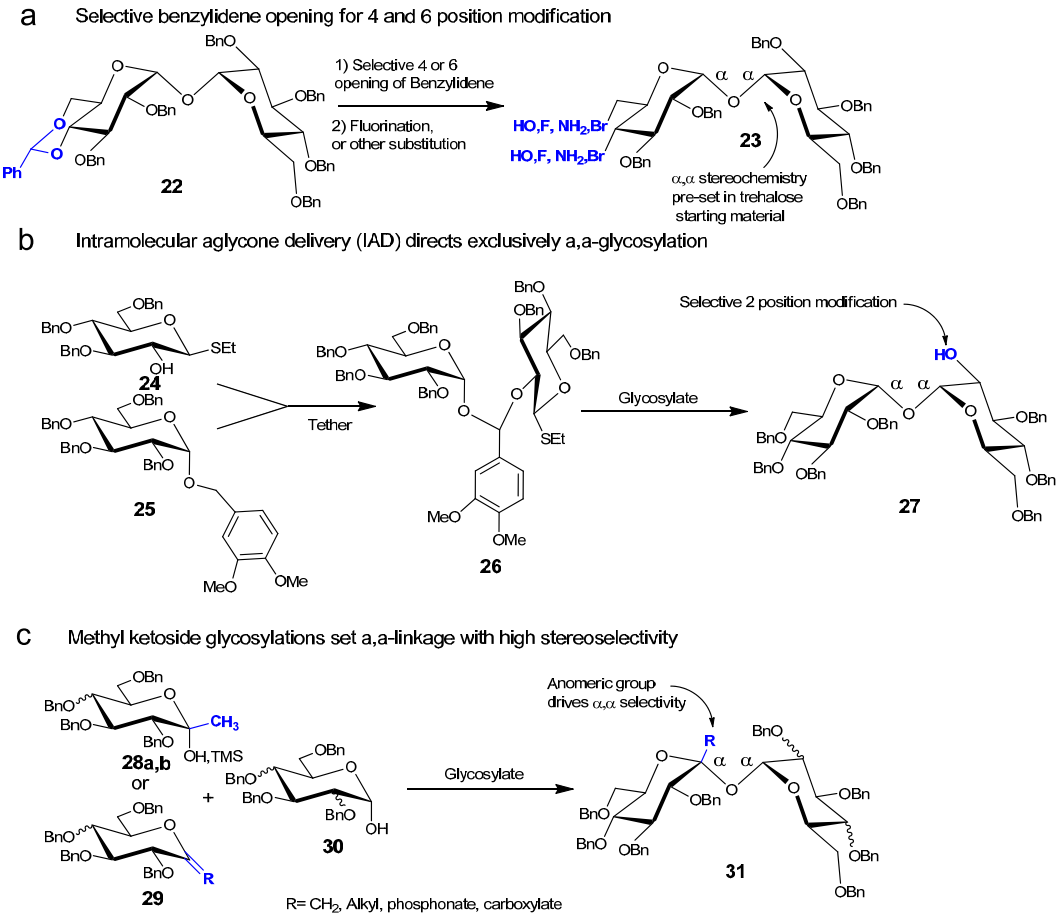


Figure 1.9. Selected examples of approaches to synthesis of trehalose analogues. (a) Benzylidene protection affords selective modification of 4 and 6 position of trehalose[133,134,141] (b) Intramolecular aglycone delivery (IAD) employs a tethering strategy to yield exclusively *α,α* stereochemistry and free C-2 hydroxyl[140]. (c) Methyl ketoside or exoglycal donors can be stereoselectively glycosylated to yield mainly the *α,α* trehalose product R=alkyl, phosphonate, carboxylate[138,142-144] .

1.8 Uptake of trehalose probes into *Mtb in vitro* and *in vivo*

The uptake of trehalose and/or trehalose analogues has not been fully evaluated in mycobacteria. *Mtb* infects and is able to replicate and persist within macrophage cells. Therefore, in these studies, trehalose and trehalose analogues not only have to be taken up by the mycobacterium but also must transverse the macrophage membrane. There is some precedent that extracellular trehalose should be taken up by mycobacteria[49,149]. While this thesis was ongoing a trehalose transporter in mycobacteria was reported, further substantiating that trehalose can be taken up by mycobacteria *in vitro*[55]. Studies of uptake of

tritiated mannose and fucose and radiolabeled peptidoglycan, do suggest that uptake and incorporation is possible both *in vitro* and in macrophages[150]. Trafficking of glycolipids has also been assessed via markers attached via hydrazide linkage[30]. Pinocytosis has been suggested as a possible mechanism for uptake but before developing trehalose analogs. The full scope of the uptake and incorporation of trehalose into *Mtb* remains to be assessed.

1.8.1 Small molecule labeling of mycobacteria

Prior to the publication of the manuscript detailing much of the work reported in this thesis[151], to our knowledge no small molecule, TB-selective probes had been reported. Previous imaging work with *Mtb* in infected macrophages has utilized several strategies[30,152] but none allow for non-toxic imaging of bacteria *in vivo*. Bacteria can be labeled with Texas-Red through oxidation with sodium periodate prior to infection. However, these harsh labeling conditions are nonselective and would be expected to oxidize carbohydrates that may be structurally important as well as peripherally associated glycolipids that may be important for specific virulence attributes[30]. Fluorescently-labeled vancomycin has been elegantly used to track the cell division patterns of *M. smegmatis* and *M. bovis* BCG[152] but its toxicity towards *Mtb*[153], may affect experimental results. Antibodies can be utilized to label *Mtb* upon fixation, but are often troubled with high background and inconsistent labeling[154], and are not useful for live imaging. Staining with auramine and rhodamine, the gold standard for labeling mycobacteria in fixed cells, tissue and sputum, only labels a fraction of the bacteria found in preparations[155]. β -lactamase probes have been elegantly employed *in vivo*, but may also label other β-lactamase expressing bacterium[156]. The absence of TB selective probes reveals a need within the tuberculosis community for novel ways to easily and reproducibly label mycobacteria within samples for both diagnostic and experimental purposes.

1.8.2 Uptake of trehalose into animals

Animal models are essential for the assessment of the uptake, metabolism and tissue distribution of any mycobacterial drug or probe. A variety of animal models of mycobacterial infection are available, including mice, guinea pigs, rabbits and cynomolgus macaque monkeys. Each of these models provides its own advantages and disadvantages. The mouse model is convenient and easily contained. But, murine tuberculosis is distinct from the human infection in that the *Mtb* produces a chronic, rather than latent, infection in mice. The granulomas in mice are smaller, more disorganized and much less well formed than in humans[157]. Guinea pigs[158] and rabbits[159] offer more realistic animal models of disease,

featuring organized granulomas, caseous and consolidated lesions and necrosis. Rabbits are only susceptible to the phenolic glycolipid (PGL) positive clades of the Beijing strain[159]. Macaques offer a manifestation of the TB with comparably complex lesion formation to humans, but the monkey studies are costly and require specialized primate facilities[160].

1.9 Positron emission tomography and computed tomography (PET/CT)

PET (positron emission tomography) and CT (computed tomography) are two, state of the art, imaging techniques that are widely used in other research areas, but, have not received much attention in the field of tuberculosis research. PET is a technique that acquires images from the emission of positrons and is particularly useful in characterizing the metabolic biochemical changes, often those associated with cancer. CT utilizes x-rays to provide a structural image. Used concurrently, PET and CT provide a powerful tool to image both structure and function of a disease. ^{18}F-2-fluoro-deoxy-D-glucose (FDG) is a widely employed PET probe that provides a readout for regions of increased metabolic activity, often those associated with inflammation[161]. Many different PET isotope labels could be chosen, including ^{18}F, ^{14}C, ^{76}Br, $^{123/124}$I. Bromine and Iodine both release higher levels of radiation and have longer half-lives (1 and 4 days respectively) as compared with ^{11}C and ^{18}F. The reasonable half-life of just under two hours for ^{18}F ,compared with 20 minutes for ^{11}C, is one of the reasons that ^{18}F is becoming the most widely employed isotope in radiochemistry and nuclear medicine. As an additional benefit, when used with carbohydrates, the electron withdrawing nature of the ^{18}F allows the fluorine to serve two purposes, both as a radiolabel and as a potential mechanistic trap [91].

At the inception of this research, no studies had been published on the use of a combined PET/CT system with TB and only one report has been published since this research commenced[162]. CT has been widely used to confirm successful treatment, but has typically not been utilized to monitor the progression of the lesions during treatment[163]. The Barry lab has begun to work with ^{18}F-2-deoxy-2-fluoroglucose (^{18}FDG) in a PET/CT system to monitor the progress of treatment with normal combined treatment with isoniazid, rifampicin, pyrazinamide and ethambutol (HRZE) and with the addition of metronidazole (MTZ) to the regimen. This preliminary work has demonstrated that the PET/CT system provides a very detailed, accurate map of the anatomical and functional activity of lesions within the lung. This work has also supported that PET/CT can be employed to evaluate the efficacy of novel chemotherapeutics both in animals and humans. One of the challenges confronting the FDG

project is that the glucose is not specific for mycobacteria. However, what FDG does suggest is that the combined CT/PET system with a probe that is specific for mycobacterium should provide a readout of both the structural and metabolic states of the tuberculosis.

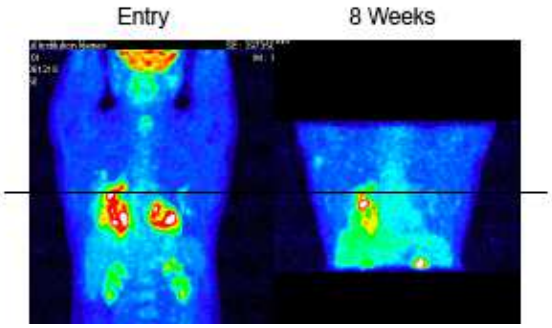


Figure 1.10. PET imaging of human patient under treatment for tuberculosis was conducted utilizing ^{18}F-2-deoxy-2-fluoroglucose (FDG). Scans before and after 8 week treatment course reveals a decrease in the size and radioactivity of FDG positive lesions, suggesting that the treatment was effective[164].

1.10 Conclusions

Tuberculosis is a leading killer of people worldwide; therefore, development of new treatments and diagnostics for TB is of the utmost importance. At the beginning of this dissertation, no current small molecules offered specificity for TB *in vivo*. Trehalose, a disaccharide found in mycobacterial membranes, plays an important role in the survival and persistence of the mycobacteria worldwide. Analogues of trehalose offer a unique synthetic challenge. Modified trehalose molecules may be incorporated into the cell envelope and/or inhibit of membrane biosynthesis, affording a readout of cell wall synthesis as well as inhibition of bacterial growth. ^{18}F labeled trehalose potentially will target the mycobacterium and allow for selective PET/CT imaging and may find use in may find use in PET studies in mice, rabbits and non-human primates. Simultaneously, radiolabeled or fluorescently trehalose could be used to better understand the roles of trehalose containing glycolipids, particularly the substrates of antigen 85 (trehalose mono and dimycolate, TMM and TDM) in the mechanism of macrophage infection, immunosuppression as well as granuloma formation[165]. The full synthetic pathway of TDM has yet to be characterized and probes, such as biotinylated-trehalose, may find use in dissection of this essential pathway. A better understanding of the synthesis of TDM and the functions of the Ag85 enzymes is vital to the rational design of trehalose-based probes. In summary, trehalose analogues may find use as TB specific probes that could help fill the TB selective labeling and diagnostic void.

1.11 Thesis Objectives

The general goal of this thesis is to address the need for improved TB diagnostics through the development of novel methods to selectively label and monitor *Mtb*. At the inception of this project it was hypothesized that analogues of trehalose would be incorporated into the cell wall of *Mtb*, through transesterification by the Ag85 enzymes, allowing for TB-selective labeling. Chemistry and biology are described that validate and test this hypothesis. Chapter 2 details the synthesis of a library of trehalose analogues, featuring fluorescent, ^{14}C- and ^{18}F-labels. This synthetic effort required the development of ketoside glycosylations as a means to stereoselectively synthesize modified trehalose compounds. Chapter 3 reveals a novel, high-throughput Ag85 enzyme assay that demonstrates the broad substrate specificity of the Ag85 enzymes, suggesting that modified trehalose will be processed by these enzymes *in vivo*. In Chapter 4 and 5 the scope of the *in vitro* and *in vivo* labeling of mycobacteria by trehalose and analogues is investigated, revealing that a fluorescent trehalose (FITC-Tre) is the first small molecule to selectively label mycobacteria in cell culture. Chapter 6 considers the differential functions of the three isoforms of Ag85, A, B and C. An improved understanding of the Ag85 enzymes is required for further development of the trehalose labeling methods. This thesis proposes that the Ag85 enzymes and trehalose will offer improved diagnostics and a better understanding of *Mtb* pathogenesis.

1.12 References

1. WHO. Global Tuberculosis Control Report 2008. (2008).
2. Young, D.B. Confronting the scientific obstacles to global control of tuberculosis. *The Journal of Clinical Investigation* **118**, 1255-1265 (2008).
3. Harries, A.D. & Dye, C. Tuberculosis. *Ann. Trop. Med. Parasitol.* **100**, 415-31 (2006).
4. Dye, C. Global epidemiology of tuberculosis. *Lancet* **367**, 938-40 (2006).
5. Corbett, E.L. et al. The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Arch. Intern. Med.* **163**, 1009-21 (2003).
6. Daffé, M. & Etienne, G. The capsule of *Mycobacterium tuberculosis* and its implications for pathogenicity. *Tuber. Lung Dis.* **79**, 153-69 (1999).
7. Koul, A., Arnoult, E., Lounis, N., Guillemont, J. & Andries, K. The challenge of new drug discovery for tuberculosis. *Nature* **469**, 483-490.
8. Thaiss, C.A. & Kaufmann, S.H. Toward novel vaccines against tuberculosis: current hopes and obstacles. *Yale J. Biol. Med.* **83**, 209-15.
9. Iseman, M.D., Cohn, D.L. & Sbarbaro, J.A. Directly Observed Treatment of Tuberculosis -- We Can't Afford Not to Try It. *N. Engl. J. Med.* **328**, 576-578 (1993).
10. Minnikin, D.E. & Goodfellow, M. Lipid composition in the classification and identification of acid-fast bacteria. *Soc. Appl. Bacteriol. Symp. Ser.* **8**, 189-256 (1980).
11. Daffé, M. & Draper, P. The envelope layers of mycobacteria with reference to their pathogenicity. *Adv. Microb. Physiol.* **39**, 131-203 (1998).
12. Brennan, P.J. & Nikaido, H. The envelope of mycobacteria. *Annu. Rev. Biochem.* **64**, 29-63 (1995).
13. Barry, C.E. & Mdluli, K. Drug sensitivity and environmental adaptation of myocobacterial cell wall components. *Trends Microbiol.* **4**, 275-281 (1996).

14. Puech, V. et al. Structure of the cell envelope of corynebacteria: importance of the non-covalently bound lipids in the formation of the cell wall permeability barrier and fracture plane. *Microbiology* **147**, 1365-1382 (2001).
15. Mineda, T., Ohara, N., Yukitake, H. & Yamada, T. The ribosomes contents of mycobacteria. *New Microbiol.* **21**, 1-7 (1998).
16. Paul, T.R. & Beveridge, T.J. Reevaluation of envelope profiles and cytoplasmic ultrastructure of mycobacteria processed by conventional embedding and freeze-substitution protocols. *J. Bacteriol.* **174**, 6508-6517 (1992).
17. Rastogi, N., Hellio, R. & David, H.L. A new insight into the mycobacterial cell envelope architecture by the localization of antigens in ultrathin sections. *Zentralbl Bakteriol* **275**, 287-302 (1991).
18. Liu, J., Rosenberg, E.Y. & Nikaido, H. Fluidity of the lipid domain of cell wall from *Mycobacterium chelonae*. *Proceedings of the National Academy of Sciences* **92**, 11254-11258 (1995).
19. Tonge, P.J. Another brick in the wall. *Nat. Struct. Mol. Biol.* **7**, 94-96 (2000).
20. Engelhardt, H., Heinz, C. & Niederweis, M. A tetrameric porin limits the cell wall permeability of *Mycobacterium smegmatis*. *J. Biol. Chem.*, 37567-37572 (2002).
21. Faller, M., Niederweis, M. & Schulz, G.E. The Structure of a Mycobacterial Outer-Membrane Channel. *Science* **303**, 1189-1192 (2004).
22. Mahfoud, M., Sukumaran, S., Hulsmann, P., Grieger, K. & Niederweis, M. Topology of the porin MspA in the outer membrane of *Mycobacterium smegmatis*. *J. Biol. Chem.* **281**, 5908-15 (2006).
23. Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J.M. & Engelhardt, H. Disclosure of the mycobacterial outer membrane: Cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proceedings of the National Academy of Sciences* **105**, 3963-3967 (2008).
24. Zuber, B. et al. Direct Visualization of the Outer Membrane of Mycobacteria and Corynebacteria in Their Native State. *J. Bacteriol.* **190**, 5672-5680 (2008).
25. Daffé, M. & Reyrat, J.-M. *The mycobacterial cell envelope*, (ASM Press, Washington, 2008).
26. Slayden, R.A. & Barry, C.E. Analysis of the Lipids of *Mycobacterium tuberculosis*. Vol. 54 229-245 (2001).
27. Ohta, M., Pan, Y.T., Laine, R.A. & Elbein, A.D. Trehalose-based oligosaccharides isolated from the cytoplasm of *Mycobacterium smegmatis*. Relation to trehalose-based oligosaccharides attached to lipid. *Eur. J. Biochem.* **269**, 3142-9 (2002).
28. Fujita, Y., Naka, T., McNeil, M.R. & Yano, I. Intact molecular characterization of cord factor (trehalose 6,6'-dimycolate) from nine species of mycobacteria by MALDI-TOF mass spectrometry. *Microbiology* **151**, 3403-3416 (2005).
29. Fujita, Y., Doi, T., Sato, K. & Yano, I. Diverse humoral immune responses and changes in IgG antibody levels against mycobacterial lipid antigens in active tuberculosis. *Microbiology* **151**, 2065-74 (2005).
30. Beatty, W.L. et al. Trafficking and release of mycobacterial lipids from infected macrophages. *Traffic* **1**, 235-47 (2000).
31. Ortalo-Magne, A. et al. Identification of the surface-exposed lipids on the cell envelopes of *Mycobacterium tuberculosis* and other mycobacterial species. *J. Bacteriol.* **178**, 456-61 (1996).
32. Noll, H. The chemistry of some native constituents of the purified wax of *Mycobacterium tuberculosis*. *J. Biol. Chem.* **224**, 149-164 (1957).
33. Noll, H. & Bloch, H. Studies on the chemistry of the cord factor of *Mycobacterium tuberculosis*. *J. Biol. Chem.* **214**, 251-265 (1955).
34. Noll, H., Bloch, H., Asselineau, J. & Lederer, E. The chemical structure of the cord factor of *Mycobacterium tuberculosis*. *Biochim. Biophys. Acta* **20**, 299-309 (1956).
35. Toriyama, S., Ikuya, Y., Masui, M., Masamichi, K. & Kusunose, E. Separation of C50-60 and C70-80 mycolic acid molecular species and their changes by growth temperatures in *Mycobacterium phlei*. *FEBS Lett.* **95**, 111-115 (1978).
36. Toubiana, R., Ribi, E., McLaughlin, C. & Strain, S.M. The effect of synthetic and naturally occurring trehalose fatty acid esters in tumor regression. *Cancer Immunol. Immunother.* **2**, 189-193 (1977).
37. Liav, A. & Goren, M.B. An improved synthesis of 6-O--mycoloyl- and 6-O-corynomycoloyl-α,α-trehalose with observations on the permethylation analysis of trehalose glycolipids. *Carbohydr. Res.* **155**, 229-235 (1986).
38. Numata, F., Ishida, H., Nishimura, K., Sekikawa, I. & Azuma, I. Syntheses of Trehalose Monomycolate and Related Compounds, and Their Lethal Toxicity and Adjuvant Activity. *J. Carbohydr. Chem.* **5**, 127 - 138 (1986).
39. Kilburn, J.J.O., Takayama, K.K. & Armstrong, E.E.L. Synthesis of trehalose dimycolate (cord factor) by a cell-free system of *Mycobacterium smegmatis*. *Biochem. Biophys. Res. Commun.* **108**, 132-9 (1982).

40. Laszlo, A., Baer, H.H., Goren, M.B., Handzel, V. & Papa, F. Screening of synthetic trehalose 6,6'-diesters and trehalose 6-monoesters as potential immunoreactants for the serodiagnosis of tuberculosis. *Res. Microbiol.* **145**, 563-572 (1994).
41. Nishizawa, M. et al. Effective Synthesis of Four Isomeric Trehalose Dicorynomycolates (TDCMs) and Their Immunoadjuvant Activities. *Synlett* **1996**, 452-454 (1996).
42. Watanabe, R. et al. Inhibitory effect of trehalose dimycolate (TDM) and its stereoisometric derivatives, trehalose dicorynomycolates (TDCMs), with low toxicity on lung metastasis of tumour cells in mice. *Vaccine* **17**, 1484-92 (1999).
43. Jenkins, I.D. & Goren, M.B. Improved synthesis of cord factor analogues via the mitsunobu reaction. *Chem. Phys. Lipids* **41**, 225-235 (1986).
44. Chapleur, Y.C., B. and Toubiana, R. Le BOP reagent and imidazole for selective O-acylation of trehalose. *J. Chem. Soc., Perkin Trans. 1*, 1940 - 1943 (1980).
45. Liav, A., Flowers, H.M. & Goren, M.B. Cord-factor analogs: synthesis of 6,6'-di-O-mycoloyl- and -corynomycoloyl-(α-galactopyranosyl α-galactopyranoside). *Carbohydr. Res.* **133**, 53-58 (1984).
46. Retzinger, G., Meredith, S., Takayama, K., Hunter, R. & Kezdy, F. The role of surface in the biological activities of trehalose 6,6'- dimycolate. Surface properties and development of a model system. *J. Biol. Chem.* **256**, 8208-8216 (1981).
47. Hunter, R.L., Venkataprasad, N. & Olsen, M.R. The role of trehalose dimycolate (cord factor) on morphology of virulent *M. tuberculosis in vitro*. *Tuberculosis (Edinb.)* **86**, 349-56 (2006).
48. Murphy, H.N. et al. The OtsAB pathway is essential for trehalose biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **280**, 14524-9 (2005).
49. Woodruff, P.J. et al. Trehalose is required for growth of *Mycobacterium smegmatis*. *J. Biol. Chem.* **279**, 28835-43 (2004).
50. Pan, Y.T., Carroll, J.D. & Elbein, A.D. Trehalose-phosphate synthase of *Mycobacterium tuberculosis*. Cloning, expression and properties of the recombinant enzyme. *Eur. J. Biochem.* **269**, 6091-100 (2002).
51. Pan, Y.T. et al. Trehalose synthase of *Mycobacterium smegmatis*: purification, cloning, expression, and properties of the enzyme. *Eur. J. Biochem.* **271**, 4259-69 (2004).
52. Klutts, S. et al. Purification, cloning, expression, and properties of mycobacterial trehalose-phosphate phosphatase. *J. Biol. Chem.* **278**, 2093-100 (2003).
53. Gibson, R.P., Turkenburg, J.P., Charnock, S.J., Lloyd, R. & Davies, G.J. Insights into Trehalose Synthesis Provided by the Structure of the Retaining Glucosyltransferase OtsA. *Chem. Biol.* **9**, 1337-1346 (2002).
54. Carroll, J.D., Pastuszak, I., Edavana, V.K., Pan, Y.T. & Elbein, A.D. A novel trehalase from *Mycobacterium smegmatis* - purification, properties, requirements. *FEBS J.* **274**, 1701-14 (2007).
55. Kalscheuer, R., Weinrick, B., Veeraraghavan, U., Besra, G.S. & Jacobs, W.R., Jr. Trehalose-recycling ABC transporter LpqY-SugA-SugB-SugC is essential for virulence of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 21761-6 (2010).
56. Bhatt, A., Molle, V., Besra, G.S., Jacobs, W.R., Jr. & Kremer, L. 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 (2007).
57. Takayama, K., Wang, C. & Besra, G.S. Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin. Microbiol. Rev.* **18**, 81-101 (2005).
58. Besra, G.S. et al. Identification of the apparent carrier in mycolic acid synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12735-12739 (1994).
59. Wang, C., Hayes, B., Vestling, M.M. & Takayama, K. Transposome mutagenesis of an integral membrane transporter in *Corynebacterium matruchotii*. *Biochem. Biophys. Res. Commun.* **340**, 953-960 (2006).
60. Domenech, P., Kobayashi, H., LeVier, K., Walker, G.C. & Barry, C.E., III. BacA, an ABC Transporter Involved in Maintenance of Chronic Murine Infections with *Mycobacterium tuberculosis*. *J. Bacteriol.* **191**, 477-485 (2009).
61. Puech, V., Bayan, N., Salim, K., Leblon, G. & Daffé, M. Characterization of the *in vivo* acceptors of the mycoloyl residues transferred by the corynebacterial PS1 and the related mycobacterial antigens 85. *Mol. Microbiol.* **35**, 1026-41 (2000).
62. Puech, V. et al. Evidence for a partial redundancy of the fibronectin-binding proteins for the transfer of mycoloyl residues onto the cell wall arabinogalactan termini of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **44**, 1109-1122 (2002).
63. Jackson, M. et al. 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 (1999).

64. Fukui, Y., Hirai, T., Uchida, T. & Yoneda, M. Extracellular proteins of tubercle bacilli. IV. Alpha and beta antigens as major extracellular protein products and as cellular components of a strain (H37Rv) of *Mycobacterium tuberculosis*. *Biken J.* **8**, 189-99 (1965).
65. Harth, G., Lee, B.Y., Wang, J., Clemens, D.L. & Horwitz, M.A. Novel insights into the genetics, biochemistry, and immunocytochemistry of the 30-kilodalton major extracellular protein of *Mycobacterium tuberculosis*. *Infect. Immun.* **64**, 3038-47 (1996).
66. De Sousa-D'Auria, C. et al. New insights into the biogenesis of the cell envelope of corynebacteria: identification and functional characterization of five new mycoloyltransferase genes in *Corynebacterium glutamicum*. *FEMS Microbiol. Lett.* **224**, 35-44 (2003).
67. Rafidinarivo, E. et al. Trafficking pathways of mycolic acids: structures, origin, mechanism of formation, and storage form of mycobacteric acids. *J. Lipid Res.* **50**, 477-490 (2009).
68. Gebhardt, H. et al. The key role of the mycolic acid content in the functionality of the cell wall permeability barrier in *Corynebacterineae*. *Microbiology* **153**, 1424-34 (2007).
69. Kacem, R. et al. Importance of mycoloyltransferases on the physiology of *Corynebacterium glutamicum*. *Microbiology* **150**, 73-84 (2004).
70. Ronning, D.R. et al. Crystal structure of the secreted form of antigen 85C reveals potential targets for mycobacterial drugs and vaccines. *Nat. Struct. Biol.* **7**, 141-6 (2000).
71. Anderson, D.H., Harth, G., Horwitz, M.A. & Eisenberg, D. An interfacial mechanism and a class of inhibitors inferred from two crystal structures of the *Mycobacterium tuberculosis* 30 kDa major secretory protein (Antigen 85B), a mycolyl transferase. *J. Mol. Biol.* **307**, 671-81 (2001).
72. Ronning, D.R., Vissa, V., Besra, G.S., Belisle, J.T. & Sacchettini, J.C. *Mycobacterium tuberculosis* antigen 85A and 85C structures confirm binding orientation and conserved substrate specificity. *J. Biol. Chem.* **279**, 36771-7 (2004).
73. Sathyamoorthy, N. & Takayama, K. Purification and characterization of a novel mycolic acid exchange enzyme from *Mycobacterium smegmatis*. *J. Biol. Chem.* **262**, 13417-23 (1987).
74. Belisle, J.T. et al. Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. *Science* **276**, 1420-2 (1997).
75. Boucau, J., Sanki, A.K., Voss, B.J., Sucheck, S.J. & Ronning, D.R. A coupled assay measuring *Mycobacterium tuberculosis* antigen 85C enzymatic activity. *Anal. Biochem.* **385**, 120-127 (2009).
76. Sanki, A.K., Boucau, J., Ronning, D.R. & Sucheck, S.J. Antigen 85C-mediated acyl-transfer between synthetic acyl donors and fragments of the arabinan. *Glycoconj. J.* (2008).
77. Elamin, A.A., Stehr, M., Oehlmann, W. & Singh, M. The mycolyltransferase 85A, a putative drug target of *Mycobacterium tuberculosis*: development of a novel assay and quantification of glycolipid-status of the mycobacterial cell wall. *J. Microbiol. Methods* **79**, 358-63 (2009).
78. Hunter, R.L., Olsen, M.R., Jagannath, C. & Actor, J.K. Multiple roles of cord factor in the pathogenesis of primary, secondary, and cavitary tuberculosis, including a revised description of the pathology of secondary disease. *Ann. Clin. Lab. Sci.* **36**, 371-86 (2006).
79. Saunders, B.M. & Britton, W.J. Life and death in the granuloma: immunopathology of tuberculosis. *Immunol. Cell Biol.* **85**, 103-11 (2007).
80. Nguyen, L., Chinnapapagari, S. & Thompson, C.J. FbpA-Dependent Biosynthesis of Trehalose Dimycolate Is Required for the Intrinsic Multidrug Resistance, Cell Wall Structure, and Colonial Morphology of *Mycobacterium smegmatis*. *J. Bacteriol.* **187**, 6603-6611 (2005).
81. Gomez, J.E. & McKinney, J.D. M. tuberculosis persistence, latency, and drug tolerance. *Tuberculosis* **84**, 29-44 (2004).
82. Kalscheuer, R. et al. Self-poisoning of *Mycobacterium tuberculosis* by targeting GlgE in an α-glucan pathway. *Nat. Chem. Biol.* **6**, 376-384 (2010).
83. Hui, Y. & Chang, C.W. Convenient divergent synthesis of a library of trehalosamine analogues. *Org. Lett.* **4**, 2245-8 (2002).
84. Rose, J.D. et al. Synthesis and biological evaluation of trehalose analogs as potential inhibitors of mycobacterial cell wall biosynthesis. *Carbohydr. Res.* **337**, 105-20 (2002).
85. Wang, J. et al. Synthesis of trehalose-based compounds and their inhibitory activities against *Mycobacterium smegmatis*. *Bioorg. Med. Chem.* **12**, 6397-6413 (2004).
86. Gobec, S. et al. Design, synthesis, biochemical evaluation and antimycobacterial action of phosphonate inhibitors of antigen 85C, a crucial enzyme involved in biosynthesis of the mycobacterial cell wall. *Eur. J. Med. Chem.* **42**, 54-63 (2007).
87. Gobec, S. et al. Phosphonate inhibitors of antigen 85C, a crucial enzyme involved in the biosynthesis of the *Mycobacterium tuberculosis* cell wall. *Bioorg. Med. Chem. Lett.* **14**, 3559-62 (2004).
88. Frlan, R., Gobec, S. & Kikelj, D. Synthesis of ethyl 3-(hydroxyphenoxy)benzyl butylphosphonates as potential antigen 85C inhibitors. *Tetrahedron* **63**, 10698-10708 (2007).

89. Sanki, A.K. et al. Synthesis of methyl 5-S-alkyl-5-thio-D-arabinofuranosides and evaluation of their antimycobacterial activity. *Bioorg. Med. Chem.* **16**, 5672-82 (2008).
90. Lee, S.S. et al. Mechanistic evidence for a front-side, SNi-type reaction in a retaining glycosyltransferase. *Nat. Chem. Biol.* **7**, 631-638.
91. Hartman, M.C.T. & Coward, J.K. Synthesis of 5-Fluoro-N-Acetylglucosamine Glycosides and Pyrophosphates via Epoxide Fluoridolysis: Versatile Reagents for the Study of Glycoconjugate Biochemistry. *J. Am. Chem. Soc.* **124**, 10036-10053 (2002).
92. Takimoto, H. et al. Interferon-gamma independent formation of pulmonary granuloma in mice by injections with trehalose dimycolate (cord factor), lipoarabinomannan and phosphatidylinositol mannosides isolated from *Mycobacterium tuberculosis*. *Clin. Exp. Immunol.* **144**, 134-41 (2006).
93. Benaroudj, N., Lee, D.H. & Goldberg, A.L. Trehalose accumulation during cellular stress protects cells and cellular proteins from damage by oxygen radicals. *J. Biol. Chem.* **276**, 24261-7 (2001).
94. Ozeki, Y. et al. In vivo induction of apoptosis in the thymus by administration of mycobacterial cord factor (trehalose 6,6'-dimycolate). *Infect. Immun.* **65**, 1793-9 (1997).
95. Bekierkunst, A. Acute granulomatous response produced in mice by trehalose-6,6-dimycolate. *J. Bacteriol.* **96**, 958-61 (1968).
96. White, R.G., Jolles, P., Samour, D. & Lederer, E. Correlation of adjuvant activity and chemical structure of wax D fractions of mycobacteria. *Immunology* **7**, 158-71 (1964).
97. Bekierkunst, A. et al. Immune Response to Sheep Red Blood Cells in Mice Pretreated with Mycobacterial Fractions. *Infect. Immun.* **4**, 256-263 (1971).
98. Moore, V.L., Myrvik, Q.N. & Kato, M. Role of cord factor (trehalose-6,6'-dimycolate) in allergic granuloma formation in rabbits. *Infect. Immun.* **6**, 5-8 (1972).
99. Bekierkunst, A. & Yarkoni, E. Granulomatous hypersensitivity to trehalose-6,6'-dimycolate (cord factor) in mice infected with BCG. *Infect. Immun.* **7**, 631-8 (1973).
100. Meyer, T.J., Ribi, E. & Azuma, I. Biologically active components from mycobacterial cell walls. V. Granuloma formation in mouse lungs and guinea pig skin. *Cell. Immunol.* **16**, 11-24 (1975).
101. McLaughlin, C.A., Parker, R., Hadlow, W.J., Toubiana, R. & Ribi, E. Moieties of mycobacterial mycolates required for inducing granulomatous reactions. *Cell. Immunol.* **38**, 14-24 (1978).
102. Guidry, T.V., Hunter, R.L., Jr. & Actor, J.K. CD3+ cells transfer the hypersensitive granulomatous response to mycobacterial glycolipid trehalose 6,6'-dimycolate in mice. *Microbiology* **152**, 3765-75 (2006).
103. Pathan, A.A. et al. Boosting BCG with recombinant modified vaccinia ankara expressing antigen 85A: different boosting intervals and implications for efficacy trials. *PLoS One* **2**, e1052 (2007).
104. Sanchez-Rodriguez, C. et al. An IgG antibody response to the antigen 85 complex is associated with good outcome in Mexican Totonaca Indians with pulmonary tuberculosis. *Int. J. Tuberc. Lung Dis.* **6**, 706-12 (2002).
105. D'Souza, S. et al. Mapping of Murine Th1 Helper T-Cell Epitopes of Mycolyl Transferases Ag85A, Ag85B, and Ag85C from *Mycobacterium tuberculosis*. *Infect. Immun.* **71**, 483-493 (2003).
106. Matsunaga, I. et al. Mycolyltransferase-mediated Glycolipid Exchange in Mycobacteria. *J. Biol. Chem.* **283**, 28835-28841 (2008).
107. Indrigo, J., Hunter, R.L., Jr. & Actor, J.K. Influence of trehalose 6,6'-dimycolate (TDM) during mycobacterial infection of bone marrow macrophages. *Microbiology* **148**, 1991-8 (2002).
108. Indrigo, J., Hunter, R.L., Jr. & Actor, J.K. Cord factor trehalose 6,6'-dimycolate (TDM) mediates trafficking events during mycobacterial infection of murine macrophages. *Microbiology* **149**, 2049-59 (2003).
109. Yamagami, H. et al. Trehalose 6,6'-dimycolate (cord factor) of *Mycobacterium tuberculosis* induces foreign-body- and hypersensitivity-type granulomas in mice. *Infect. Immun.* **69**, 810-5 (2001).
110. Syed, S.S. & Hunter, R.L., Jr. Studies on the toxic effects of quartz and a mycobacterial glycolipid, trehalose 6,6'-dimycolate. *Ann. Clin. Lab. Sci.* **27**, 375-83 (1997).
111. Ishikawa, E. et al. Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle. *The Journal of Experimental Medicine* **206**, 2879-2888 (2009).
112. Baba, T., Natsuhara, Y., Kaneda, K. & Yano, I. Granuloma formation activity and mycolic acid composition of mycobacterial cord factor. *Cell. Mol. Life Sci.* **53**, 227-32 (1997).
113. Lima, V.M. et al. Role of trehalose dimycolate in recruitment of cells and modulation of production of cytokines and NO in tuberculosis. *Infect. Immun.* **69**, 5305-12 (2001).
114. Ulrichs, T. & Kaufmann, S.H. New insights into the function of granulomas in human tuberculosis. *J. Pathol.* **208**, 261-9 (2006).
115. Russell, D.G. Who puts the tubercle in tuberculosis? *Nat. Rev. Microbiol.* **5**, 39-47 (2007).
116. Ryll, R., Kumazawa, Y. & Yano, I. Immunological properties of trehalose dimycolate (cord factor) and other mycolic acid-containing glycolipids-a review. *Microbiol. Immunol.* **45**, 801-11 (2001).

117. Turneer, M. et al. Humoral immune response in human tuberculosis: immunoglobulins G, A, and M directed against the purified P32 protein antigen of *Mycobacterium bovis* bacillus Calmette-Guerin. *J. Clin. Microbiol.* **26**, 1714-1719 (1988).
118. Launois, P. et al. T-cell-epitope mapping of the major secreted mycobacterial antigen Ag85A in tuberculosis and leprosy. *Infect. Immun.* **62**, 3679-3687 (1994).
119. Comas, I. et al. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat. Genet.* **42**, 498-503 (2010).
120. Naito, M., Fukuda, T., Sekiguchi, K. & Yamada, T. The domains of human fibronectin mediating the binding of alpha antigen, the most immunopotent antigen of mycobacteria that induces protective immunity against mycobacterial infection. *Biochem. J.* **347**, 725-731 (2000).
121. Abou-Zeid, C. et al. Characterization of fibronectin-binding antigens released by *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG. *Infect. Immun.* **56**, 3046-3051 (1988).
122. Peake, P., Gooley, A. & Britton, W.J. Mechanism of interaction of the 85B secreted protein of *Mycobacterium bovis* with fibronectin. *Infect. Immun.* **61**, 4828-4834 (1993).
123. Pasula, R., Wisniowski, P. & Martin II, W.J. Fibronectin Facilitates *Mycobacterium tuberculosis* Attachment to Murine Alveolar Macrophages. *Infect. Immun.* **70**, 1287-1292 (2002).
124. Andersen, C.S. et al. A Simple Mycobacterial Monomycolated Glycerol Lipid Has Potent Immunostimulatory Activity. *J. Immunol.* **182**, 424-432 (2009).
125. Horwitz, M.A., Lee, B.W., Dillon, B.J. & Harth, G. Protective immunity against tuberculosis induced by vaccination with major extracellular proteins of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 1530-4 (1995).
126. Baldwin, S.L. et al. Evaluation of New Vaccines in the Mouse and Guinea Pig Model of Tuberculosis. *Infect. Immun.* **66**, 2951-2959 (1998).
127. Olsen, A.W., Williams, A., Okkels, L.M., Hatch, G. & Andersen, P. Protective Effect of a Tuberculosis Subunit Vaccine Based on a Fusion of Antigen 85B and ESAT-6 in the Aerosol Guinea Pig Model. *Infect. Immun.* **72**, 6148-6150 (2004).
128. Lima, K.M. et al. Single dose of a vaccine based on DNA encoding mycobacterial hsp65 protein plus TDM-loaded PLGA microspheres protects mice against a virulent strain of *Mycobacterium tuberculosis*. *Gene Ther.* **10**, 678-85 (2003).
129. Khader, S.A. et al. IL-23 and IL-17 in the establishment of protective pulmonary CD4+ T cell responses after vaccination and during *Mycobacterium tuberculosis* challenge. *Nat. Immunol.* **8**, 369-377 (2007).
130. Lu, D. et al. Pulmonary Immunization Using Antigen 85-B Polymeric Microparticles to Boost Tuberculosis Immunity. *The AAPS Journal* **12**, 338-347.
131. Tang, C. et al. Efficacy of Recombinant Bacille Calmette-Guérin Vaccine Secreting Interleukin-15/Antigen 85B Fusion Protein in Providing Protection against *Mycobacterium tuberculosis*. *J. Infect. Dis.* **197**, 1263-1274 (2008).
132. Brennan, P.J. Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb.)* **83**, 91-7 (2003).
133. Hadfield, A.F., Hough, L. & Richardson, A.C. The synthesis of 6-deoxy-6-fluoro-α,α-trehalose and related analogues. *Carbohydr. Res.* **63**, 51-60 (1978).
134. Hadfield, A.F., Hough, L. & Richardson, A.C. The syntheses of 4,6-dideoxy-4,6-difluoro- and 4-deoxy-4-fluoro-α,α-trehalose. *Carbohydr. Res.* **71**, 95-102 (1979).
135. Lin, F.L., van Halbeek, H. & Bertozzi, C.R. Synthesis of mono- and dideoxygenated α,α-trehalose analogs. *Carbohydr. Res.* **342**, 2014-2030 (2007).
136. Hanessian, S. & Lavallee, P. Synthesis of 6-amino-6-deoxy-α,α-trehalose: a positional isomer of trehalosamine *The Journal of Antibiotics* **25**, 683-684 (1972).
137. Leigh, C.D. & Bertozzi, C.R. Synthetic Studies toward *Mycobacterium tuberculosis* Sulfolipid-I. *J. Org. Chem.* **73**, 1008-1017 (2008).
138. Namme, R., Hideyo, T.M. & Ikegami, T.S. Development of Ketoside-Type Analogues of Trehalose by Using α-Stereoselective O-Glycosidation of Ketose. *Eur. J. Org. Chem.* **2007**, 3758-3764 (2007).
139. Rodriguez, M.A. et al. Synthesis of 2-Iodoglycals, Glycals, and 1,1'-Disaccharides from 2-Deoxy-2-iodopyranoses under Dehydrative Glycosylation Conditions. *The Journal of Organic Chemistry* **72**, 8998-9001 (2007).
140. Pratt, M.R., Leigh, C.D. & Bertozzi, C.R. Formation of 1,1-α,α-glycosidic bonds by intramolecular aglycone delivery. A convergent synthesis of trehalose. *Org. Lett.* **5**, 3185-8 (2003).
141. Defaye, J., Driguez, H., Henrissat, B., Gelas, J. & Bar-Guilloux, E. Asymmetric acetalation of α,α-trehalose: Synthesis of α-galactopyranosyl α-glucopyranoside and 6-deoxy-6-fluoro-α-glucopyranosyl α-glucopyranoside. *Carbohydr. Res.* **63**, 41-49 (1978).

142. Li, X., Ohtake, H., Takahashi, H. & Ikegami, S. A facile synthesis of 1'-C-alkyl-α-disaccharides from 1-C-alkyl-hexopyranoses and methyl 1-C-methyl-hexopyranosides. *Tetrahedron* **57**, 4297-4309 (2001).
143. Li, X., Ohtake, H., Takahashi, H. & Ikegami, S. An efficient synthesis of new 1'-C-methyl-α-O-disaccharides using 1-methylenesugars as the glycosyl donors. *Tetrahedron* **57**, 4283-4295 (2001).
144. Gomez, A.M., Pedregosa, A., Valverde, S. & Lopez, J.C. Stereoselective synthesis of substituted exo-glycals from 1-exo-methylene pyranoses. *Tetrahedron Lett.* **44**, 6111-6116 (2003).
145. Horlacher, R., Peist, R. & Boos, W. Improved method for the preparative synthesis of labeled trehalose of high specific activity by *Escherichia coli*. *Appl. Environ. Microbiol.* **62**, 3861-3 (1996).
146. Stambuk, B.U., Crowe, J.H., Crowe, L.M., Panek, A.D. & de Araujo, P.S. A dependable method for the synthesis of [14C]trehalose. *Anal. Biochem.* **212**, 150-3 (1993).
147. Tran-Dinh, S., Courtois, A., Wietzerbin, J., Bouet, F. & Herve, M. Reciprocal effects of 2-fluoro-2-deoxy-glucose and glucose on their metabolism in *Saccharomyces cerevisiae* studied by multi-nuclear NMR spectroscopy. *Biochimie* **77**, 233-239 (1995).
148. Bida, G.T., Satyamurthy, N. & Barrio, J.R. The Synthesis of 2-^{18}F-2-Fluoro-2-Deoxy-D-Glucose Using Glycals: A Reexamination. *J. Nucl. Med.* **25**, 1327-1334 (1984).
149. Tzvetkov, M., Klopprogge, C., Zelder, O. & Liebl, W. Genetic dissection of trehalose biosynthesis in *Corynebacterium glutamicum*: inactivation of trehalose production leads to impaired growth and an altered cell wall lipid composition. *Microbiology* **149**, 1659-73 (2003).
150. Woodbury, J.L. & Barrow, W.W. Radiolabeling of *Mycobacterium avium* oligosaccharide determinant and use in macrophage studies. *J. Gen. Microbiol.* **135**, 1875-84 (1989).
151. Backus, K.M. et al. Uptake of unnatural trehalose analogs as a reporter for *Mycobacterium tuberculosis*. *Nat. Chem. Biol.* **7**, 228-35 (2011).
152. Thanky, N.R., Young, D.B. & Robertson, B.D. Unusual features of the cell cycle in mycobacteria: polar-restricted growth and the snapping-model of cell division. *Tuberculosis (Edinb.)* **87**, 231-6 (2007).
153. Provvedi, R., Boldrin, F., Falciani, F., Palu, G. & Manganelli, R. Global transcriptional response to vancomycin in *Mycobacterium tuberculosis*. *Microbiology* **155**, 1093-1102 (2009).
154. Glatman-Freedman, A., Martin, J.M., Riska, P.F., Bloom, B.R. & Casadevall, A. Monoclonal antibodies to surface antigens of *Mycobacterium tuberculosis* and their use in a modified enzyme-linked immunosorbent spot assay for detection of mycobacteria. *J. Clin. Microbiol.* **34**, 2795-2802 (1996).
155. Strumpf, I.J., Tsang, A.Y., Schork, M.A. & Weg, J.G. The reliability of gastric smears by auramine-rhodamine staining technique for the diagnosis of tuberculosis. *Am. Rev. Respir. Dis.* **114**, 971-6 (1976).
156. Kong, Y. et al. Imaging tuberculosis with endogenous Î²-lactamase reporter enzyme fluorescence in live mice. *Proceedings of the National Academy of Sciences*.
157. Hunter, R.L., Jagannath, C. & Actor, J.K. Pathology of postprimary tuberculosis in humans and mice: Contradiction of long-held beliefs. *Tuberculosis* **87**, 267-278 (2007).
158. Turner, O.C., Basaraba, R.J. & Orme, I.M. Immunopathogenesis of Pulmonary Granulomas in the Guinea Pig after Infection with *Mycobacterium tuberculosis*. *Infect. Immun.* **71**, 864-871 (2003).
159. Tsenova, L. et al. Virulence of selected *Mycobacterium tuberculosis* clinical isolates in the rabbit model of meningitis is dependent on phenolic glycolipid produced by the bacilli. *J. Infect. Dis.* **192**, 98-106 (2005).
160. Flynn, J.L. et al. Non-human primates: a model for tuberculosis research. *Tuberculosis (Edinb.)* **83**, 116-8 (2003).
161. Pauwels, E.K. et al. FDG accumulation and tumor biology. *Nucl. Med. Biol.* **25**, 317-22 (1998).
162. Sathekge, M., Maes, A., Kgomo, M., Stoltz, A. & Van de Wiele, C. Use of 18F-FDG PET to predict response to first-line tuberculostatics in HIV-associated tuberculosis. *J. Nucl. Med.* **52**, 880-5.
163. Curvo-Semedo, L., Teixeira, L. & Caseiro-Alves, F. Tuberculosis of the chest. *Eur. J. Radiol.* **55**, 158-172 (2005).
164. Barry, C. *Unpublished Results* (2008).
165. Hamasaki, N. et al. In vivo administration of mycobacterial cord factor (Trehalose 6, 6'-dimycolate) can induce lung and liver granulomas and thymic atrophy in rabbits. *Infect. Immun.* **68**, 3704-9 (2000).

2 Chapter 2: Synthesis of trehalose probes

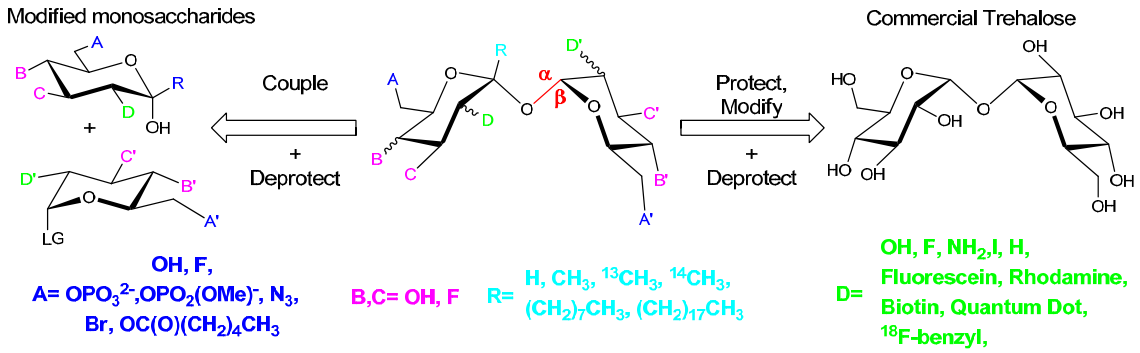
2.1 Introduction

Chapter 2 details the synthesis of a library of trehalose analogues and the development of a robust method for the synthesis of modified trehalose, with α,α stereoselectivity, through methyl ketoside glycosylations. Retrosynthetic strategies for trehalose synthesis were first developed with consideration given to syntheses that began from trehalose as well as from monosaccharide building blocks. Then several alternative synthetic approaches (radical bromination and enzymatic) were briefly explored and proved successful to varying degrees. Confronted with challenges in both of these initial approaches, the ketoside trehalose glycosylation was developed as a means to set the α,α,1,1-linkage stereoselectively and allow for late stage installation of fluorescent and radioactive tags. After initial difficulties with the glycosylation chemistry and deprotection steps, these reactions were optimized and allowed for the synthesis of a library of trehalose analogues, including, trehalose with fluorine, fluorescein, quantum dots, biotin and even ^{18}F. This chemistry is a robust method to rapidly synthesize trehalose derivatives.

2.2 Results

2.2.1 Retrosynthetic analysis

Trehalose can be synthesized through two complementary strategies (**Scheme 2.1**). The first is through direct manipulation of the trehalose scaffold, utilizing asymmetric protection strategies to desymmetrize the starting scaffold. The second is a glycosylation strategy, where two monosaccharides are coupled together to form the α,α-1,1-linkage. There are strengths and weaknesses to both synthetic approaches. The advantage of beginning with trehalose is that the difficult α,α-1,1-linkage is preset in the starting material. Direct modification of the trehalose scaffold is challenged by the innate insolubility of trehalose in most organic solvents and the inherent difficulties with introducing asymmetry in the symmetric trehalose molecule. In contrast, starting with monosaccharides, allows for greater diversification of starting material, better solubility and circumvents the desymmetrization step. However, setting two glycoside anomeric centers in the relatively unfavourable α-configuration through glycosylation chemistry can prove challenging. Both of these approaches were employed and the corresponding hurdles circumvented in the synthesis of the trehalose library.



Scheme 2.1 Retrosynthetic strategies for synthesizing trehalose derivatives either from monosaccharides or through direct modifications to the trehalose scaffold. LG-leaving group.

2.2.2 Open active site of Antigen 85 (Ag85) suggests tolerance to modifications

Early clues to the Ag85 enzymes tolerance of sugar modifications were found in the crystal structures, co-crystallized with trehalose and other ligands[1-3]. The crystal structure of the Ag85B enzyme bound to trehalose (PDB 1F0P) was viewed PYMOL[4] (**Figure 2.1**). The structure suggested an open active site that would tolerate modifications to the sugar scaffold, prompting the synthesis of a library of trehalose analogues, which would have been unnecessary if the Ag85 enzymes had no tolerance towards modifications to the trehalose scaffold. A full survey of the Ag85 enzyme promiscuity has been detailed in **Chapter 3**.

The structural analysis revealed approximately 3.9 Å of free space in the binding pocket between C-1, and the neighboring residues, suggesting that the anomeric position could be modified. In opposite orientation, there is 6.1 Å between C-1' and the amine of the adjacent tryptophan residue. This data suggested that there would be enough space in the binding pocket to accommodate the methyl modification on trehalose. Considerable space was also observed between the nearest amino acids and both C-4 (4.38 Å) and C-2 (4.54 Å). The C-2 and C-4 hydroxyls were found to project outwards into the solvent, suggesting that these positions could be appended with bulky groups without a dramatic loss of activity. The open active site revealed by the Ag85 crystal structure suggests that these enzymes could tolerate a variety of modifications to the sugar scaffold.

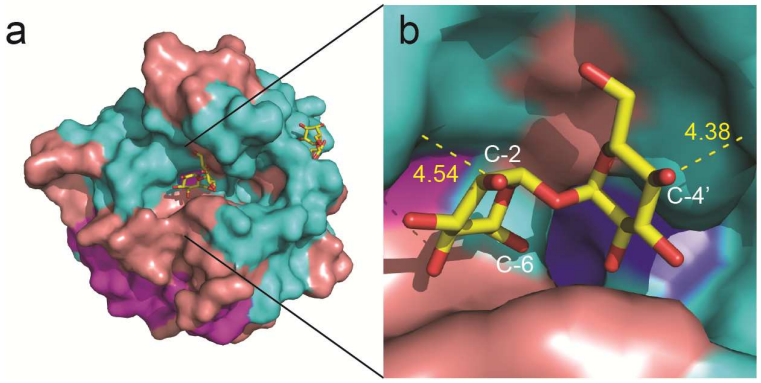


Figure 2.1. The X-ray crystal structure of Ag85B from *Mycobacterium tuberculosis* (PDB 1F0P) suggests significant modification to trehalose substrates may be tolerated without altering binding contacts. (a) Ag85B is bound to trehalose (yellow) (PDB 1F0P)[3]. (b) The active site of Ag85B bound to trehalose. C-2 and C-4' hydroxyls both point out of the active site. C-6 points into the hydrophobic tunnel (blue) that is thought to accommodate the long fatty acyl chains of the mycolic acids. Measurements to nearest residue are indicated in angstroms[3]. Figure generated in Pymol[4].

2.2.3 Overview of the trehalose library

The trehalose library (**Figure 2.2**) shows all important trehalose derivatives that have been taken forward for further evaluation, described in **Chapters 3-6**. This library achieved modifications at every position of the trehalose scaffold, including charged bulky groups as well as fluorophors and biotin for *in vivo* labeling and incorporation of fluorine at multiple positions in order to facilitate the development of ^{18}F positron emission tomography (PET) probes. With this wide variety of modifications, a full evaluation of the scope and limitations of trehalose labeling both *in vitro* and *in vivo* will be possible.

It is important to note that this project was the result of a collaborative grant, funded by the Bill and Melinda Gates Foundation TB Drug Accelerator Program for the development of novel methods to image tuberculosis. As such, this work was the result of the synthesis of several chemists in addition to the author. Throughout this section I will focus mainly on my own body of work and will briefly describe contributions by co-workers where it is relevant or adds to the completeness of the work. These additions will be indicated by footnotes. For complete synthesis of all compounds in **Figure 2.2** please refer to the publication of this work[5].

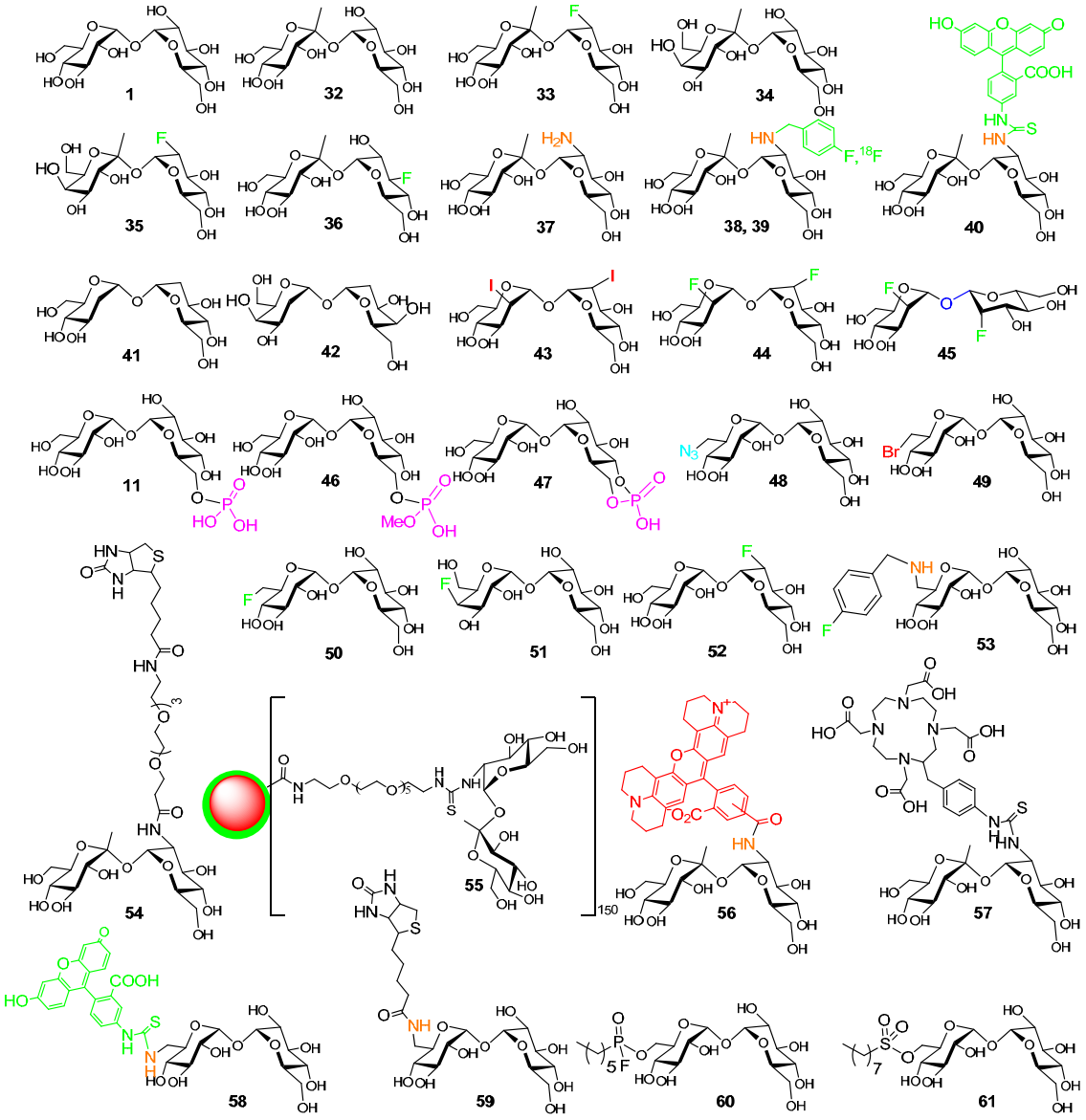


Figure 2.2. The trehalose library. This panel reflects all significant compounds discussed for the remainder of this thesis. Compounds 11, 32-40, 54-57, 60 and 61 were synthesized by the author. Compounds 41-45, 47-53, 58 and 59 were synthesized by coworkers at the University of Oxford[i]. Colors highlight the diverse modifications incorporated into the trehalose scaffold.

2.2.4 Synthetic strategies

The following sections present a brief discussion of each of the synthetic approaches attempted for the synthesis of fluorinated (particularly ^{18}F) trehalose and trehalose with unnatural stereochemistry and compounds appended by various labeling tags. Initially the

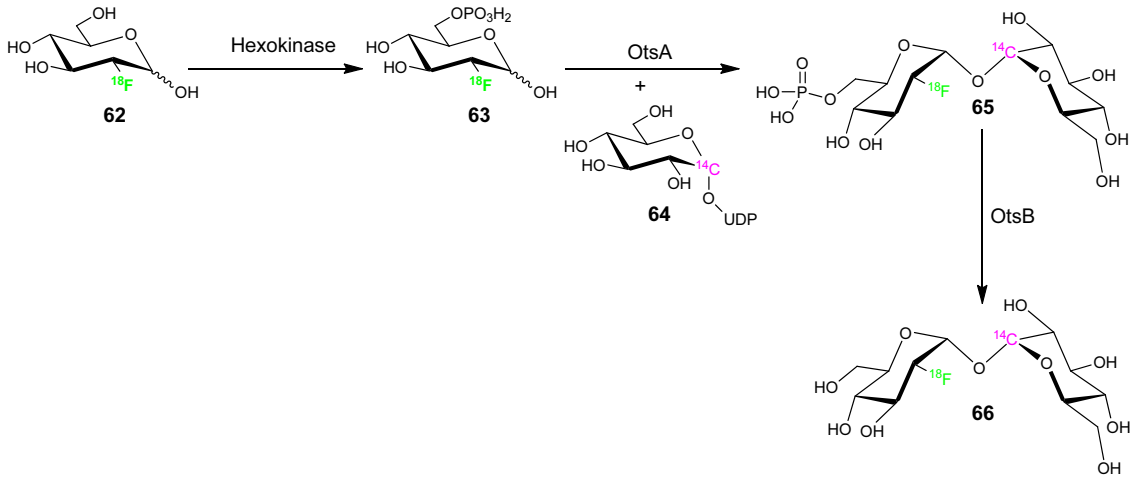
[i] **41-45** were synthesized by Dr. Omar Boutureira. **46-53** were synthesized by Mitul Patel. **49-51** and **53** were synthesized by Dr. François D'hooge. **52** was synthesized by Dr. Seung Lee.

synthesis focused on an enzymatic route and this work led to the preliminary reaction conditions that were then improved by others in the laboratory[ii] to successfully synthesize 2-^{18}F-trehalose. Although the enzymatic synthesis offered promise for synthesis of some trehalose analogues, this route was challenged by the substrate specificity of the enzymes, which limited the breadth of library that could be generated. To expand the library scope, four additional trehalose modification strategies were explored: 1) A radical bromination strategy was investigated and proved unsuccessful at modifying trehalose. 2) Intramolecular aglycone delivery (IAD) was hindered by the difficulty of synthesizing the target donor stereoselectively. 3) A mannose directed glycosylation afforded trehalose as a complex mixture of stereoisomers. 4) A methyl ketoside based synthesis, where a methyl group on the anomeric position of the donor directs the stereochemistry of the glycosylation, successfully yielded a library of trehalose analogues.

2.2.5 Enzymatic synthesis

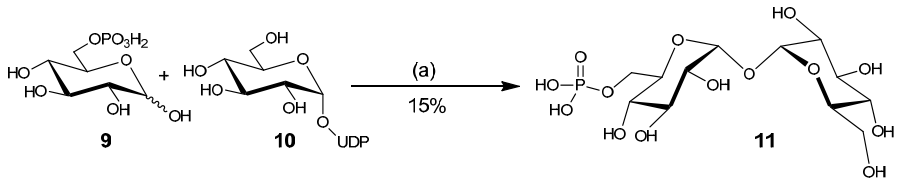
Syntheses that facilitate the late stage incorporation of both ^{18}F and ^{14}C labels were desirable for this project to generate both PET probes and the corresponding ^{14}C versions of compounds for autoradiography, respectively. ^{14}C labeled D-glucose-6-phosphate is readily available from Sigma Aldrich. Likewise, 2-^{18}F-2-deoxy-D-glucose (^{18}FDG) is available from most cyclotron sources. ^{13}C-labeled glucose is also readily available and could prove useful as a isotopic label for mass spectrometry and hyperpolarized MRI. As such, a synthesis that utilized these commercial precursors and linked them to either glucose or fluoro-glucose was envisioned. The trehalose 6-phosphate synthase and phosphatase enzymes, OtsA and OtsB respectively, had been cloned in the laboratory of Professor Ben Davis and had previously been demonstrated to synthesize trehalose from UDP-glucose and glucose-6-phosphate. The trehalose-6-phosphate synthase (OtsA) catalyzes the transfer of glucose from UDP-glucose to glucose-6-phosphate to form trehalose-6-phosphate **11**. The trehalose phosphatase (OtsB) then cleaves off the phosphate group from **11** to generate trehalose[6]. These enzymes were employed to enzymatically produce ^{18}F-trehalose and in future work they could be used to generated ^{14}C-trehalose and a dual labeled ^{14}C-^{18}F -trehalose (**Scheme 2.2**).

[ii] Dr Seung Seo Lee, Dr. Francois D'Hooge (University of Oxford) and Dr. Gwendolyn Mariner (TRS, NIH, NIAID)



Scheme 2.2. Synthetic approach to incorporate either or both ^{18}F and ^{14}C radioactive labels into the trehalose scaffold, utilizing the trehalose-6 phosphate synthetase (OtsA) and the phosphatase (OtsB)

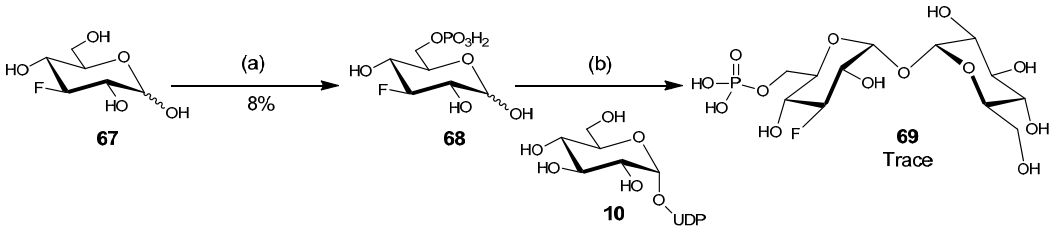
It was found that trehalose-6-phosphate could be synthesized in 15% yield by OtsA from **9** and **10** under un-optimized conditions (**Scheme 2.3**). Initial problems with contamination by glucose-6-phosphate were eliminated by the addition of a slight excess of UDP-glucose **10**.



Scheme 2.3. Synthesis of trehalose-6-phosphate (b) OtsA, 20 mM $MgCl_2$ and 100 mM TEA, 37 ˚C, 16 h.

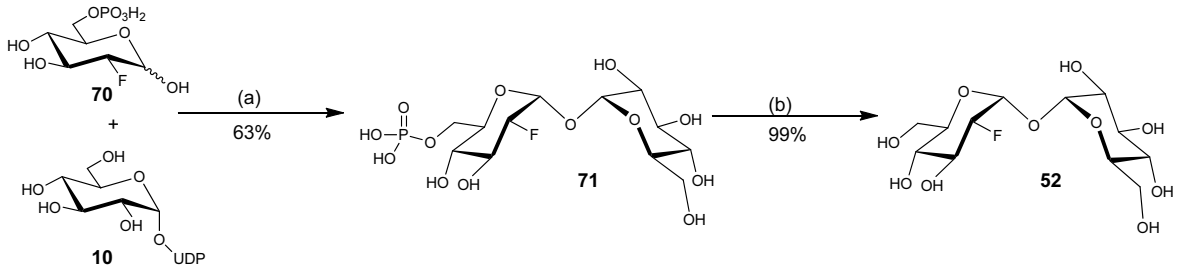
This enzymatic approach was also utilized in an attempt to synthesize 3-F-3-deoxy-D-trehalose-6-phosphate and 2-F-2-deoxy-D-trehalose-6-phosphate **69**. 3- and 2-F-2- or 3-F-3-deoxy-D-glucose-6-phosphate **68, 70** were synthesized enzymatically with commercially available hexokinase, pyruvate kinase, phosphoenolpyruvate and ATP and were purified by anion exchange in somewhat low yield with large quantities of salt (**Scheme 2.4**). Pure 3-F-3-deoxy-D-glucose-6-phosphate **68** was obtained in 8% yield. 2-F-2-deoxy-D-glucose-6-phosphate **70** was detected by mass spectrometry and NMR, but was found to be a mixture of manno- and gluco- epimers, which made full characterization difficult (see **Scheme 2.27** for synthesis of pure 3,4,5-tri-*O*-acetyl-2-F-2-deoxy-D-glucose). **68** and **70** were then reacted with **10** in the presence of OtsA. A small peak at M+1 423, which corresponded to the fluorinated trehalose derivatives **69** and **71**, was detected by ESI mass spectrometry. More **10**

and additional enzyme were added to the reaction in order to try to force completion. However, after 36 hours, upon HPLC purification and lyophilization of all column fractions, no product was isolated.



Scheme 2.4. Attempted synthesis of 3-F-3-deoxy-D-trehalose-6-phosphate, 69 (a) ATP, phosphorenol-pyruvate, 20 mM TEA buffer, 10 mM $MgCl_2$ and 5 mM KCl pH 7.6, pyruvate kinase and hexokinase , 37 ˚C, 18 h, 8% (b) OtsA, 20 mM $MgCl_2$ and 100 mM TEA, 37 ˚C, 18 h.

The lack of reactivity was surprising. It was hypothesized that either the batch of enzyme was old and less active than it had been originally or that the 3-F-3-deoxy- D-glucose-6-phosphate might inhibit the glycosylation reaction catalyzed by the OtsA. Dr. Seung Lee subsequently optimized this reaction and found that vast excess of enzyme and UDP-glucose were required to force these reactions to completion (see **Appendix 1** for details).



Scheme 2.5. Synthesis of 2-F-2-deoxy-trehalose 52.(a) OtsA, HEPES buffer, pH 7.4, 20 mM $MgCl_2$, 30 °C, 16 h. (b) Alkaline phosphatase, pH 8.0, 37 °C, 2 h[iii].

2.2.6 Radical bromination

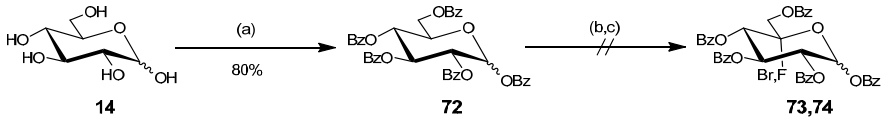
Selective radical halogenation[7,8] has previously been utilized to brominate the 5- position and following bromination, fluorination can easily be achieved by treatment with $AgBF_4$. The C-5 fluoro modification has been shown to provide inhibitors of glycosyl transferases[9]. It was hypothesized that modifications at the 5-position could modulate the electronics of the trehalose, potentially serving as novel inhibitors and/or substrates of the Ag85 enzymes or

[iii] Synthesis carried out by Dr. Seung Lee

OtsA. With the generation of $AgB^{18}F_4$ or $Ag^{18}F$, this strategy would allow for the late stage incorporation of the radiolabel.

Prior successful work with radical bromination by Dr. Katie Doores in Professor Ben Davis's lab suggested that radical bromination of peracetylated mannose could be carried out successfully, but produced a fairly unstable product. In contrast with the acetylated mannose, Dr. Doores found that bromination of the corresponding benzoylated mannose proceeded smoothly[10]. Following the mannose precedent it was hoped that radical bromination of benzoylated glucose and/or trehalose would afford a stable product and that the bromine could be replaced with a fluorine using $AgBF_4$ in diethyl ether.

Numerous conditions were tested to radically brominate both acetylated or benzoylated glucose and none afforded the brominated compound in our hands (**Scheme 2.6**). This lack of reactivity may be explained by lower reactivity or greater instability of the gluco-radical compared with the corresponding mannose radical investigated previously. Reactions were heated in carbon tetrachloride under irradiation with a 300 W lamp. A mixture of α- and β-perbenzoylated glucose **72** was treated with N-bromosuccinimide (NBS) and reaction was monitored by NMR and TLC. By NMR, starting material was incompletely consumed. Peracetylated glucose was also utilized and a complex mixture of products was obtained. Addition of more NBS did not aid the reaction. Likewise, addition of catalytic benzoyl peroxide did not aid in the reaction. Ultimately a brominated product **73**, hydrolysis product and an elimination product were detected by mass spectrometry, but chromatography failed to afford pure compounds. Attempts to fluorinate the crude reaction mixture also did not afford the desired product **74**.

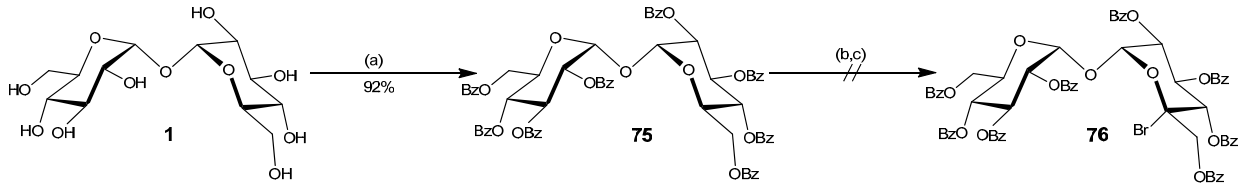


14 **72** **73,74**

Scheme 2.6. Radical bromination of benzoylated glucose 72 showed no reactivity. (a) Benzoyl chloride, pyridine, DMAP, RT, 16 h (b) NBS, CCl_4, UV light, 3-8 h, 100 ˚C, complex mixture (c) $AgBF_4$, no reaction.

Radical bromination was also attempted on trehalose (**Scheme 2.7**). Trehalose was benzoylated in 62% yield with benzoylchloride and pyridine, affording benzoylated **75**. **75** was then irradiated in the presence of NBS, NBS and benzoyl peroxide or Br_2 in carbon

tetrachloride. Br_2 was suggested to be a more powerful brominating agent[7]. All reaction conditions revealed a new less polar spot at R_f = 0.8 (2:1 Petrol:EtOAc), but no new peaks were seen by crude NMR. Starting material was re-isolated in near quantitative yield. Reflux at 100 °C with bromine for 10 h produced multiple product spots by TLC. As of yet, no 5-Br-trehalose has been isolated (See **Appendix 1** for further details).



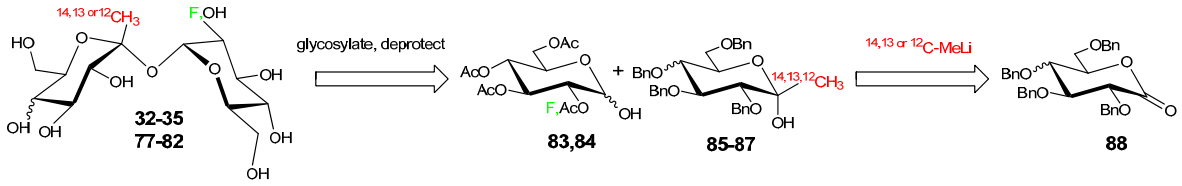
Scheme 2.7. Radical bromination of trehalose (a) Benzoyl chloride, pyridine, RT, 16 h (b) NBS, CCl_4, UV light, 3-8 h, 100 ˚C, no reaction (c) Br_2, CCl_4, UV light, 10 hours, 100 ˚C, complex mixture.

This project was continued by Dr. François D'Hooge, who optimized bromination conditions for peracetylated glucose. Following literature precedent, he successfully found that the β-acetyl protected sugar proved more reactive than the benzoyl protected **72** described above[11]. Consistent with the attempts detailed in **Scheme 2.7** all conditions tested by Dr. D'Hooge for the bromination of trehalose (both peracetyl and perbenzoyl) did not yield the desired 5-Br-trehalose. It was hypothesized that the clamshell-like conformation that trehalose favors[12] may block the bromination reaction at C-5.

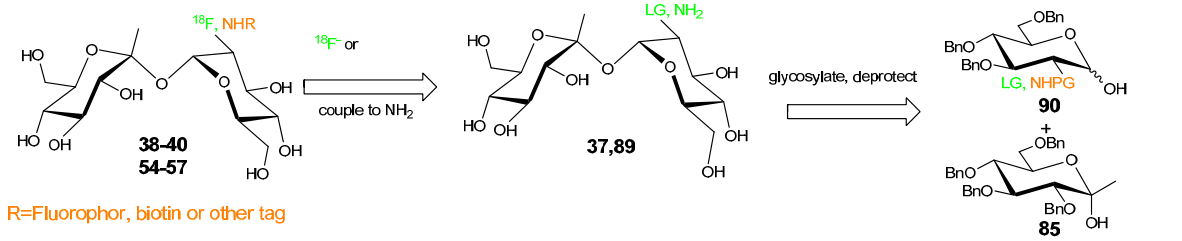
2.2.7 Ketoside trehalose synthesis

2.2.7.1 Retrosynthetic analysis

Ketoside-type analogues of trehalose were identified as possible candidate molecules for ^{14}C-F-trehalose because their synthesis offered α,α-selective glycosylation and late installation of an isotopic ^{13}C or radioactive ^{14}C label from the reaction between isotopically labeled[14,13, or 12]C-MeLi and gluconolactone analogue **88** (**Scheme 2.8**). Displacement of a good leaving group and/or coupling reactions to aminotrehalose compound **37** would also allow for the late stage placement of ^{18}F-label or other tag (fluorophores, biotin, etc) (**Scheme 2.9**). In this section we describe the successful synthesis of a library of methyl ketoside analogues, including, ^{13}C, fluoro, amino, fluorescent, biotinylated and quantum dot trehalose derivatives.



Scheme 2.8. Retrosynthetic analysis of ^{14}C-trehalose and ^{14}C-2-F-2-deoxy trehalose by methyl ketoside glycosylation, starting from readily available precursors 83, 84 and 85-87.

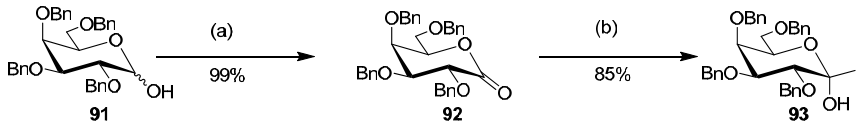


Scheme 2.9. Retrosynthetic analysis of ^{18}F-trehalose and fluorescent/biotinylated-trehalose by methyl ketoside glycosylation. Commercially available ^{18}F can be used to displace good leaving group (LG) in a late stage addition of the radioactive tag. 2-Amino trehalose can be used as a coupling partner for a variety of amino-reactive tags (fluorophors, biotin, etc). PG-Protecting group.

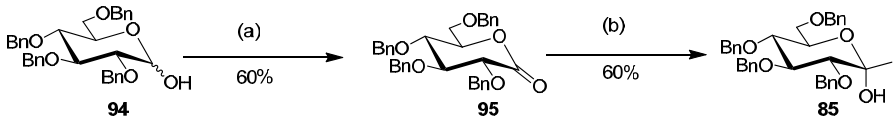
This work builds on previous investigations of ketosides[13,14] and exoglycals[15,16] as donors for α,α-selective trehalose synthesis. Previous groups had found that protected heptulo-pyranosides with gluco **85**, galacto **93** and manno configurations all served as excellent donors for fairly α-selective glycosylations, affording stereoselectivity between 6:1 (α,α:α,β) and exclusively α,α-disaccharide, using bismuth triflate[17] as catalyst. We found that modified trehalose analogues, can readily be synthesized utilizing Lewis acid catalyzed α-selective-*O*-glycosylation of ketoses and exoglycals.

2.2.7.2 Syntheses of donors and acceptors

Synthesis of the methyl-hexopyranose derivatives of glucose **85** and galactose **93**, required as donors for glycosylation, proceeded smoothly. Swern oxidation was employed to synthesize both 2,3,4,6-tetra-*O*-benzyl-D-gluconolactone **95** and 2,3,4,6-tetra-*O*-benzyl-D-galactonolactone **92** in near quantitative yields from commercially available precursors. Treatment with methyllithium stereoselectively generated **85** and **93** in good yields (**Scheme 2.10** and **Scheme 2.11**).



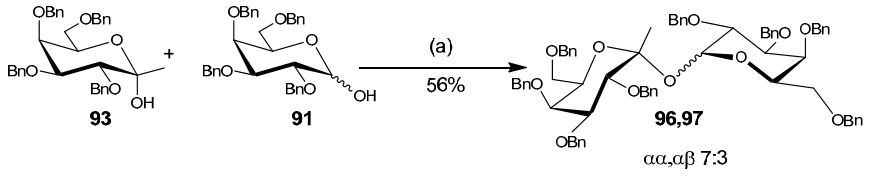
Scheme 2.10. Synthesis of galacto donor 93. (a) Ac_2O, DMSO, RT, 16 h (b) Methyllithium, -78 °C, 1.5 h.



Scheme 2.11. Synthesis of gluco donor 85. (a) Ac_2O, DMSO, RT, 16 h (b) Methyllithium, -78 °C, 1 h.

2.2.7.3 Glycosylation of ketosides

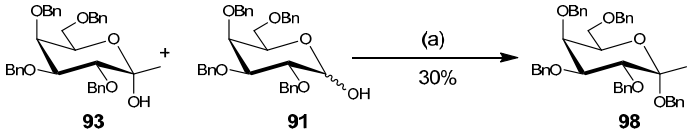
A ketoside glycosylation (**Scheme 2.12**) was first attempted with the 3,4,5,7-tetra-*O*-benzyl-1-deoxy-D-galacto-heptulopyranose **93** as donor and commercially available 2,3,4,6-tetra-*O*-benzyl-D-galactose **91** as acceptor. Bismuth triflate (5% mmol) was added to an anhydrous solution of **91** in DCM. Then, the **93** as added via syringe and the solution stirred. This reaction afforded product as a mixture of anomers **96** and **97**, even though the bismuth triflate had been reported to give exclusively the α product.



Scheme 2.12. Glycosylation with bismuth triflate catalysis. (a) Bismuth triflate 5% mmol, DCM, 4 Å molecular sieves, 0 °C, 3 h.

Subsequently, when the reaction was repeated, it was dominated by an overwhelming side reaction (**Scheme 2.13**), which was most pronounced when too much triflate was added or when reaction was left for too long or warmed to room temperature. In this side reaction, discovered by Li, 2001[18], the dissaccharide decomposes into the corresponding glucoside **98**. The glucoside runs slightly higher on TLC than the disaccharide and has a methyl NMR shift more downfield and is readily detected by ESI M+1 =662. In repeat experiments, before full understanding of this side reaction, the reaction was left overnight, allowed to warm to RT and this side product was isolated in 30% yield. In a second repeat experiment, near

quantitative amounts of triflate were added and substantial amounts of side product was detected. This side product repeatedly lowered the yield of the ketoside reactions, so we will briefly discuss some of the efforts taken to optimize the glycosylation and minimize the production of the unwanted glycoside **98.**

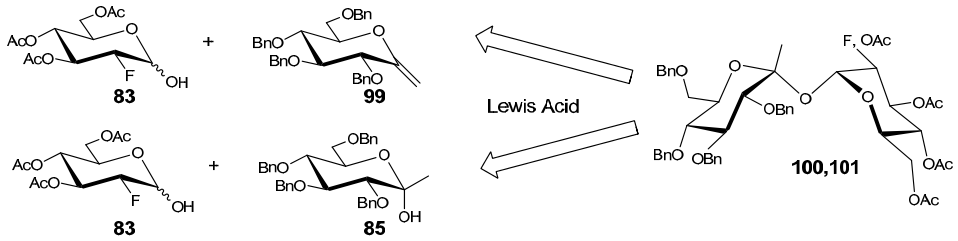


Scheme 2.13. Side reaction of ketoside glycosylations. (a) Bismuth triflate, 4 Å molecular sieves, 3 h at 0 °C and then 16 h at RT.

2.2.7.4 Glycosylation of exoglycals

A review of the literature on glycosylations with heptulo-pyranosides suggested several solutions to the problem of the pyranoside side product[17-19]. Several other triflate catalysts have been successfully utilized, including TMSOTf, Tf_2NH, TfOH, $Sc(OTf)_3$, and BF_3[13,14,20]. Potentially a different catalyst might reduce side reactions. However, these other catalysts have reportedly lower levels of α,α-selectivity than Bi(III)OTf[17].

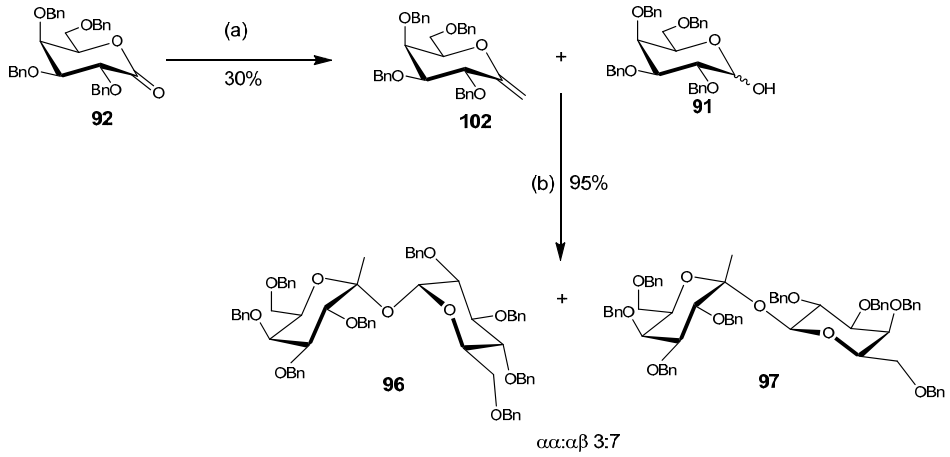
Alternatively, 1-methylene donors (**Scheme 2.14**) have been utilized as donors for the synthesis of methyl trehalose analogues. Li et al[19] synthesized the exoglycal in one step from the protected lactone, utilizing the Tebbe reagent[21]. The glycosylation of methylene sugars has been successfully carried out with Lewis acid catalysis. Typically the reported glycosylation with exomethylene sugars is complete within 15 minutes. This quicker reaction time was hypothesized to help avoid side product formation. The drawbacks of exomethylene sugars reported in the literature include acid sensitivity and instability; exoglycals cannot be stored for long periods of time and product often degrades during column chromatography. Another disadvantage of the exomethylene approach was a more difficult introduction of the ^{14}C label. Previously, the introduction of ^{14}C was planned through the use of ^{14}C-methyllithium during the synthesis of the **85**. Generation of ^{14}C labeled Tebbe reagent is less straightforward than the synthesis of ^{14}C-methyllithium. An alternative to Tebbe is the use of the Petasis reagent[22], which is stable at room temperature and easily be generated from methyl iodide.



Scheme 2.14. Alternative glycosylation strategies to prevent side product formation.

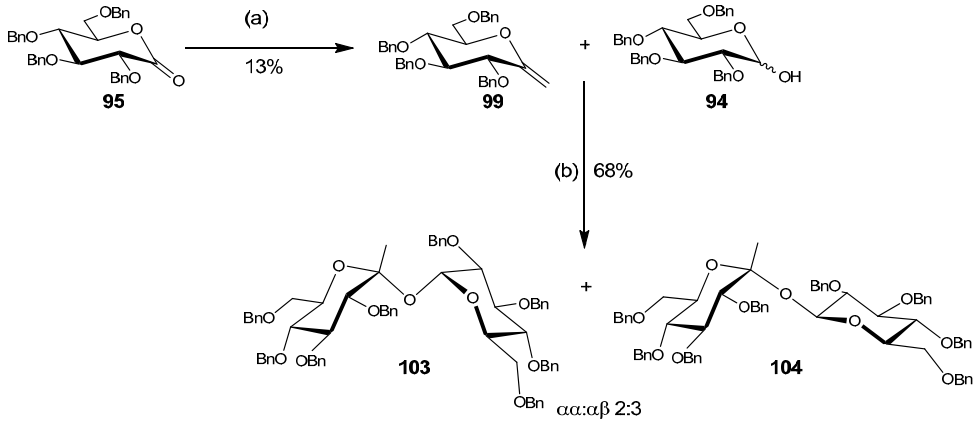
Stereoselectivity was also considered during the synthetic strategy. The catalyst may aid in the α,α-selectivity. As noted previously, Bi(III)OTf supposedly yielded exclusively the α,α-trehalose analogue. However, other catalysts were reported to afford in a mixture of α,α- and α,β-products. One can consider the nature of the protecting group to increase the α,α selectivity. For example acyl or benzoyl groups at the 2- position usually drives the formation of the 1,2-*trans* linked O-glycosides. With non-participating groups, such as benzyl- or fluoro-substitution, typically the α-glycoside is typically formed, due to the anomeric effect. Additionally, the nature of the solvent can affect α-selectivity, with less polar solvents benefitting α-selectivity. However, Li[19] found that solvent does not affect the α,β ratio of glycosylation of exomethylene sugars. It was thought that the potentially longer reaction time with reduced temperature and the use of a less polar solvent, such as ethyl ether, could increase the α-selectivity of the ketoside glycosylation. Finally, Namme[13] found that the use of trimethyl-silyl-α-glycoside acceptor sugars yielded exclusively the α,α linkage. Trimethyl-silyl-α-glycosides have been investigated by Jansson et al and were found to be excellent donors for α -selective synthesis[23]. They are more difficult to synthesize and, therefore, would be considered if other approaches to increase the reaction stereoselectivity failed.

After several unsuccessful attempts to glycosylate with the **85** and **93**, the exoglycal route was investigated. Tebbe's reagent was commercially available and both the exomethylene-glucal **99** and the exomethylene-galactal **102** were synthesized according to previous protocols[24] from the corresponding lactones **92** and **95** (**Scheme 2.15, Scheme 2.16**). These reactions proceeded to completion on TLC. However, upon purification, yields were quite low. The exomethylene sugars are known to be acid sensitive, so it was thought that the product was decomposing on the silica. Even with the use of 1% NEt_3 to neutralize the silica, column yields of donors **99** and **102** were only roughly 30%.



Scheme 2.15. Exomethylene glycosylation synthesis of 96 and 97. (a) Tebbe, -78 °C, 1 h, (b) TMSOTf, DCM 4 Å molecular sieves, -78 °C, 40 min.

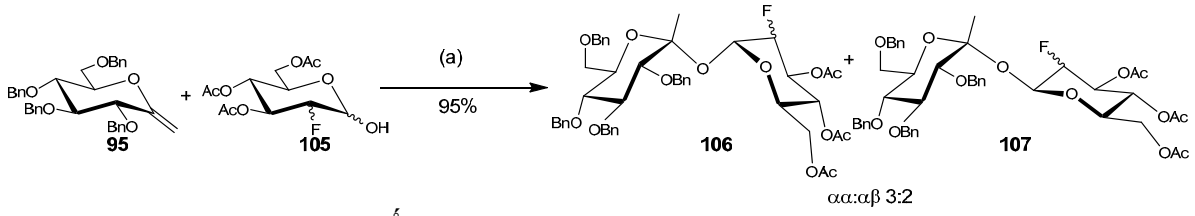
The synthesis of 1-methyl-galacto-trehalose compounds **96 and 97** proceeded smoothly through the exomethylene route, with a 95% yield for the glycosylation step (**Scheme 2.15**). The glycosylation was not stereoselective, producing a 3:7 ratio of α,α:α,β. The total yield over three steps was 38%, with a 12% yield of the desired α,α product.



Scheme 2.16. Exomethylene glycosylation synthesis of 103 and 104. (a) Tebbe, -78 °C, 1h, (b) TMSOTf, DCM 4 Å molecular sieves, -78 °C, 45 min.

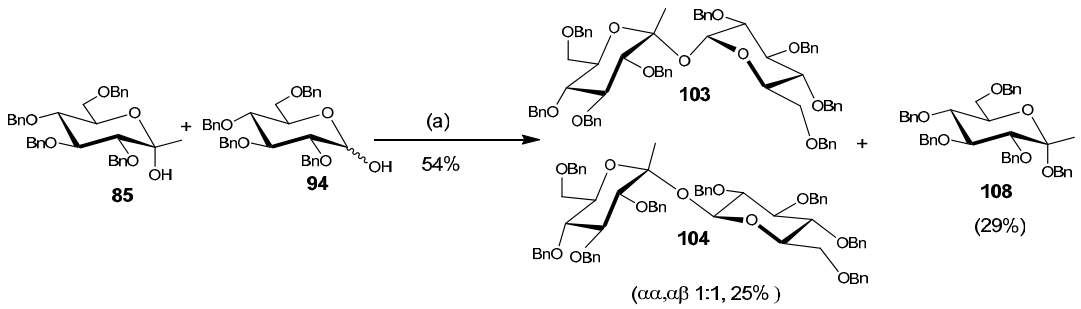
As with the exomethylene-galactal **102**, the exomethylene-glucal **99** glycosylation proceeded well (**Scheme 2.17**), with a 68% yield. The lower yield relative to the galactose may reflect a smaller scale reaction or the lower reactivity of the glucal relative to galactal. The overall yield of **103** and **104** over three steps was 23%. As with galactose, no stereoselectivity was observed with a product ratio of 2:3 α,α:α,β.

To determine whether the exomethylene reactions would tolerate acetylated and fluorinated sugars as acceptors, a mixture of 3,4,5-tri-*O*-acetyl-2-F-2-deoxy-D-glucose/mannose **105** was glycosylated with the exomethylene-glucose **99**. Gratifyingly, the glycosylation proceeded well, with a 95% yield. The selectivity was still non-existent with a 3:2 of α,α:α,β. During this reaction it was discovered that the **105** (synthesized via Selectfluor™) contained a mixture of the manno- and gluco-compounds. Only the α,β-2-F-2-deoxy-gluco-product **107** could be isolated to purity from the reaction mixture (**Scheme 2.17**).



Scheme 2.17. (a) TMSOTf, DCM, 4 Å molecular sieves, -78 °C, 30 min.

As the glycosylations with exomethylene sugars did not yield high stereoselectivity, glycosylations with the 1-alkyl-sugars were revisited. In place of bismuth triflate, fresh TMSOTf Lewis acid was employed and the glycosylation of **85** with **94** (**Scheme 2.18**) was attempted at -40 °C and afforded a 25% yield of benzyl protected 1-methyl-trehalose analogues **103** and **104** and a 29% yield of side product **108** as a 1:1 mixture of α,α and α,β anomers.



Scheme 2.18. Ketoside glycosylation with 2,3,4-6-tetra-*O*-benzyl-D-glucose 94. (a) TMSOTf, DCM 4 Å molecular sieves, 0 °C, 4 h.

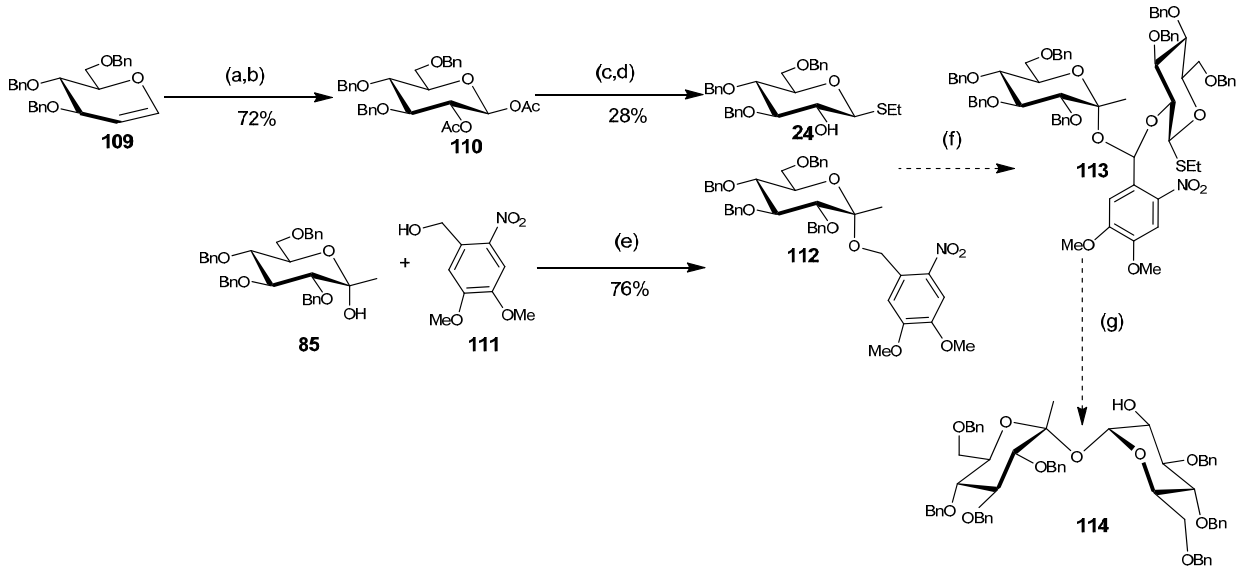
Compared to the galactose analogue **93** (**Scheme 2.12**), the reaction of glucose analogue **85** proceeded more slowly and was not complete after 6 hours. Side product formation was more significant for gluco- than for galacto-product. This could be explained by the distortion of the oxocarbenium species. The axial hydroxyl of the galactose configuration would make the

ring distortion of oxocarbenium species formation occur more quickly than with the equatorial hydroxyl of glucose. Cooling the reaction from 0 °C to -40 or -78 °C reduced the side product formation but did not drive the formation of more of the desired α,α-product. In summary, glycosylation with benzylated 1-methyl-glucose and galactose donors **85** and **93** and tetrabenzyl acceptors **91** and **94** proceeded slowly with production of large quantities of side product and a lack of stereoselectivity.

2.2.8 Intramolecular aglycone delivery (IAD)

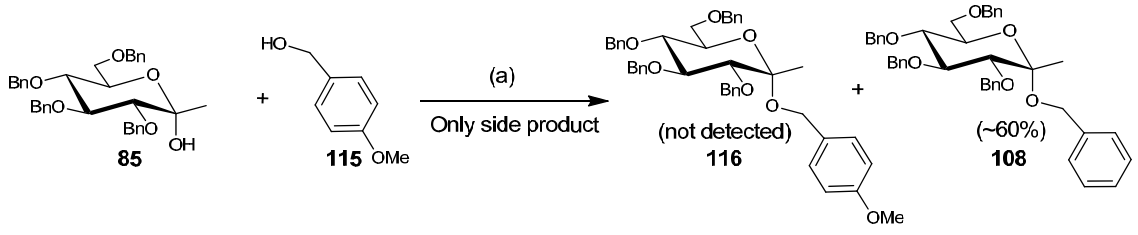
Because the benzylated methyl ketoside approach appeared to offer poor stereoselectivity, we considered an intramolecular aglycone (IAD) synthesis (see **Appendix 1** for details) of trehalose analogues, based on Bertozzi's work[25]. The donor sugar, ethyl 2-*O*-acetyl-3,4,6-tri-*O*-benzyl-1-thio-β-D-glucopyranoside **24**, was synthesized from the 3,4,6-tri-*O*-benzyl-D-glucal **109**, utilizing an innovative dihydroxylation reaction with sodium periodate, ruthenium chloride and cerium chloride[26]. The diol was then acetylated and glycosylated with ethanethiol (**Scheme 2.19**) to afford the desired donor **27** in good yield.

For the acceptor sugar, two intramolecular aglycone syntheses of asymmetric trehalose were considered. Following the work of Bertozzi[25], the trehalose could be synthesized without an anomeric methyl modification. This can be accomplished by making the glycosyliodide[27], followed by stereoselective glycosylation with *para*-methoxybenzyl alcohol (PMB) **115** or 4,5-dimethoxybenzyl alcohol (DMB) **117** . The acceptor can then be oxidized with DDQ and trapped by the known thioglucoside[28]. Alternatively, it was thought that the methyl ketoside sugars **93** and **85** could be employed to stereoselectively generate the α-methoxybenzyl-glucose acceptor, circumventing the synthesis of the glycosyliodide. This second strategy was chosen and the acceptor was synthesized in one step by glycosylating **85** with 4,5-dimethoxy-2-nitrobenzyl (DMNB) alcohol **111** (**Scheme 2.19**).

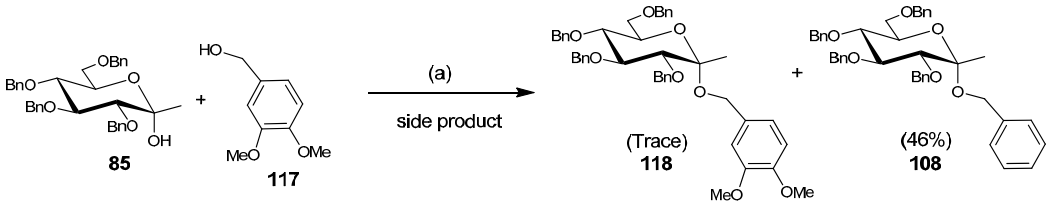


Scheme 2.19. Intramolecular aglycone delivery synthesis. (a) $RuCl_3$, $CeCl_3$, $NaIO_4$, EtOAc/CH_3CN/H_2O (3:3:1), RT, 15 min, 81% (b) pyridine, Ac_2O, 2 h, 90%, (c) BF_3, EtSH, RT, 4 h, 76% (d) NaOMe, methanol, 2 h, 37% (e) TMSOTf, -40 ˚C, 5 min, (f) DDQ, 15 min (g) MeOTf, DTBMP, 60 °C, 16 h.

This glycosylation of **85** with DMNB proceeded extremely quickly to give exclusively the α-product **112** (**Scheme 2.19**). The nitro group was not necessary for the reaction series, but was chosen for initial test reactions due to ready availability. The electron withdrawing nature and bulk of the *ortho*-nitro-group could interfere with the DDQ oxidation. PMB and DMB were obtained and glycosylations with the methyl ketoside donors were attempted but found to proceed poorly relative to the DMNB reaction. With the PMB, only the benzyl alcohol side product was detected by ESI MS and TLC (**Scheme 2.20**). Therefore, no purification was carried out. For the DMB, trace product **116** was detected. However, purification yielded no pure product **118** and side product in 46% yield (**Scheme 2.21**).

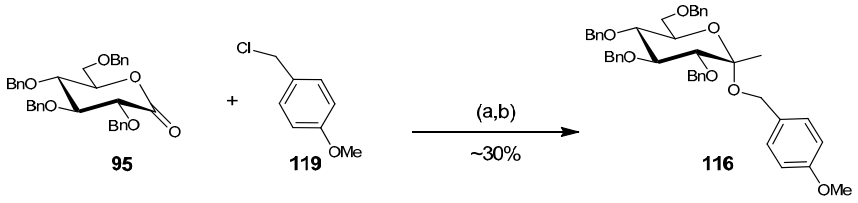


Scheme 2.20. Glycosylation with *para*-methoxybenzyl alcohol (PMB). (a) TMSOTf, -78 ˚C, 2 h, produced exclusively side product 108.



Scheme 2.21. Glycosylation with dimethoxybenzyl alcohol (DMB). (b) TMSOTf, -78 ˚C, 2 h, 46% side product 108 and trace product 118.

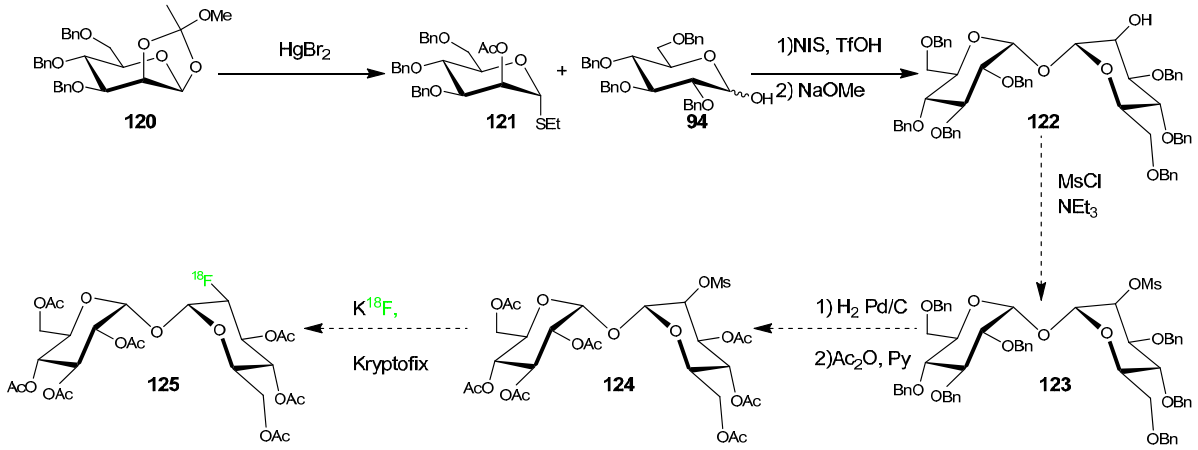
We postulated that it might be possible to trap the product *in situ* from the lithiated anion, formed during the methyllithium addition to the gluconolactone compound **95**. This reaction was attempted but no reaction was observed between anion and *para*-methoxybenzyl chloride (PMBCl). However, when sodium hydride was added, product **118** was formed (**Scheme 2.22**) and none of the side product was produced. However, removal of the remaining PMBCl proved challenging. These encouraging results suggested to us that the *para*-methoxybenzyl bromide (PMBBr) might be a more reactive PMB source. Reaction with PMBBr failed to afford more than trace product.



Scheme 2.22. Reaction of 95 with PMBCl. (a) Methyllithium, THF, 78 ˚C, 30 min (b) *in situ*, PMBCl added, NaH, 78 ˚C to RT, 2 h , approximately 30% contaminated with PMBOH.

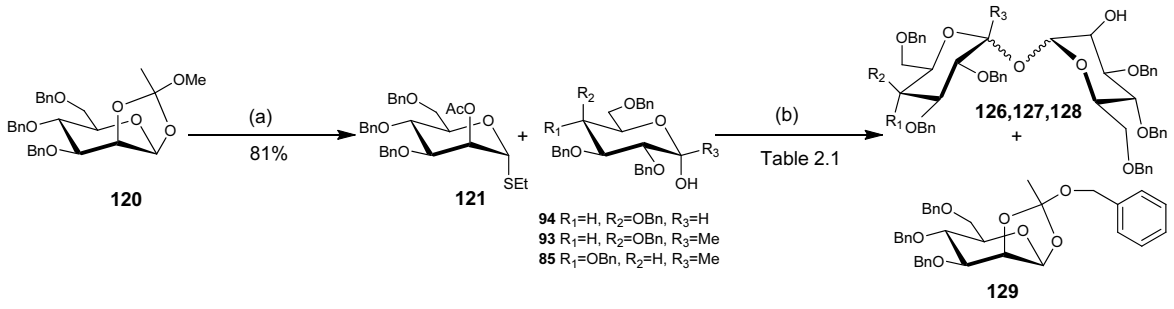
2.2.9 Mannose directed glycosylations

As an alternative to the IAD, a synthesis of asymmetric trehalose, beginning with mannose was considered (**Scheme 2.23**). The advantage of this approach was that the 2'-OH did not need to be epimerized prior to introduction of the ^{18}F label. It was thought that, because of the neighboring group effect, if 2-*O*-acetyl-3,4,6-tri-*O*-benzyl-1-thio-α-D-mannopyranoside **121** was glycosylated with **94** only the α,α- and α,β-products would be obtained.



Scheme 2.23. Progress towards and proposed synthesis of manno-trehalose analogues, using neighboring group participation to form the α,α-linkage. Dotted arrows indicate reactions that were not attempted.

When this glycosylation was carried out mainly the α,α- and α,β-products were produced (as detected by HMBC NMR). However, β,α- and β,β-products were obtained in small amounts. The α,α-product was largely separable from the reaction mixture, affording a somewhat pure sample. Upon replication of this reaction, production of side product **129** became problematic and reaction optimization failed to afford α,α selectivity or high yield of the desired product (**Scheme 2.24**). With synthesis of the methyl ketoside sugars, glycosylations with **93** and **85** were attempted with the anomeric methyl group to direct the stereochemistry. The use of the ketoside failed to increase the stereoselectivity and instead drove the reaction towards side product **129** formation (**Table 2.1**).



Scheme 2.24. Mannose directed glycosylations. (a) $HgBr_2$, 60 ˚C, MeCN, 5 h, (b) NIS, TfOH, 30 min - 1 h, Table 2.1 for yields.

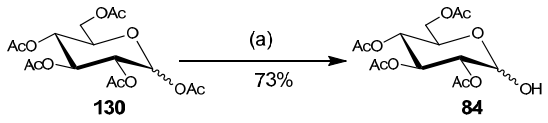
Table 2.1. Manno-thioglycoside 121 glycosylation reactions

Entry	Acceptor	Time	Yield (anomeric mixture)	Yield side pdt 129	Recovered Donor
1	94	1 h	24%	10%	none
2	93	30 min	10%	48%	26%
3	85	1 h	26%	40%	53%

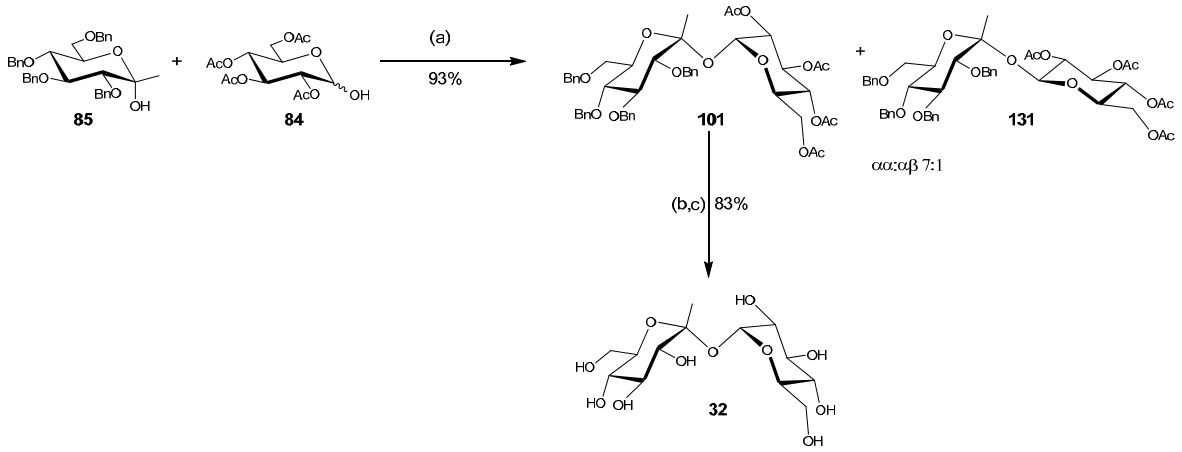
2.2.10 Methyl ketoside glycosylations with acetylated acceptors

Because the benzylated donors afforded poor stereoselectivity and low yields and the IAD and mannose strategy proved challenging, the ketoside glycosylations were reconsidered with acetylated donors. In all previous work on ketoside trehalose analogues, benzylated acceptor sugars were utilized. However, after significant reaction optimization it was found that the tetrabenzyl-ketosides could be glycosylated with acetylated acceptors. Acetylated acceptors provided quicker reaction time, approximately 1 hour versus 6 hours, higher yields (80-95%) instead of 30-40% and greater alpha selectivity (7:1) than benzylated sugars (1:1). Fortuitously, fluorinated compounds also proved to be an excellent acceptor for these glycosylations and allowed for the synthesis of fluoro trehalose analogues. See **Table 2.1** and **Appendix 1** for full optimization work. The following sections describe the synthesis of the acetylated acceptors and the successful synthesis of a ketoside-trehalose library.

2,3,4,6-Tetra-*O*-acetyl-D-glucose **84** was considered as an alternative acceptor that possibly would have greater reactivity than tetra benzyl glucose and reduce the amount of side product **108** formed during the ketoside glycosylations. To synthesize **84**, 1,2,3,4,6-penta-*O*-acetyl-D-glucose **130** was selectively deprotected at the anomeric position with the use of hydrazine acetate as previously reported in Kahn et al (**Scheme 2.25**)[29].

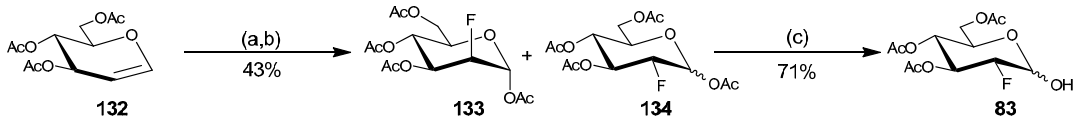


Scheme 2.25. Synthesis of 2,3,4,6-tetra-O-acetyl-D-glucose 84. (a) Hydrazine acetate, DMF, RT, 18 h.



Scheme 2.26. Synthesis of methyl-trehalose compound 32. (a) TMSOTf, 4 Å molecular sieves, anhydrous DCM, –40 °C, 3 h (b) NaOMe, anhydrous MeOH, RT, 1 h (c) H_2 , $Pd(OH)_2$/C, basic alumina, RT, 72 h.

The glycosylation of **85** was attempted with 2,3,4,6-tetra-*O*-acetyl-D-glucose **84** instead of with 2,3,4,6-tetra-*O*-benzyl-D-glucose **94**. This reaction proceeded smoothly, going almost to completion within one hour in a 93% yield. The product was found to be a 7:1 mixture of α,α and α,β anomers **101, 131** (**Scheme 2.26**). This is an improved yield and better stereoselectivity than was found with the benzylated acceptors. After optimization of the deprotection steps (see discussion in the next section), deacetylation and hydrogenation afforded the desired products in near quantitative yields.

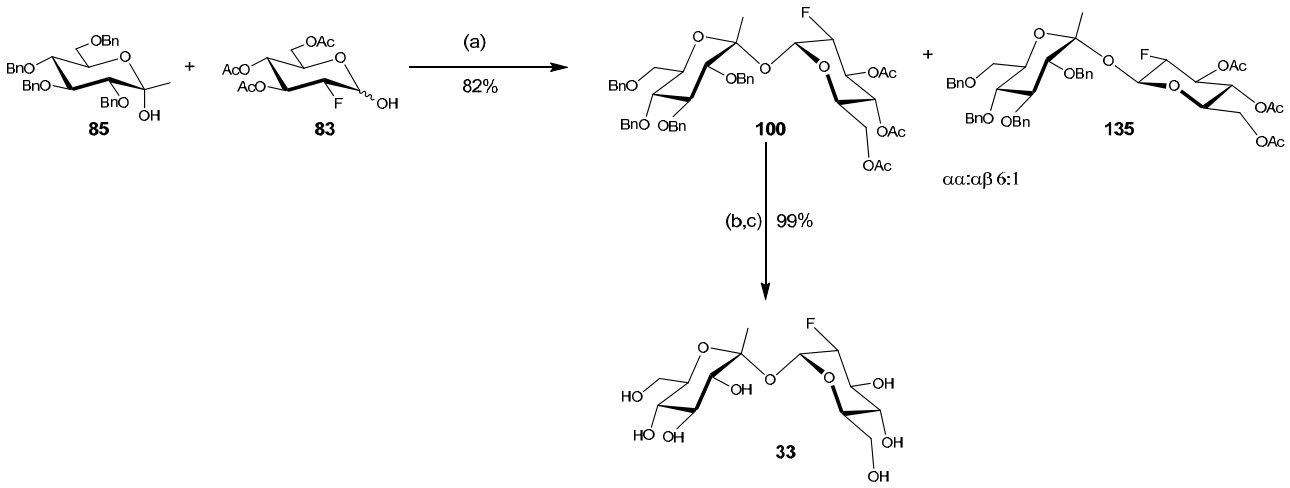


Scheme 2.27. Synthesis of 2-F-2-deoxy-D-glucose 83. (a) Selectfluor, $MeNO_2$, 43% i) RT, 6 h ii) Reflux 30 min (b) Ac_2O, Pyridine, 1 h, RT, (c) hydrazine acetate, DMF, 20 h, 71%.

Glycosylations with **83** and the exomethylene donors had previously hinted that fluoro sugars might serve as possible acceptors (**Scheme 2.17**). 3,4,5-tri-*O*-acetyl-2-F-2-deoxy-D-glucose was synthesized with some difficulty (**Scheme 2.27**) following procedures developed previously[30-32]. Treatment of the triacetyl glucal **132** produced a 1:1 mixture of the acetylated mannose **133** and glucose **134** in 43% yield. The gluco- and manno-sugars were then acetylated with acetic anhydride and pyridine to facilitate separation of the mixture of epimers, following which the anomeric hydroxyl was selectively deprotected, utilizing hydrazine acetate to afford **83** in 71% yield. Overall yield of **83** was only 10%. This was

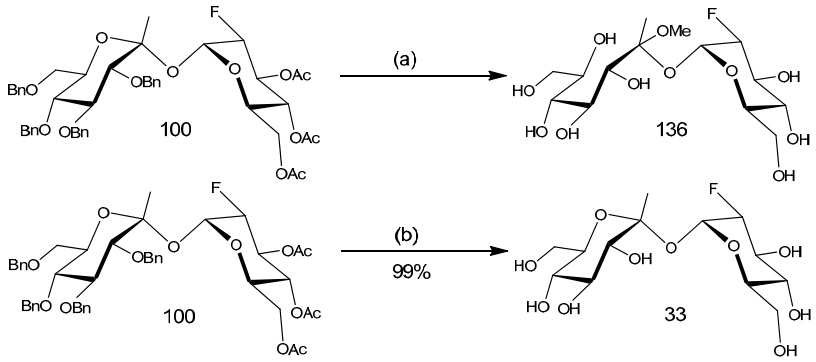
sufficient to continue with the glycosylations. The manno and gluco sugars were differentiated by their fluorine NMR spectra. These spectra are well characterized[33] and the observed chemical shifts matched those reported previously. Length of reaction, solvent and temperature can all affect the ratio of gluco- to manno-sugars[31]. However, optimization of this reaction was not explored further.

Use of acetylated acceptors greatly accelerated the rate of glycosylation and the stereoselectivity of the reaction. When the ketoside glycosylation was carried out with the **83**, glycosylation proceeded in an 82% yield, with a 6:1 ratio of α,α- and α,β-anomers **100** and **135** (**Scheme 2.28**).



Scheme 2.28. Synthesis of compound 33. (a) TMSOTf, 4 Å molecular sieves, anhydrous DCM, –40 °C, 2 h (b) NaOMe, anhydrous MeOH, RT, 3 h (c) H_2 , $Pd(OH)_2$/C, basic alumina, RT, 48 h.

Before synthesizing a library of analogues we turned our attention towards optimization of the deprotection of these compounds to ensure that full synthesis of deprotected ketosides was feasible. Namme[13] and Li[18] have both explored the deprotection of benzyl protected 1-methyl-disaccharides. Both groups found that these disaccharides are acid sensitive and are best deprotected with H_2 $Pd(OH)_2$/C. Namme found that hydrogenation in THF, followed by the addition of MeOH with basic alumina, yielded the desired products in nearly quantitative yields within 3 hours. Li found that nearly the same reaction conditions required 20-30 hours for completion and that a debenzylated, ring-open sugar **136** proved to be a significant side product.



Scheme 2.29. Deprotection of 100. (a) i) NaOMe, MeOH ii) $Pd(OH)_2/C$ NEt_3, MeOH, 14h, decomposed on silica, (b) i) NaOMe, MeOH ii) $Pd(OH)_2/C$, basic alumina MeOH, 30 h.

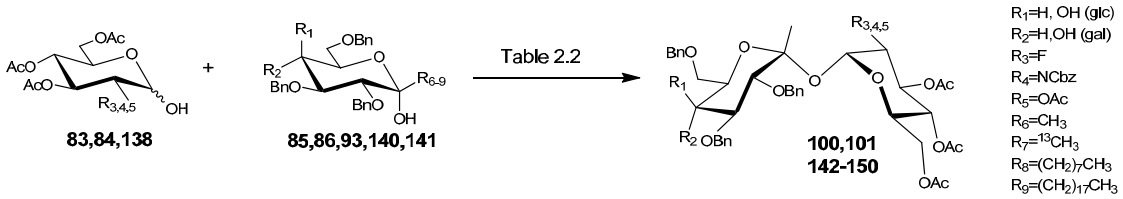
Deacetylation of these compounds proceeded smoothly with sodium methoxide in methanol (**Scheme 2.29**). Hydrogenation was extremely sluggish. For a first attempt, hydrogenation with $Pd(OH)_2$ and NEt_3, produced an unstable baseline product with M+1= 408, which corresponded to the ring open product **136**, reported by Li et al[18]. For the second attempt, basic aluminium oxide was employed. After several additional attempts it was found that the reaction proceeds slowly. After 18 hours the hydrogenation was not complete and partially benzylated trehalose was isolated. After 30 hours, the hydrogenation of **100** was complete, with complete disappearance of the benzyl groups by NMR. Addition of acid to accelerate the reaction was not initially employed due to the reported acid sensitivity of methyl ketosides.

In subsequent experiments it was found that Pd/C or $PD(OH)_2/C$ 20% both catalyzed the hydrogenation reaction and that the basic alumina was not required for successful deprotection, contrary to previous reports[13,18]. Hydrogenation could be accelerated through the use of elevated pressure (H-Cube®). The use of ethanol or trifluoroethanol as a solvent, rather than methanol, eliminated the formation of the ring open product. Optimization of the deprotection step will be discussed further in **Section 2.2.13**.

2.2.11 Methyl ketosides: Reaction scope and library synthesis

Having validated the feasibility of the synthesis of methyl ketoside analogues, a library of modified ketosides was synthesized and the scope of this reation was investigated. Ketoside reactions allowed for the facile introduction of an isotopic label (^{13}C) at the anomeric postion. Addionally this glycosylation tolerated the introduction of longer alkyl chains (C8 and C18)

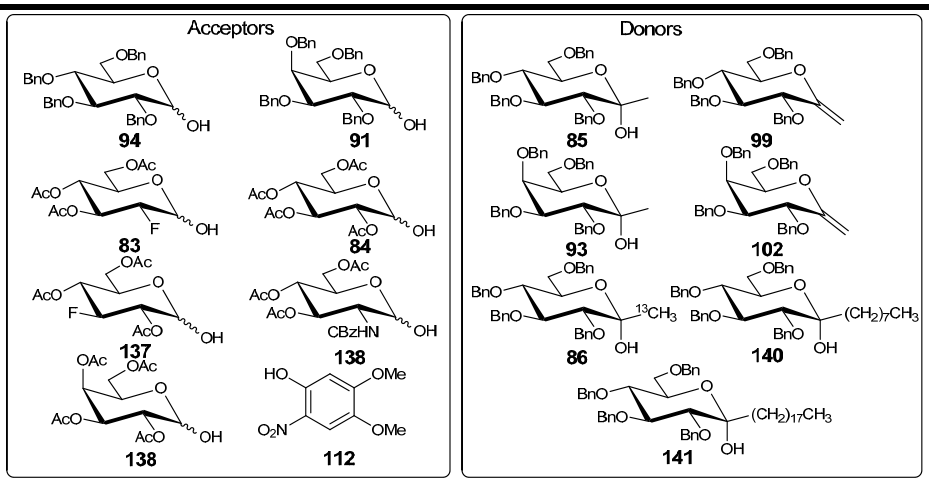
at the anomeric position. These glycosylations allowed for the synthesis of 2 and 3 fluoro, gluco and galacto α,α- and α,β-analogues of trehalose and afforded moderate to good α,α: α,β stereoselectivity. Scaling of the reaction to more than gram scale only modestly reduced the yield and does not affect stereoselectivity. Glycosylation with non-carbohydrate alcohols as acceptors was also found to be possible in good yield. Glycosylation and deprotection with protected glucosamine acceptor sugars was also successful and allowed for the sythesis of 2-amino-2-deoxy-trehalose. The amine handle permitted the late stage coupling to a variety of amino reactive probe groups (Fluorescein, rhodamine, ^{18}F-fluorobenzaldehye, quantum dots, biotin). **Scheme 2.30** and **Table 2.2** summarizes the optimized reaction conditions employed to synthesis the ketoside trehalose library. In sum, this reaction provides a rapid means to synthesized a library of modified ketoside trehalose analogues.



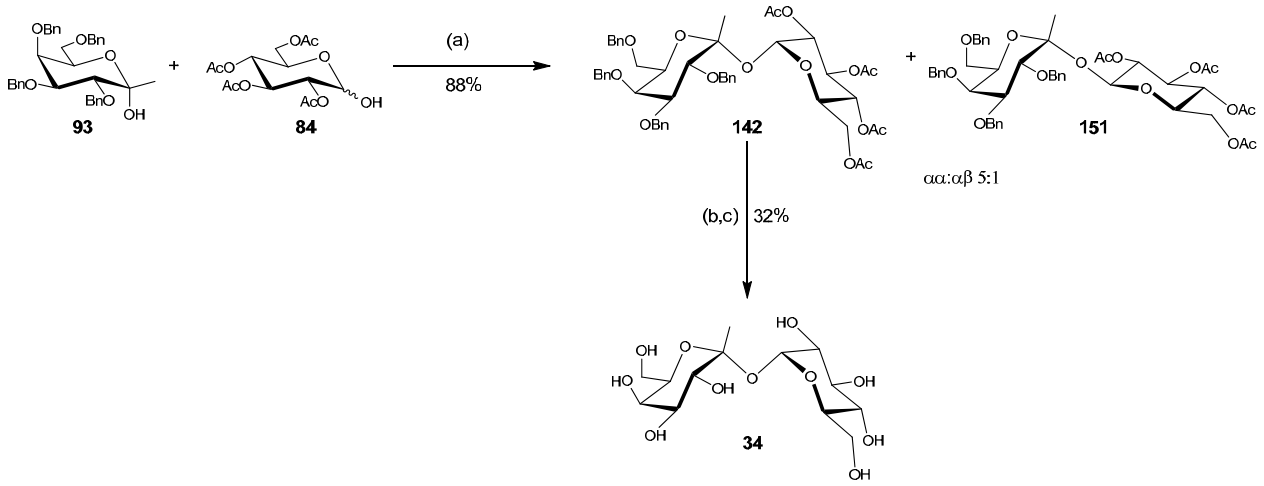
Scheme 2.30. Methyl ketoside glycosylation. R groups indicate all modifications that were tolerated by the reaction.

Table 2.2. Summary of methyl ketoside glycosylation conditions.

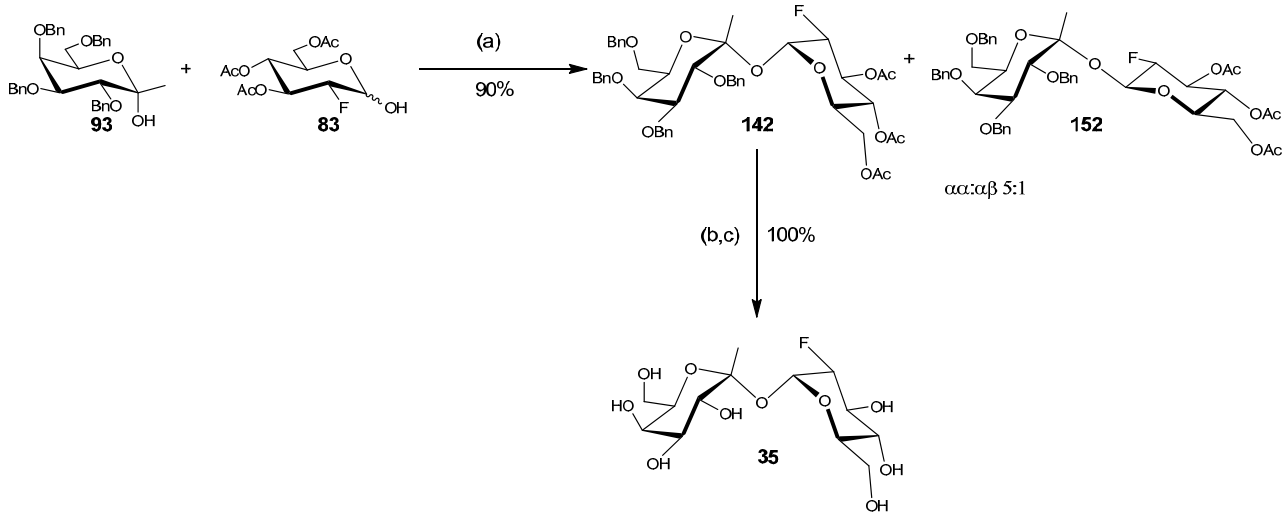
Entry	Donor	mmol	Acceptor	mmol	Catalyst	T ˚C	Time	Yield	α,α: α,β
1	93	0.13	91	0.09	BiOTf	0	3 h	56%	7:3
2	102	0.1	91	0.09	TMSOTf	-78	1 h	95%	3:7
3	99	0.07	94	0.07	TMSOTf	-78	30 min	68%	2:3
4	85	0.17	94	0.16	TMSOTf	-40	4 h	25%	1:1
5	102	0.31	105	0.58	TMSOTf	–78	45 min	30%	3:7
6	99	0.04	105	0.09	TMSOTf	–78	45 min	95%	3:2
7	85	0.24	84	0.25	TMSOTf	–40	3 h	93%	7:1
8	85	0.14	83	0.22	TMSOTf	–40	2 h	82%	6:1
9	93	0.32	84	0.42	TMSOTf	–40	30 min	88%	5:1
10	93	0.14	83	0.16	TMSOTf	–40	1 h	90%	6:1
11	85	0.21	137	0.18	TMSOTf	–40	3 h	52%	4:1
12	85	0.31	138	0.38	TMSOTf	–78	15 min	81%	6:1
13	85	0.7	139	0.04	TMSOTf	–78	3 h	76%	1:6
14	85	3.17	138	3.5	TMSOTf	-40	2.5 h	46%	5:1
15	86	0,071	84	0.082	TMSOTf	-78	2 h	92%	1:25
16	140	0.055	84	0.055	TMSOTf	-40	30 min	73%	6:1
17	141	0.19	84	0.29	TMSOTf	-40	30 min	61%	α only
18	85	0.29	112	0.27	TMSOTf	-40	5 min	76%	α only



The syntheses of the methyl galacto-gluco-trehalose analogue **34** as well as the methyl 2-F-2-deoxy-galacto-gluco-trehalose analogue **35** proceeded smoothly, affording comparable yield to the gluco compounds (**Scheme 2.31** and **Scheme 2.32)**. Stereoselectivity for glycosylations with the galacto donor were slightly more modest (5:1), but still sufficient for synthetic purposes. Hydrogenation of both **142** and **143** required long reaction times and excess palladium.



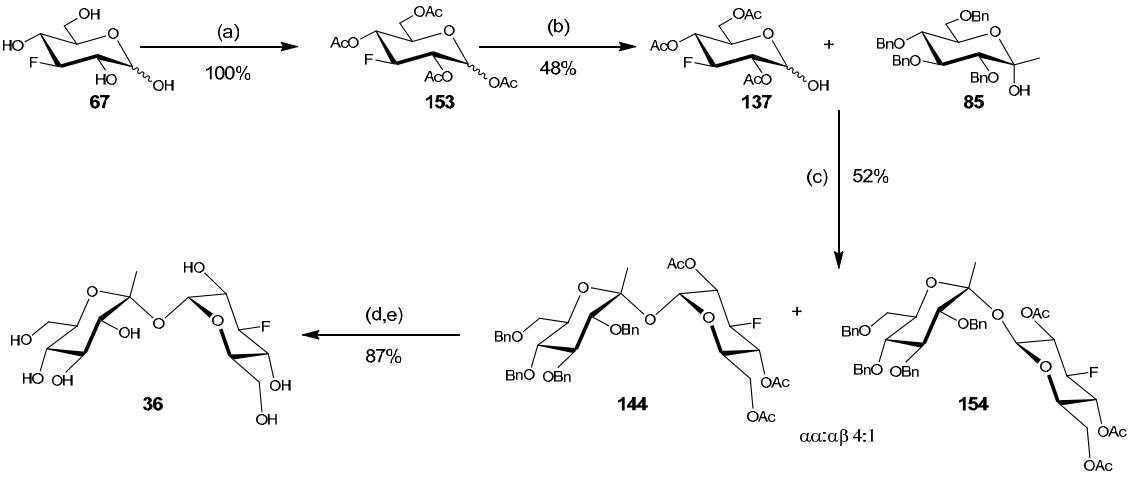
Scheme 2.31. Synthesis of 34. (a) TMSOTf, 4 Å molecular sieves, anhydrous DCM, −40 °C, 30 min (b) NaOMe, anhydrous MeOH, RT, 1 h (c) H_2 , $Pd(OH)_2$/C, basic alumina, RT, 72 h.



Scheme 2.32. Synthesis of 35. (a) TMSOTf, 4 Å molecular sieves, anhydrous DCM, -40 °C, 1 h (b) NaOMe, anhydrous MeOH, RT, 3 h (c) H_2 , $Pd(OH)_2$/C, basic alumina, RT, 96 h.

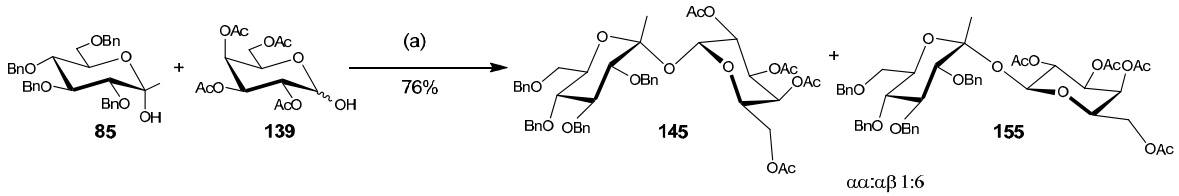
Methyl 3-F-3-deoxy-trehalose analogue **36** was synthesized from 3-F-3-deoxy-D-glucose **67** in moderate yield and stereoselectivity (**Scheme 2.33**). **67** was obtained as a kind gift from Dr. François D'Hooge. **67** was acetylated without problem, utilizing pyridine and acetic anhydride and then the anomeric acetate was selectively removed with hydrazine acetate to afford **137** in a somewhat low yield. Glycosylation with **137** proceeded in a fairly low yield (52%) and only a 4:1 ratio of αα:αβ. Typically the ratios of selectivity were found to be closer to 6:1 and one would think that the 2-OAc would have a directing influence on the

glycosylation. Potentially, the 3-F counteracts this effect. Deprotection proceeded without problems, utilizing the H-Cube® Pd/C cartridge at 70 Bar for the hydrogenation.



Scheme 2.33. Synthesis of 36. (a) Acetic anhydride, pyridine, RT, (b) Hydrazine acetate, DMF, RT (c) TMSOTf, 4 Å molecular sieves, DCM, –40 °C, 5.5 h (d) NaOMe, anhydrous MeOH, RT, 1.5 h (e) H Cube® 10% Pd/C cartridge at 70 bar, 25 °C, 1 h.

Methyl gluco-galacto-trehalose analogue **145** was also synthesized to look at the anomeric selectivity when the galacto moiety is in the acceptor rather than the donor. This compound was also synthesized with a ^{14}C label at the NIH (described in **Section 2.2.11.3** below) and the unlabeled compound was employed to confirm synthesis of ^{14}C analogue. Both the α,α-**145** and α,β-**155** products were obtained to purity. Reaction proceeded in 76% yield of **145** and **155** with 20% recovery of donor **85**, affording the reverse of the usual stereochemistry, primarily the α,β-compound as indicated by NMR coupling constants (α,β:α,α 6.6:1). Preference for the α,β-compound was found for the glycosylation with ^{13}C-methyl-glucose **86** as well (**Scheme 2.34**). It is unclear why this sudden switch in stereochemistry occurred, it could be related to some change in the solvent or the particular batch of TMSOTf, as both of these compounds were synthesized concurrently.

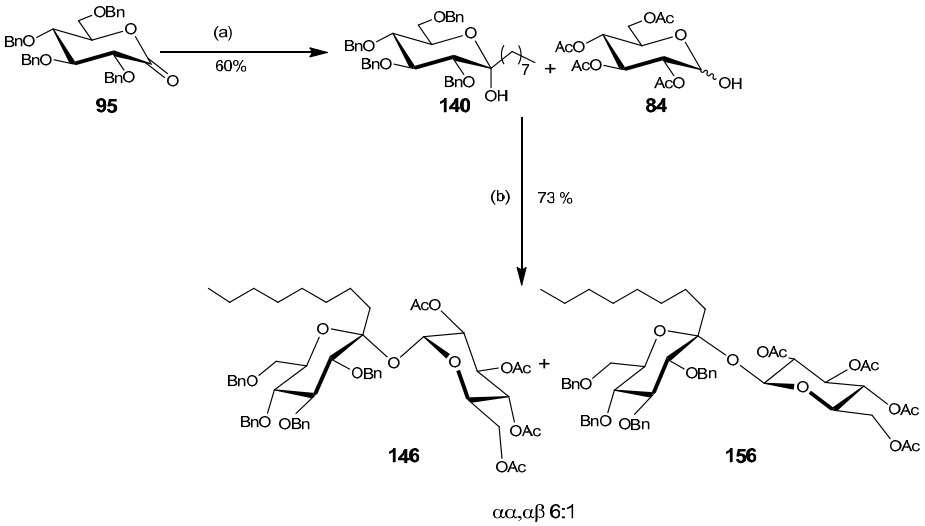


Scheme 2.34. Synthesis of methyl-galacto-gluco-trehalose 145. (a) TMSOTf, -78 ˚C, 4Å molecular sieves, 2 h, 6.6:1 α,β: α,α.

2.2.11.1 C18-trehalose synthesis

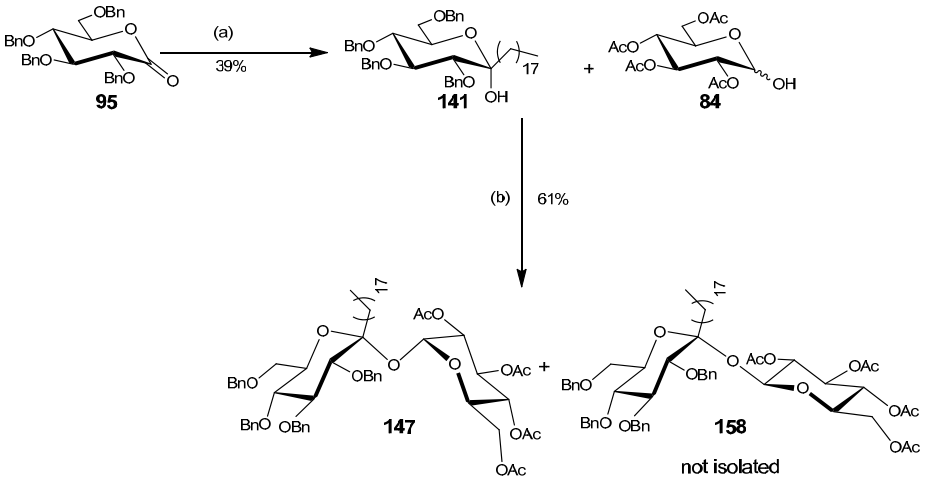
There currently exists interest in microarray carbohydrate display for the purpose of screening for novel binding proteins[34,35]. To facilitate binding studies trehalose with a C18 lipid chain attached from the anomeric hydroxyl was synthesized, utilizing the previously discussed ketoside glycosylation technique.

Previous work in this area had attempted to synthesize the lipidated amino linked trehalose derivatives, utilizing the Kochetkov reaction, which treats the sugar with 50 times excess of ammonium bicarbonate for 6 days at room temperature[36]. However, this approach had been plagued by poor stereoselectivity and low yields. Glycosyl amines are relatively unstable[28], which makes the stereocontrol at the anomeric position of sugar difficult to achieve due to the mutarotation of the hemiaminal. Syntheses of the α,α-1,1-linked disaccharides, and purification from the mixture of α,α-, α,β- and β,β-conformers was found by others in the lab to be difficult. Ketoside trehalose was proposed as an alternative method for the attachment of trehalose to binding arrays, while still preserving the native trehalose linkage, which could be essential for binding.



Scheme 2.35. Synthesis of C8-trehalose analogue 146 (a) *n*-Octylmagnesium chloride, THF, -78 °C, 30 min (b) TMSOTf, 4 Å molecular sieves, anhydrous DCM, −40 °C, 30 min

In order to develop this chemistry we first conducted a test reaction, where the C8-glucose compound **140** was synthesized from octylmagnesium chloride (**Scheme 2.35**). When this reaction was conducted with methyllithium only the α-heptulose was generated. However, with the chloride, the C8-glucose was obtained as a mixture of α and β products, from which the α product was separated. The α donor was glycosylated in 88% yield with **84** to afford a 6:1 separable mixture of the α,α- and α,β-products. This synthesis with **140** confirmed that the ketosides with longer lipid chains would glycosylate comparably to the methyl ketosides.



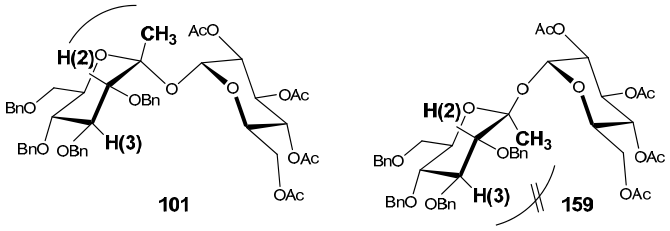
Scheme 2.36. Synthesis of C18-trehalose analogue 147. (a) *n*-Octadecylmagnesium bromide 157, THF, -78 °C, 3 h (b) TMSOTf, 4 Å molecular sieves, anhydrous DCM, −40 °C, 30 min.

C18-glucose compound **141** was synthesized from gluconolactone and 1-bromo-octadecane, which was first converted into the corresponding octadecyl Grignard reagent **157**. **141** was then glycosylated to produce the C18-trehalose analogue **147** (**Scheme 2.36**). The Grignard **157** was generated from magnesium turnings following the standard procedure, but problems were encountered with Grignard solubility. It precipitated out of solution at room temperature and was particularly insoluble at the lower temperatures of the addition to the lactone. As a result of this insolubility, the lactone addition stopped before completion. With the addition of excess Grignard the reaction proceeded in a modest 39% yield of **147** and **158**. Purification proved difficult, as removal of the excess remaining lipid required multiple rounds of column chromatography.

The synthesis of longer alkyl ketosides revealed for the first time that ketoside glycosylations tolerate the inclusion of much longer anomeric modifications. The addition of lipids to the trehalose C-1 could allow for the generation of trehalose-resins and microarrays for affinity capture and purification of trehalose binding proteins.

2.2.11.2 NOE assignment of stereochemistry

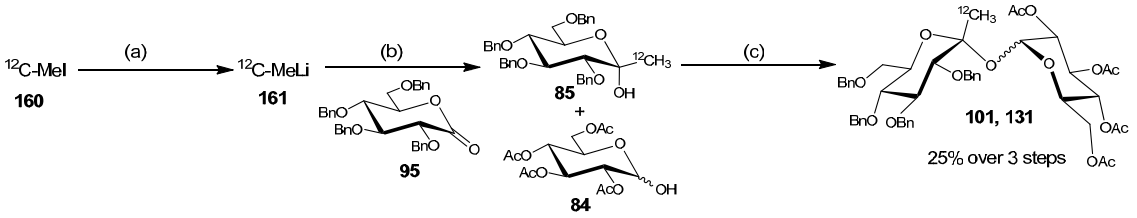
NMR analysis was used to unambiguously assign the stereochemistry of the ketoside derivatives as α,α. Stereochemistry for the glucose half of the disaccharide was confirmed to be α by the coupling constants of the anomeric proton. Due to the absence of anomeric proton and corresponding coupling constants for the ketoside half of the molecule, Nuclear Overhauser Effect (NOE) analysis was employed to confirm the stereochemistry of the glycosidic linkage as α. NOE between H-2 and the methyl group (**Scheme 2.37**) was observed. No NOE between the methyl group and H-3 was observed. This experiment confirmed that this trehalose analogue is the α,α- linked glycoside. Likewise, NOE experiments were conducted on the C18-trehalose, which confirmed that this compound was also α,α. Comparison of the coupling constants from the α,α-gluco-gluco methyl ketoside analogue **101** to all other ketosides synthesized, allowed for unambiguous assignment of the compounds as α,α.



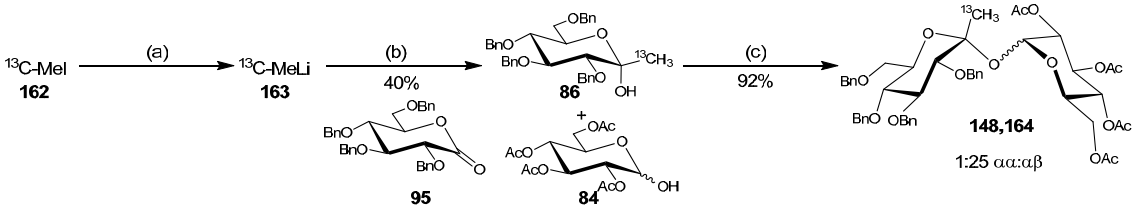
Scheme 2.37. NOE observed between methyl protons and H-2, but not between methyl group and H-3 confirms that the stereochemistry of the ketoside anomeric center is α.

2.2.11.3 Isotopic ^{13}C and ^{14}C labeling of trehalose

Based on previous work[37,38], it was thought that ^{14}C-methyllithium could be generated *in situ* from *tert*-butyllithium (*t*-BuLi) and methyl iodide. Previously, groups had used *n*-butyllithium (*n*-BuLi), but for this work *t*-BuLi was chosen because of concerns that *n*-BuLi would add into the lactone **95**, generating unwanted side products. One complication of this approach was that prior lithium halogen exchanges had been conducted on a much larger scale than was available for this radiochemical synthesis.



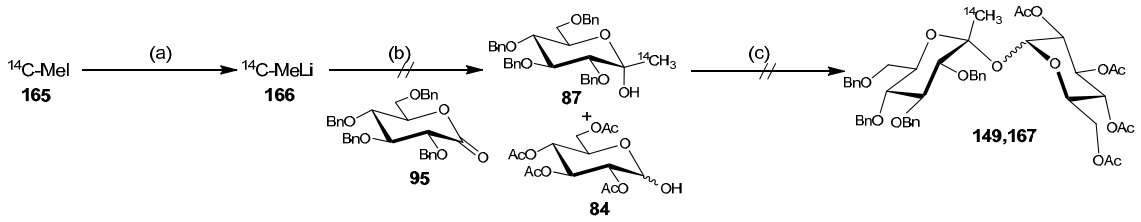
Scheme 2.38. A test synthesis with unlabeled methyl iodide was carried through in 25% overall yield, based on the methyl iodide. (a) *t*-BuLi 2.2 eq., THF, -78 ˚C, 15 min, (b) 95, THF, -78 ˚C (c) TMSOTf, 4 Å molecular sieves, DCM, -40 ˚C, 30 min.



Scheme 2.39. Synthesis of ^{13}C-methyl-trehalose analogues 148 and 149. (a) *t*-BuLi 2.2 eq., THF, -78 ˚C, 15 min, (b) 95, THF, -78 ˚C (c) TMSOTf, 4 Å molecular sieves, DCM, -40 ˚C, 2 h.

To validate the proposed ^{14}C-trehalose synthesis, the lithium halogen exchange was optimized on both ^{12}C- and ^{13}C-methyl iodide, **160** and **162** (**Scheme 2.38** and **Scheme 2.39)**. During optimization of the lithium halogen exchange, several observations were made regarding the finicky nature of this reaction. The order of reactant addition and whether the methyl iodide was added neat or as a solution dramatically affected reaction success. Reactions worked most consistently and with the best yield when neat methyl iodide was added to the *t*-BuLi, consistent with the previously mentioned protocols. When the order of addition was reversed and the *t*-BuLi was added to neat methyl iodide the reaction still proceeded, although not as consistently. It was found that dissolving the methyl iodide in any amount of solvent reduced the yield of the methyl-glucose compound **85**. Potentially trace water in the solvent could quench the methyllithium upon formation. It was also found that the concentration of lactone did not make a significant difference upon addition. Likewise, the amount of time for the exchange did not seem to be significant. Lithium halogen exchange appeared to take place almost instantaneously, judging by color change of the solution, although prior reports had suggested leaving the reaction for significant periods of time and warming to room temperature. **95** could be added immediately after the methyl iodide and formation of **85** was observed.

^{13}C-methyl-trehalose analogue **148** (**Scheme 2.39**) was synthesized utilizing the optimized lithium halogen exchange methodology on a larger scale. These reactions proceeded smoothly and the desired ketoside donor **87** was obtained in 40% yield. Upon glycosylation with TMSOTf, utilizing the standard conditions, **149** and **167** were obtained in 82% yield. In a departure from previous reactions, the α,β product rather than the α,α product was favoured in a 25:1 ratio, as indicated by NMR integration and coupling constants. Potentially the isotopic label affected the stereoelectronics. Alternatively, something else in the reaction mixture (solvent, TMSOTf, molecular sieves) may have altered the stereochemistry, as the synthesis of methyl gluco-galacto-trehalose analogue **145** (**Scheme 2.34**), was troubled by the same reverse in stereoselectivity.



Scheme 2.40. Unsuccessful attempted synthesis of ^{14}C-ketoside based trehalose analogues. (a) *t*-BuLi 2.2 eq., THF, -78 ˚C, 15 min, (b) 95, THF, -78 ˚C (c) TMSOTf, DCM, -40 ˚C, 3 h.

The synthesis of ^{14}C-methyl glucose donor **87** was then attempted, utilizing a lithium halogen exchange between ^{14}C-MeI and *t*-BuLi (**Scheme 2.40**). The reaction was challenged by small scale and unknown purity of the radioisotope.

At the NIH the identical conditions to those optimized at Oxford were tested with ^{12}C-methyl iodide on the same lactone **95** to ensure that the chemistry would translate, prior to conducting the radiochemistry. No product was initially detected and instead starting material was completely consumed and a large baseline spot (Petrol:EtOAc, 2:1) was produced (**Figure 2.3**). Upon LCMS analysis of this spot, it appeared to contain multiple products, which moved in pure ethyl acetate via TLC. The stoichiometry of methyl iodide to *t*-BuLi, reaction time, order of addition, concentration, rate of addition were all adjusted. However, no changes produced any product. Even when the THF was re-distilled, no product was detected. Distilled diethyl ether was evaluated and under identical reaction conditions to those used with THF, product was detected. The successful switch from THF to ether suggested that some uncharacterized difference between the THF in Oxford and NIH may have explained the increased side product formation observed at the NIH. Reaction proceeded optimally with addition of 2 equivalents of *t*-BuLi at 0.34 M. Higher concentrations increased side product formation. Upon much trial acceptable reaction conditions were developed for the NIH lithium halogen exchange. These featured addition of a solution of *t*-BuLi to neat the methyl iodide, following which the lactone was added (see **Appendix 1** for full details).

With optimized reaction conditions, the radiochemistry was attempted. The shape of source vial provided an additional problem. Initially, addition of the methyl iodide to the *t*-BuLi via syringe was planned, because this form of addition afforded the most consistent results. However, the methyl iodide was sold in source vials that were about 30 cm long and completely sealed. Two units of 1 mCi ^{14}C-methyl iodide at 1mmol/55mCi were obtained,

meaning that approximately 0.012 mmol of methyl iodide was in each source vial. This small amount of material could not be removed from the long source vial. It was decided to reverse the order of addition and add *t*-BuLi to the source vial. An additional complication was that the methyl iodide was stabilized over large amounts of copper in order to prevent degradation to iodine upon exposure to light. It was not known what effect the copper had on the reaction.

The ^{14}C-halogen exchange was attempted twice, followed by addition of gluconolactone **95**, and both attempts proved unsuccessful. The reaction was conducted in the source vial. The methyl iodide was first condensed to the bottom of the vial on dry ice. Then, a small stir bar was added. However, the vial was too narrow at the bottom to allow for proper stirring. With the optimized conditions mentioned above, no white precipitate, indicating product **87** formation, was observed. By TLC there may have been trace **87**. Also, a spot higher than the product spot was detected (**Figure 2.3**). However, the reaction mainly produced a large baseline spot.

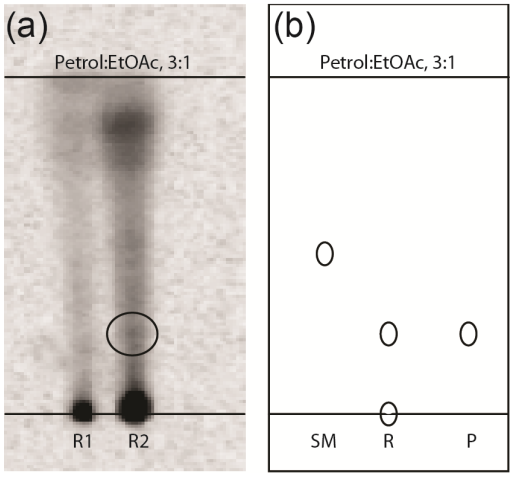
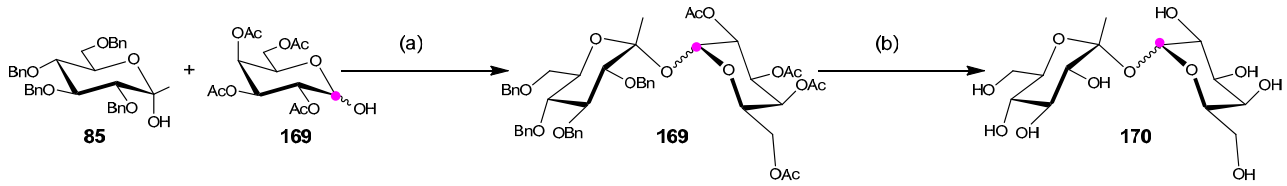


Figure 2.3. (a) Radiographic images of the TLC plate of both reactions (R1 and R2) conducted. Nearly all radioactivity is concentrated in the baseline spot. The circle indicates where the product should have been. Little radioactivity is detected in that region. (b) Model TLC indicating where the product should be observed. SM-starting material, R-Reaction, P-Product.

Very little radioactivity was detected via TLC in the region where the product **87** was expected. Nearly all radioactivity was found in the baseline spot (**Figure 2.3**). Even though little to no product was detected, purification by HPLC was attempted to determine whether there was trace product. Attempts at purification of the crude reaction mixture afforded none of the desired product (see **Appendix 1**). It is unclear as to why this reaction did not proceed.

Potentially the copper could have interfered with the exchange, forming some cuprate-like species. Water may have contaminated the source vial, leaving it less than completely dry. The lack of stirring could have affected the reactivity. The vial may have contained far less methyl iodide than indicated could have been contaminated by unknown impurities.

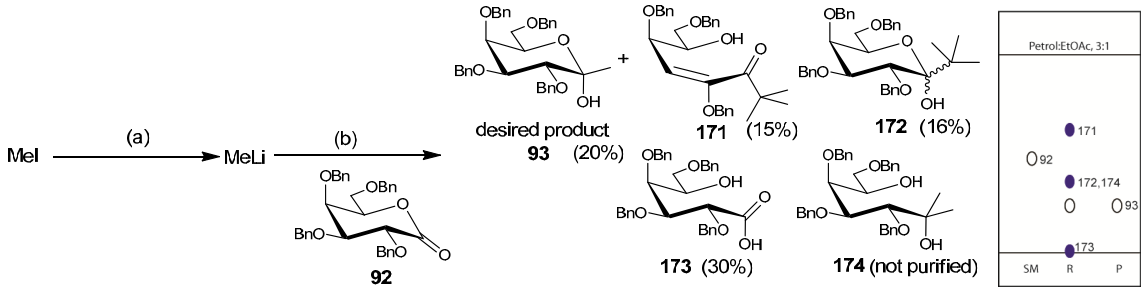
As the lithium halogen exchange failed, it was thought that a ketoside analogue of trehalose could be synthesized (**Scheme 2.41**), utilizing commercially available ^{14}C-galactose, doped with unlabeled galactose. Yields were not possible because the radioactivity prevented the use of reduced pressure to remove solvents, but product identity of the protected disaccharide **160** matched that of **145**. Deprotection was complicated by the stubborn hydrogenation step and a lack of resources (NMR, ESI MS, etc) at the NIH for full characterization of the final radiolabeled compound. The final compound was assayed for uptake into *Mtb* (see **Chapter 3**) and showed no appreciable uptake. This could be due to an incomplete deprotection, poor stereoselectivity of the galacto glycosylation (see **Scheme 2.34**) or inhibition of uptake by the galacto configuration.



Scheme 2.41. Synthesis of ^{14}C-galacto-trehalose compound 170 (a) TMSOTf, -40 ˚C, DCM 3 h (b) NaOMe, MeOH, 2 h (c) H_2, Pd $(OH)_2$/C, RT, 48 h (pink dot indicates ^{14}C label).

Upon returning to Oxford, further characterization of the nature of all side products of the methyllithium reaction was attempted in order to ascertain why the reaction did not succeed. The side products were characterized by ESI and NMR and three side products were isolated to enough purity for identification (**171, 172, 173** in **Scheme 2.42**). Despite its lack of nucleophilicity, the NMR spectra and mass revealed that the *t*-BuLi had added into the lactone to produce an anomeric mixture of both the axial and equatorial *t*-butyl groups. An elimination product where the *t*-BuLi both added to the lactone and deprotonated the lactone, resulting in elimination of one benzyl group was identified. This elimination product has been previously identified[39]. Finally, the carboxylic acid **173** was isolated. A literature review

revealed that open chain ketosides are remarkably stable and have been shown to remain open under acid workup[40], so potentially, the carboxylic acid could have been produced upon workup. The double addition product **174**, where two methyl units had added was detected by ESI MS when its spot was scraped off of a TLC plate. However, this product was not detected after flash chromatography.



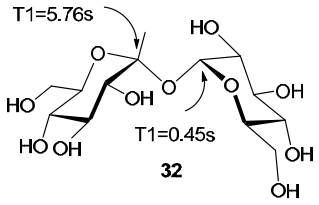
Scheme 2.42. Side products of lithium halogen exchange. (a) *t*-BuLi 2.2 eq., THF, -78 ˚C, 15 min, (b) 92, THF, -78 ˚C. TLC (petrol:EtOAc, 3:1) Numbers on TCL indicate side product produced. Sm-starting material, R-reaction, P-product.

2.2.12 Alternative uses for isotopically labeled trehalose: MRI

As the synthesis of ^{14}C-trehalose compounds proved unsuccessful, other methods for the determination of the *in vivo* uptake and labeling of trehalose analogues into mycobacteria were surveyed. ^{13}C NMR was identified as an alternative and/or complementary technique for the ^{18}F PET. Previously, hyperpolarized magnetic resonance imaging (MRI) has been utilized to measure the pharmacokinetics[41] of potential drug compounds, and to measure the metabolism of compounds such as pyruvate and glucose[42] by selected organs. MRI is a non-invasive imaging technique that has gained widespread clinical usage. One of the major drawbacks of MRI is its low sensitivity relative to other techniques (such as PET). Hyperpolarization enhances the sensitivity of MRI by employing isotopically labeled probes, which allow for polarization of nuclear spin states of isotopes, such as ^{13}C, thereby greatly increasing the labeled compounds NMR signal.

^{13}C-methyl-trehalose analogue **148** (**Scheme 2.39**), synthesized with the lithium halogen exchange methodology, was initially proposed as a possible probe for hyperpolarized MRI. A long T1 relaxation time is required for hyperpolarized MRI. It was discovered that carbon centers attached to hydrogen have a much shorter relaxation time than quaternary carbons

and carbonyl carbons. The label in the primary methyl group of ^{13}C-methyl-trehalose was not suitable for a long T1 half-life. Methyl-trehalose with the ^{13}C label on the quaternary anomeric carbon was considered as an alternative probe. This carbon should act as a masked carbonyl and should have a longer relaxation time. Prior to synthesis of this analogue, the T1 relaxation time of the selected carbon was determined in compound **32**. The T1 experiment revealed that that the relaxation time of the target carbon (**Scheme 2.43**) was found to be 5.76 s. In comparison, the relaxation time of the neighboring anomeric carbon is 0.45 s.



Scheme 2.43. T1 relaxation times of 32.

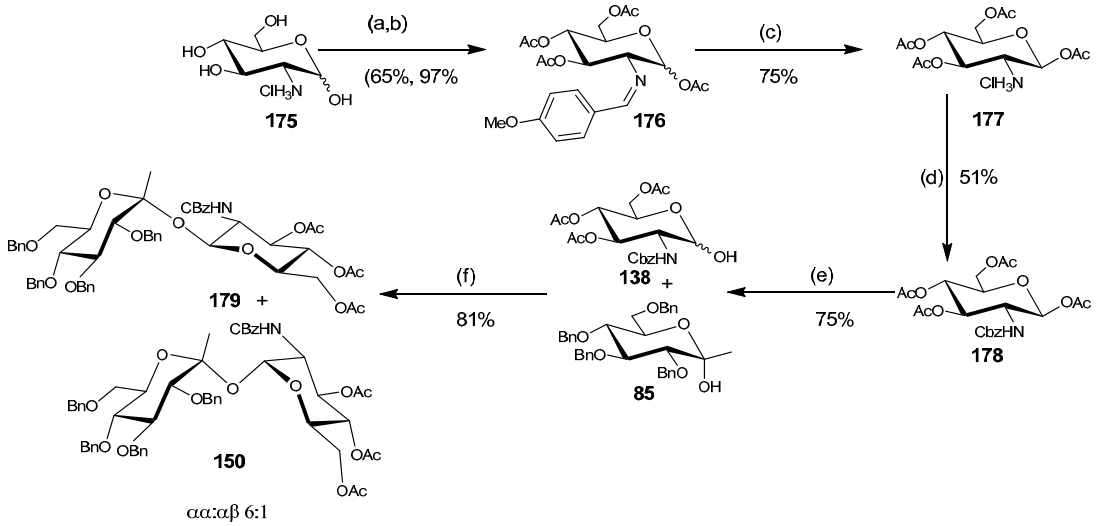
These results suggested that the introduction of the quaternary center into the trehalose backbone did lengthen the relaxation time of the target carbon 10 fold. However, 5.6 s is still far too short for this compound to be useful in hyperpolarized MRI, which requires a T1 time of upwards of 30 s. The T1 relaxation time of **32** suggested that it would not serve as a useful hyperpolarized MRI agent

2.2.13 Modification of 2-aminotrehalose analogue 37

Both ^{14}C- and ^{13}C-isotopic labeling strategies had failed to yield the desired trehalose uptake probe. We turned our attention towards an alternative to directly labeling the sugar scaffold: appending the label off of the trehalose scaffold. An example of this strategy was the attachment of a radioactive prosthetic group (^{14}C-benzyl or ^{14}C-methyl-ether) to the trehalose instead of direct incorporation of the radiolabel into the backbone as had been attempted previously. This approach offered an ease of synthesis, but was challenged by the presence of bulky groups possibly hindering the trehalose's ability to serve as a substrate of the Ag85 enzymes.

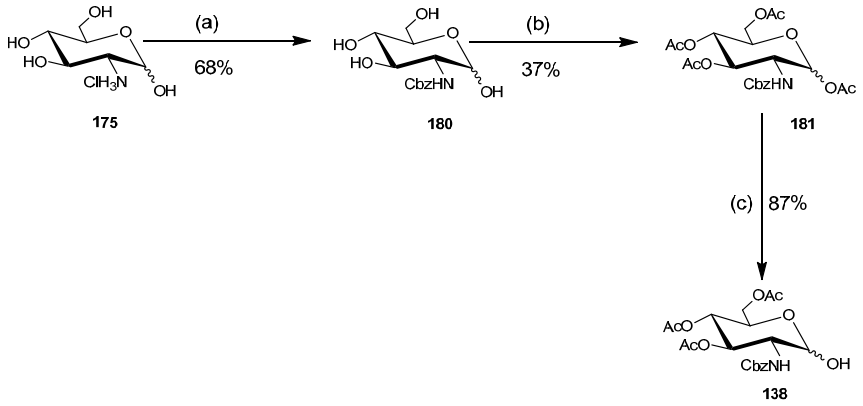
Structural analysis (**Figure 2.1**) of trehalose bound to Ag85B highlighted 2- positions (C-2 and C-4') on trehalose that were directed outward towards solvent and that might tolerate substitutions and yet retain activity as substrates for Ag85. It was envisioned that a

fluorescent, radioactive or biotin group could be attached to 2-aminotrehalose, allowing for the late stage modification of the aminotrehalose analogue by peptide coupling or reductive amination. The crystal structure suggests that these bulky groups off of the 2- position would be situated outside of the active site and would not dramatically interfere with enzymatic activity. 2-amino-2-deoxy-methyl-trehalose analogue **37** was synthesized through the ketoside glycosylation strategy, by coupling **85** with the carboxybenzyl (Cbz) protected 3,4,5-tri-*O*-acetyl-D-glucosamine **138** (**Scheme 2.44**).



Scheme 2.44. Synthesis of 150. (a) Anisaldehyde, 1M NaOH, RT, 10 min, 0 °C, 1 h (b) Acetic anhydride, anhydrous pyridine, RT, 24 h (c) 5 M HCl, acetone, 10 min (d) Benzyl chloroformate, Na_2CO_3, DCM, RT, 1.5 h (e) benzylamine, THF, RT, 72 h (f) TMSOTf, 4 Å molecular sieves, anhydrous DCM, –40 °C, 15 min.

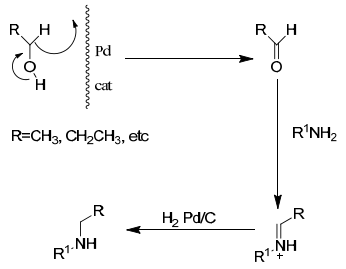
Synthesis of **37** proceeded smoothly until the final hydrogenation step. The standard reaction sequence[43] to synthesize **138** was utilized to afford the desired acceptor in good yield. Imine formation with *para*-anisaldehyde proceeded rapidly under biphasic conditions. The acetylation reaction required chilling on ice and slow addition of the acetic anhydride to prevent the formation of *N*-acetyl-glucosamine by acid catalyzed hydrolysis of the imine and subsequent acetylation. The glycosylation with **85** afforded the desired protected 2-amino ketosides **150** and **179** in good yield and stereoselectivity (**Scheme 2.44**).



Scheme 2.45. Alternate synthesis of 128. (a) Benzyl chloroformate, Na_2CO_3, DCM, RT, 1.5 h (b) acetic anhydride, anhydrous pyridine, RT, 24 h (c) benzylamine, THF, RT, 72 h.

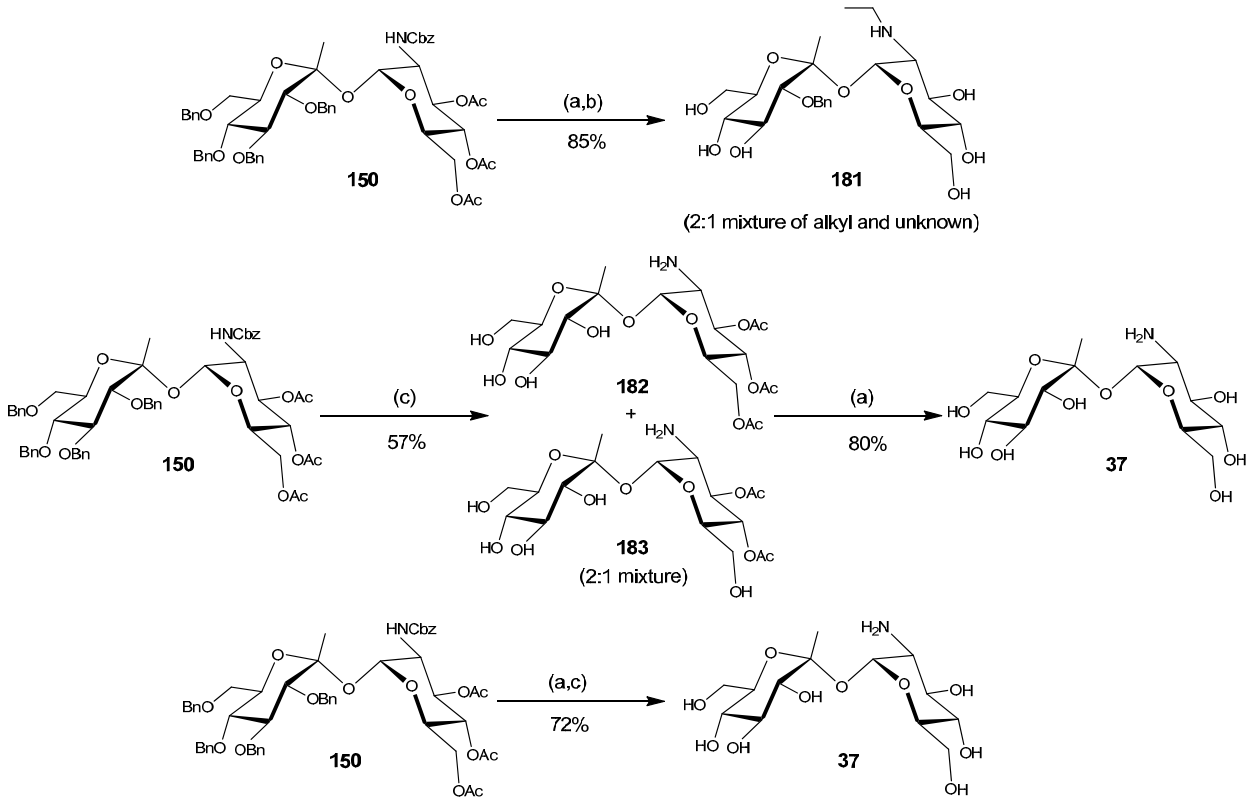
In order to streamline acceptor synthesis and avoid imine hydrolysis, a more rapid, three step synthesis was employed: acylation of glucosamine with carboxybenzyl chloride to produce **180** and was followed by peracetylation and anomeric deprotection, utilizing benzylamine to yield **138**. This reaction series proceeded smoothly in 22% yield over three steps (**Scheme 2.45**). This acceptor synthesis offered the advantage of brevity, but was less crystalline and did not yield exclusively the β-anomer, as in the previous reaction series.

With protected **150** compound in hand, we then turned our attention towards the deprotection steps. With the prior syntheses, **150** was first deacetylated and then submitted to hydrogenation with $Pd(OH)_2/C$ and basic alumina in ethanol. However, when this protocol was followed, the reductive amination product **181**, contaminated with an unknown side product (**Scheme 2.46** and **Scheme 2.47**) was obtained in good yield in lieu of the desired free amine. This transamination catalyzed by the palladium in the hydrogenation mixture (**Scheme 2.46**) has been reported previously[44].



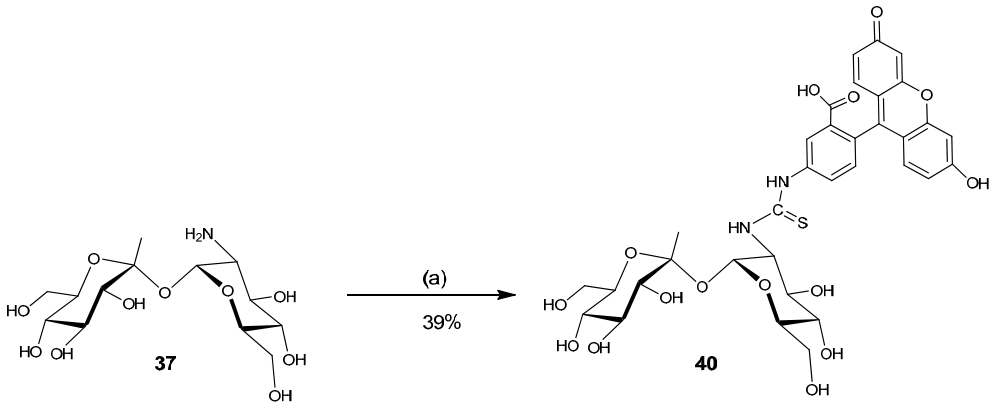
Scheme 2.46. Proposed mechanism of previously reported reductive amination during hydrogenation of protected amines[44].

It was found that use of trifluoroethanol (TFE) as solvent eliminated the reductive amination problem. Concurrently it was decided to reverse the order of deprotection and debenzylate prior to deacetylation. The basic conditions of the hydrogenation resulted in a partial deacetylation of the compounds (**Scheme 2.47**). Full deacetylation was possible, confirming that amine was not acetylated by possible acetate migration, during the hydrogenation. Subsequently, deacetylation followed by hydrogenation in TFE was employed to afford **37** in good yield.



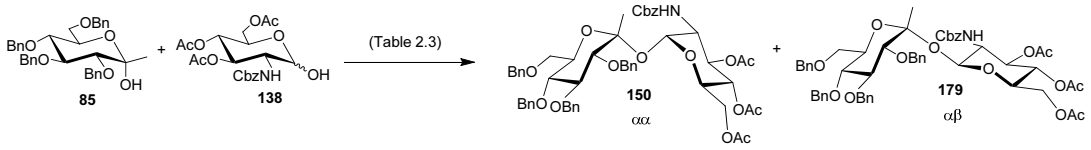
Scheme 2.47. Deprotection of 150. (a) NaOMe, MeOH, RT, 1-3 h (b) $Pd(OH)_2/C$, H_2, RT, Ethanol, 48-72 h, (c) $Pd(OH)_2/C$, H_2, RT, trifluoroethanol, 48-72 h.

The free amine in **37** was then coupled to fluorescein isothiocyanate on a small scale in modest 39% yield to afford fluorescein trehalose compound **40** (FITC-Tre) (**Scheme 2.48**). **40** was purified by HPLC, utilizing the Synergi Hydro™ semi-prep column. Initial HPLC in the presence of NH_4OH, afforded two peaks for **40** with the same mass, protentially due to separation of the protonate and deprotonated carboxylate forms on the column. This problem was resolved when purifaction employed a 0.1% TFA gradient.



Scheme 2.48. Synthesis of Fluorescein-trehalose 40. (a) Fluorescein-isothiocyanate, $NaHCO_3$ buffer (pH 9), RT, 16 h.

The synthesis of FITC-Tre was scaled up in order to obtain ca. 100 mgs of the product for biological studies. Three key steps required optimization to obtain sufficient product. First, the glycosylation of **85** with **138** had to be scaled up. On a large scale yields of this glycosylation decreased from 85% to 46% and reaction seemed to stall prior to completion (**Scheme 2.49** and **Table 2.3 entries 4-6**). This reduction is consistent with prior group members' results that suggested when glycosylations are scaled up, the yield often decreases. The 46% yield of the glycosylation on a 3 g scale (**Table 2.3 entry 6**) was sufficient to generate large quantities (1.8 g) of the protected disaccharide, which could then be easily deprotected and the amine coupled to a variety of amine reactive groups. High recovery of the donor **85** and acceptor **138** allowed for multiple glycosylation cycles to generate large quantities of the desired product. Also notable is that the α,α and α,β are very easy to separate by column chromatography.



Scheme 2.49. Optimization of the synthesis of 150 detailed in Table 2.3.

Table 2.3. Optimization of the glycosylation between 85 and 138. Cat-catalyst (TMSOTf), RSM-recovered starting material.

Entry	Donor	mmol	Acceptor	mmol	Cat. eq.	T ˚C	Time	Yield	Donor RSM	Acceptor RSM	α,α: α,β
1	85	0.31	138	0.38	0.17	–78	15 min	81%	-	-	6:1
2	85	0.11	138	0.2	0.98	-40	3 h	61%	32%	-	5:1
3	85	1.6	138	1.63	0.23	-40	2.5 h	52%	33%	20%	7:1
4	85	3.17	138	3.5	0.17	-40	2.5 h	46%	10%	18%	5:1
5	85	3.07	138	3.12	0.4	-40	7 h	36%	25%	30%	2.75:1
6	85	6.67	138	7.74	0.15	-40	7 h	46%	64%	-	4:1

Once the glycosylation was providing reasonable yields, the deacetylation and hydrogenation of **150** were optimized (**Table 2.4**). For these deprotections it was decided to revert to the original deprotection order, first deacetylate and then hydrogenate. Deacetylation proceeded smoothly in anhydrous methanol with sodium methoxide. Hydrogenation was then attempted, utilizing the H-Cube® Pd/C cartridge. Initial attempts were carried out in trifluoroethanol (TFE). These reactions proceeded extremely slowly, even with the addition of formic acid at 70 bar and gave low yields (59%) over two steps. We discovered that product **37** was precipitating out on the cartridge in the TFE. The addition of water to the hydrogenation mixture (water:TFE, 1:1) increased the yield and decreased reaction time, greatly accelerating the reaction to roughly 8-10 h at 70 bar for 200 mgs protected ketoside, circulating through the H-Cube®. It was also found that **150** and, presumably other ketosides, tolerated the addition of formic acid to further catalyze the hydrogenation, in contrast to prior reports regarding the acid labile nature of ketosides[14].

Table 2.4. Reaction optimization for deprotection of 150. Deact-Deacetylation, Hydrog-hydrogenation. TFE-trifluoroethanol.

Entry	Mg 150	Mmol 150	Solvent	Catalyst for	Time	Deprotection order	Product/side products	Yield
1	45	0.046	ethanol	$Pd(OH)_2/C$, basic Alumina	48 h	1) Deacet 2)Hydrog	**181**	85%
2	49	0.05	TFE	$Pd(OH)_2/C$, basic Alumina	48 h	1) Hydrog	**182**	100%
3	107	0.109	TFE	$Pd(OH)_2/C$, basic Alumina	72 h	1) Hydrog	**182**	66.4%
4	202	0.2	TFE	$Pd(OH)_2/C$, basic Alumina	4 days	1) Hydrog	**182, 183**	55%
5	351	0.359	TFE	$Pd(OH)_2/C$, basic Alumina	4 days	1) Deact 2) Hydrog	**37**	67%
6	300	0.307	TFE	Pd/C H-Cube® 70 bar	3 days	1) Deact 2) Hydrog	**37**	59%
7	90	0.106	TFE:H_2O (1:1) Formic acid	Pd/C H-Cube® 70 bar	10 h	1) Deact 2) Hydrog	**37**	100%
8	290	0.34	TFE:H_2O (1:1)	Pd/C H-Cube® 70 bar	15 h	1) Deact 2) Hydrog	**37**	100%
9	200	0.205	TFE:H_2O (1:1)	Pd/C H-Cube® 70 bar	30 h	1) Deact 2) Hydrog	**37**	100%
10	1.08 g	1.22	TFE:H_2O (1:1) Formic acid	Pd/C	48 h	1) Deact 2) Hydrog	**37**	91%

With sufficient **37** in hand, we turned our focus towards optimization of the coupling with fluorescein. After extensive coupling work (**Table 2.5**), it was found that reaction yields were optimal in a mixture of bicarbonate buffer and acetonitrile with heating over a short time period (2 h). Longer reaction times, higher pH, and larger reaction volume did not increase the yield.

Table 2.5. FITC-Tre coupling optimization. Sodium carbonate buffer $NaHCO_3$ was 75mM at pH 9 unless otherwise indicated. FITC-Fluorescein isothiocyanate. RT- room temperature.

Entry	Mmol 37	Eq. 37	Mmol FITC	Eq. FITC	Solvent	Vol.	Time	T °C	Yield
1	0.015	1.15	0.013	1	$NaHCO_3$	3 mL	O/N	35 °**C** for 30 min then RT	39%
2	0.045	1	0.064	1.4	$NaHCO_3$	5 mL	72 h	RT	48%
3	0.073	1	0.116	1.6	$NaHCO_3$	10 mL	48 h	RT followed by 40 °**C** for 4 h	28%
4	.045	1	0.082	1.82	$NaHCO_3$ pH 12	2 mL	24 h	RT	18%
5	.12	1	0.174	1.55	$NaHCO_3$	5 mL	24 h	RT followed by 40 °**C** for 4 h	47%
6	.15	1	0.27	1.8	$NaHCO_3$	50 mL	48 h (50 mL) then 24 h (15 mL)	RT for 44 h then 40 °**C** for 4 h	28.4%
7	.05	1	0.07	1.4	$NaHCO_3$ (1 mL) MeCN (0.5 mL)	1.5 mL	2 h	50 °**C**	64%
8	.036	1	0.043	1.2	$NaHCO_3$ (1 mL) MeCN (0.5 mL)	1.5 mL	1.5 h	50 °**C**	70%
9	.03	1	0.044	1.4	$NaHCO_3$ (1 mL) MeCN (0.5 mL)	1.5 mL	2 h	50 °**C**	72%

HPLC conditions were optimized (**Figure 2.4**) for the final purification of the FITC-Tre **40** from fluorescein isothiocyanate. An acetonitrile gradient with 0.1% TFA (0 to 50% over 15 minutes) worked well. Shorter gradients were also optimized to reduce acetonitrile consumption. Additionally, a mixture of up to 50% methanol/acetonitrile also worked to reduce acetonitrile consumption, although **40** tended to elute at a higher concentration of methanol than when purified with acetonitrile and would stick to the column and coelute with fluorescein isothiocyanate at 100% methanol in a pure aqueous methanol gradient system.

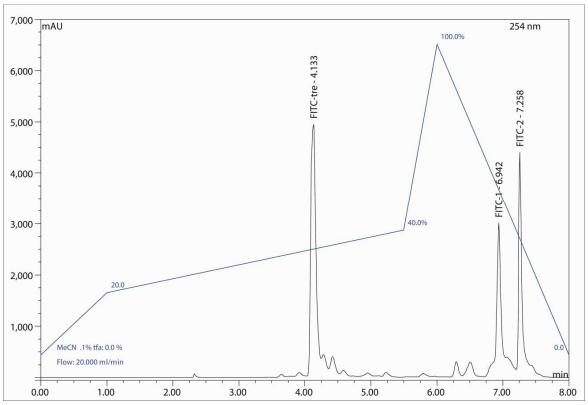
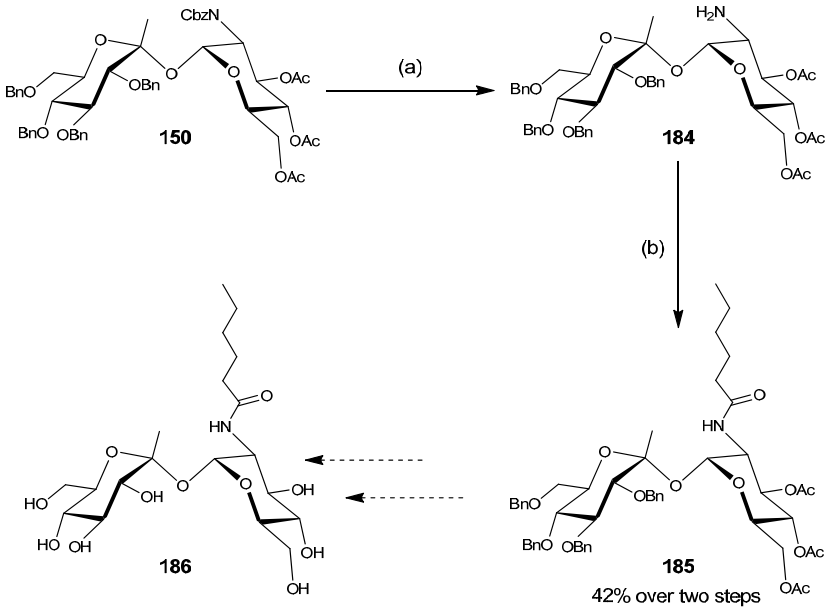


Figure 2.4. HPLC chromatograph of FITC-Tre 40 purification.

2.2.13.1 Selective removal of the Cbz protecting group

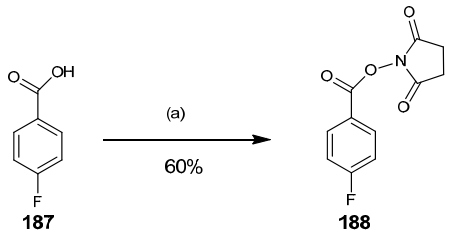
During the deprotection optimization process, it was discovered that the Cbz group could be selectively hydrogenated in the presence of all benzyl groups (**Scheme 2.50**). Capitalizing on this selectivity, the Cbz was removed and selectively acylated on the amine with hexanoyl chloride in 42% over two steps to afford **185**. This reaction series demonstrates the feasibility of selectively modifying the amine prior to deprotection, allowing the use of reagents that could react with any free hydroxyls.



Scheme 2.50.Selective removal of carboxybenzyl (Cbz) group. (a) H_2, $Pd(OH)_2$ /C, RT, 2 h (b) hexanoyl chloride, RT, DCM, 16 h.

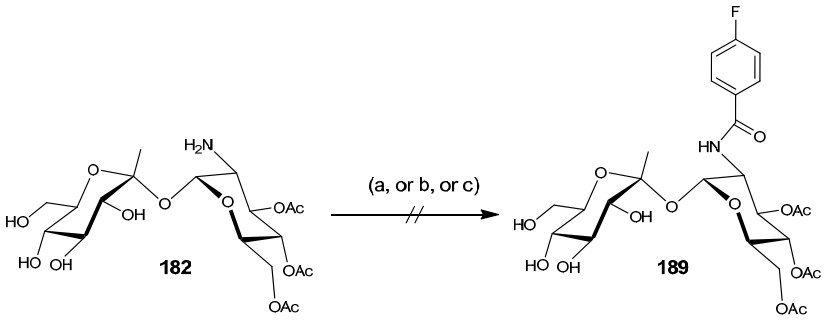
2.2.13.2 Peptide coupling

It was then thought that if fluorescein and hexanoyl chloride could be coupled to **37**, various other compounds might be coupled via peptide coupling techniques. One such compound is 4-F-benzoic acid **187**. ^{18}F-benzoic acid has already been accessed and could provide an easier route to ^{18}F based trehalose[45]. We attempted three different peptide coupling methods, NHS ester, PyBOP and Bis(tetramethylene)fluoroformamidinium Hexafluorophosphate (BTFFH). NHS ester **188** was synthesized easily from the 4-F-benzoic acid and N-hydroxysuccinimide (NHS) (**Scheme 2.51**).



Scheme 2.51. Synthesis of NHS ester 188. (a) NHS, DCC, THF, 16 h.

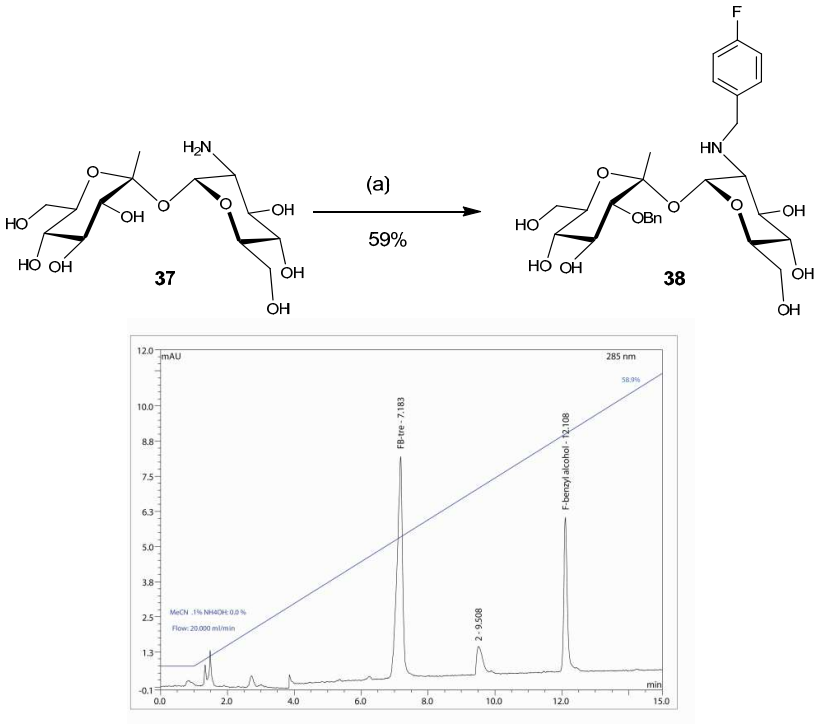
Product of coupling **189** was detected by ESI when BTFFH was utilized. However, none of the desired product was isolated through any of these methods. Coupling with a commercially available NHS ester of a biotin was subsequently found to require heating and long reaction times in anhydrous DMF. In aqueous buffer, **188** appeared to hydrolyze before any coupling could be detected. The somewhat hindered nature of the secondary amine may slow the rate of reactivity to the point that hydrolysis of the NHS ester **188** outcompeted the peptide coupling. These couplings were carried out while the trehalose analogue was still acetate protected (for increased solubility in organic solvents), so acetate migration to the amine of **37** may also have hindered the reaction.



Scheme 2.52. Attempted peptide couplings. (a) PyBop, 4-F-Benzoic acid, DIPEA, DMF, RT, 2 h (b) BTFFH, 4-F-Benzoic acid, DIPEA, DMF, RT, 2 h (c) NHS ester 188, DMF, DIPEA, 2 h – 16 h.

2.2.13.3 Reductive amination

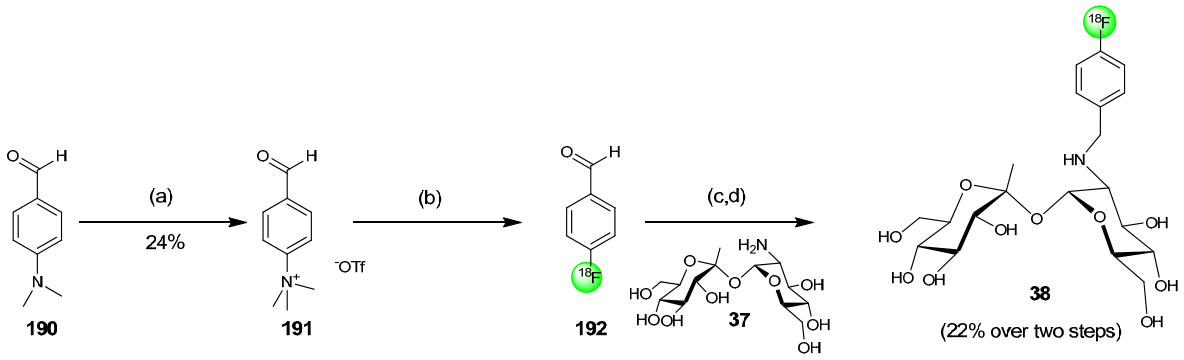
As an alternative to peptide coupling, reductive amination was chosen as a possible strategy for the synthesis of ^{18}F-trehalose. Fluoro-benzyl-trehalose analogue **38** was synthesized in two different manners. First, the imine between 4-F-benzaldehyde and **37** was formed and reduced via hydrogenation. Partitioning between water and DCM yielded the desired product as a salt with 4-F-benzoic acid. Second, reductive amination with sodium borohydride, followed by HPLC purification with a 0.1% NH_4OH acetonitrile gradient also yielded the desired product as the trifluoroacetic acid (TFA) salt (**Scheme 2.53**). TFA was identified as the salt by mass spectrometry and fluorine NMR. Possibly TFA contamination from prior HPLC purifications caused this salt formation.



Scheme 2.53. Synthesis of 4-F-benzyl-trehalose analogue 38, using reductive amination. Reagents and conditions (a) i) 4-F-benzaldehyde, MeOH, RT, 10 min, ii) sodium borohydride, MeOH, RT, 10 min. Inset shows HPLC trace. 38 elutes first at about 20% acetonitrile.

As a central premise of this project was to synthesize and test ^{18}F-trehalose PET probes, an ^{18}F-trehalose derivative **38** was synthesized through the use of a reductive amination strategy (**Scheme 2.54**). This approach is common in radiochemistry as a means of labeling peptides and proteins. The synthesis of the ^{18}F-benzaldehyde precursor has been well documented[46]. Because the radiochemistry was already developed, it was thought that the fluorobenzyl would provide the ideal handle for a preliminary synthesis of ^{18}F-trehalose derivatives.

With the guidance of Dr. Dale Kiesewetter (NIH, NIBIB) the desired compound was obtained in 22% radiochemical yield in just over one hour, including HPLC purification, starting from the trimethylaminobenzaldehyde salt **191**. This salt was synthesized from dimethylbenzaldehyde and methyl triflate. **191** crystallized from solution in modest 24% yield. A fluoride displacement yielded the ^{18}F-aldehyde **192**. Reductive amination with **37** afforded the desired product, which was synthesized in an hour, including purification steps, and was then injected into rabbits infected with *Mtb* for evaluation as a PET probe (see **Chapter 5** for further discussion).



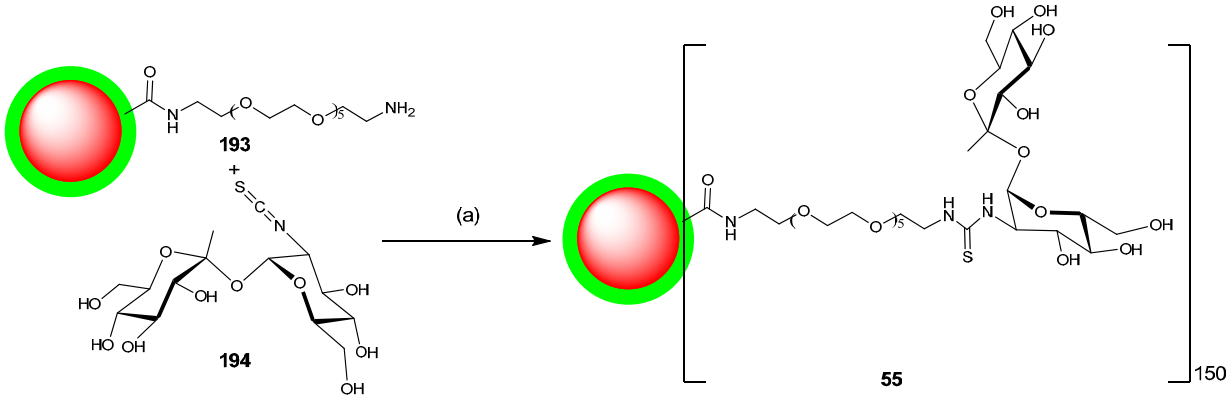
Scheme 2.54. Synthesis of ^{18}F-trehalose 39 (a) Methyl trifluoromethanesulfonate, DCM, RT, 16 h (b) $^{18}F^-$, acetonitrile, 130 °C, 5 min (c) Acetic acid, 37, methanol, 130 °C, 1 min (d) Sodium cyanoborohydride, 130 °C, 9 min.

Radiochemical purity was determined by HPLC with **39** having a retention time of 12.5 minutes, which was identical to the non-radioactive **38**. Both semi prep and analytical HPLC traces showed very little UV active contaminants with the same retention time as the radioactive peak, confirming high purity. Radioactive compound appeared as two peaks because it saturates the detector and then the detector recovers. Radiographic TLC was employed to positively identify the product and to determine how much of the desired compound was in the breakthrough of the C_{18} cartridge. None of **39** was found in the breakthrough, which had been a problem in earlier syntheses. In summary, we have developed what we believe is the first synthesis of an ^{18}F-trehalose analogue.

2.2.13.4 Trehalose coated quantum dots

As an alternative probe to FITC-Tre and ^{18}F-trehalose, quantum dots, covered in trehalose were synthesized (**Scheme 2.55**). Amino PEG coated quantum dots (QDs) **193** with an emission of 655 nm were obtained from Invitrogen. Isothiocyanate trehalose analogue **194** was synthesized from **37**, employing thiophosgene. Reaction with thiophosgene proceeded spot to spot (TLC) and with complete conversion by mass spectrometry. **194** was used without further purification to couple to the amino-peg QDs. Reaction was conducted in $NaHCO_3$ buffer pH 9, overnight at 4 °C with shaking. Upon completion, unreacted isothiocyanate trehalose compound **194** was washed away using a viva spin concentrator and QDs were assayed for the degree of coating, using a phenol, acid assay. A loading of approximately 150 sugars/dot were obtained in the modified QDs **55**. Even though QD

concentration can be determined by the UV absorbance[47]. It was difficult to calculate the yield of this reaction as the starting PEGylation of **193** was unknown.



Scheme 2.55. Synthesis of Zn-Cd-trehalose coated quantum dots 55. (a) $NaHCO_3$ buffer, 4°C, 14 h.

It has previously been demonstrated that QDs could be run on DNA gels and this can be used to demonstrate that the dots have been modified[48]. A standard 1% agarose gel in TAE buffer and added the dots **193** and **55** in loading buffer. A DNA ladder was included for comparison. QDs migrate slowly relative to DNA ladder. Glycosylated dots **55** migrated more quickly than unmodified dots **193**, consistent with the model that reaction with **194** would block many of the free amines on the dots surface, allowing for greater positive charge affinity (**Figure 2.5**). These results combined with the phenol acid assay confirmed that the dots had been modified.

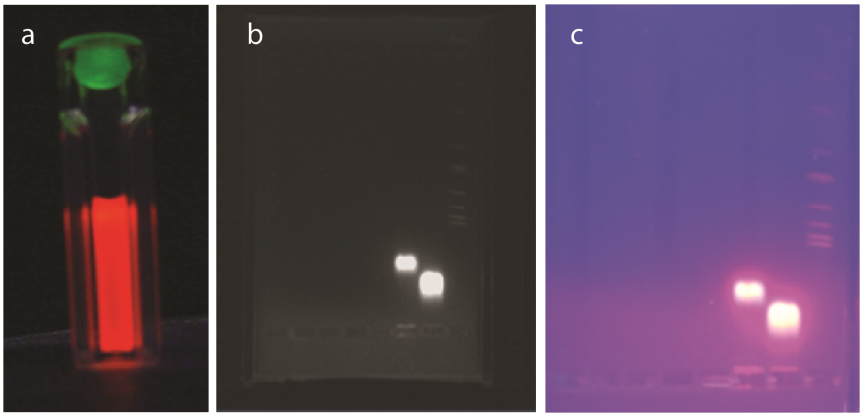
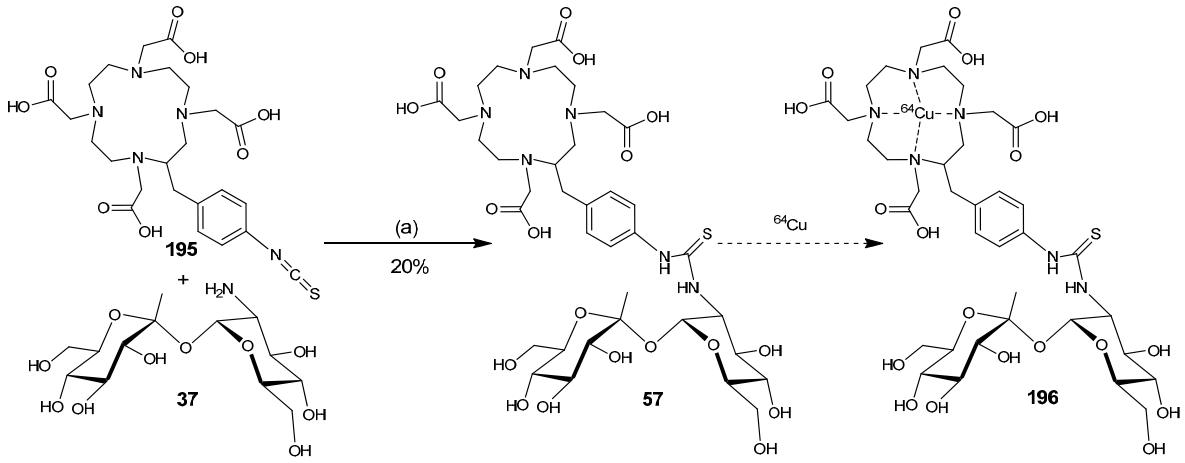


Figure 2.5. Quantum Dots 55 characterization. (a) Sample of the carbohydrate modified CdSe-ZnS quantum dot solution under UV irradiation. (b) DNA gel under UV light (254 nm) run for 60 min at 100 V (c) DNA gel under long wave 366 nm light. Left lane (modified dots) Right lane (unmodified dots).

2.2.13.5 Synthesis of DOTA-trehalose analogue 57

DOTA (1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid) is a very effective chelator for many metal ions, including copper 64. ^{64}Cu is a positron emitting radioisotope of copper, with a half life of 12.7 hours. It has found many uses in medical imaging and often antibodies are labeled with DOTA and chelated with ^{64}Cu for the purpose of PET imaging, particularly of tumors, utilizing tumor specific antibodies. Notably, ^{64} Cu has also been demonstrated to be a chemotherapeutic, prolonging the life of hamsters bearing human colon cancer cells[49]. Potentially, the radiation could have a therapeutic effect on the treatment of tuberculosis.



Scheme 2.56. Synthesis of DOTA-trehalose 196. (a) 75 mM Sodium bicarbonate buffer, pH 9, acetonitrile, 3 days, RT.

A preliminary synthesis of DOTA-trehalose analogue **57** was conducted **(Scheme 2.56**) with the hope that the compound would selectively label *Mtb* and allow for PET scanning with a longer residence time of the probe in the infected animal. Residence time was relevant because all of the preliminary enzymology and uptake assays (**Chapter 3 and Chapter 4**) suggested that uptake of trehalose into *Mtb* is time dependent and could take longer than the 100 minute half-life of ^{18}F-trehalose. The 12 hour half life of ^{64}Cu could compensate for slow uptake. Additionally, the ease of synthesis of copper chelates (mix with copper and purify by PD10 SepPak) was attractive because no extensive synthetic procedures would need to be developed.

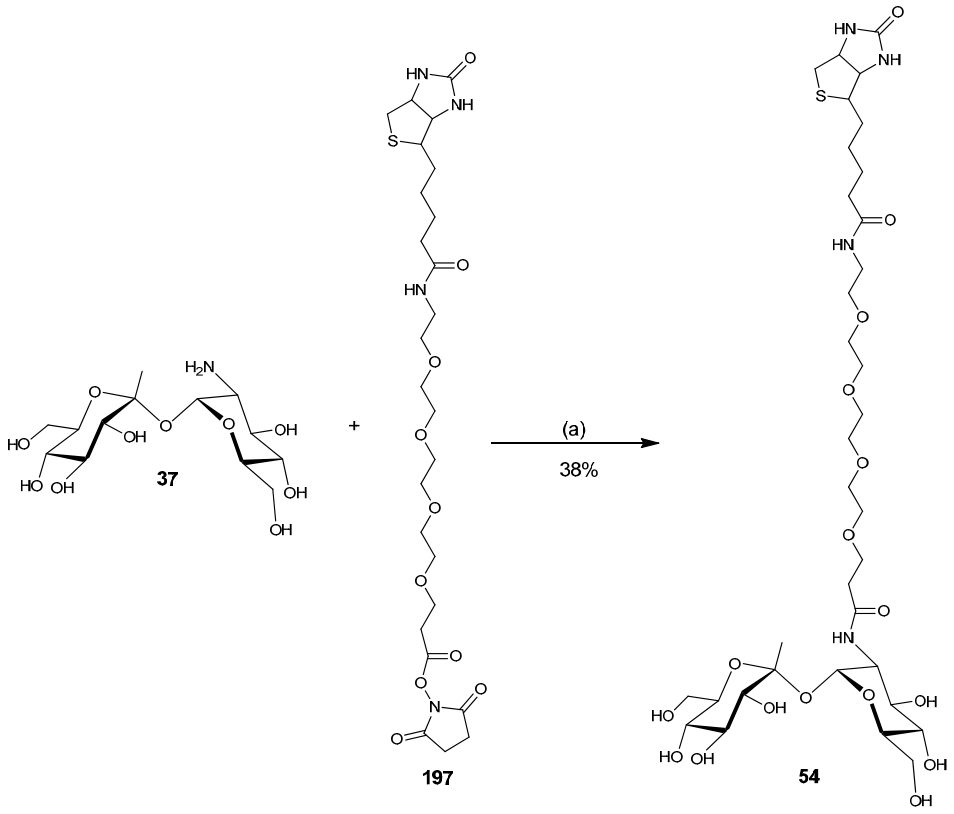
Utilizing a DOTA-isothiocyanate compound **195** obtained from Macrocyclics™, it was found that the previously synthesized **37** could be coupled to the DOTA over a period of three days

at RT in bicarbonate buffer with acetonitrile. Reaction monitoring by TLC and ESI mass spectrometry, suggested that the DOTA coupling reaction proceeded slowly, potentially due to solubility or steric bulk. All HPLC conditions we attempted on the coupled product **57** resulted in **57** eluting in the injection fraction, possibly due to the compound's zwitterionic nature. Resin-based anion exchange followed by cation exchange afforded **57** in albeit low (20% yield).

2.2.13.6 Synthesis of biotin-trehalose analogue 54

With an interest in the development of tuberculosis diagnostics and other methods for tracking *Mtb* in cell culture, we sought to synthesize a biotinylated trehalose **54**. This compound could be used to pull down and count bacteria with FACS, utilizing anti-biotin magnetic beads. Likewise, anti-biotin antibodies conjugated to fluorophors could be used to visualize biotin in cell cultures. We also considered the use of streptavidin coated nanoparticles to visualize the *Mtb* with EM microscopy.

A Biotin-PEG-NHS ester **197** was obtained from Fisher with a tetraethylene glycol linker. The spacer might be useful for increased affinity during pull-down assay and/or enhance Ag85 activity towards the compound. Coupling with **37** was conducted in anhydrous DMF to minimize the hydrolysis of the NHS ester (**Scheme 2.57)**. Coupling proceeded slowly over 4 days, but eventually reaction was deemed complete. Biotin absorbs weakly at 200 nm, so HPLC purification, using reverse phase C18 and UV detection was possibly. **54** was obtained as a white solid in 38% yield.

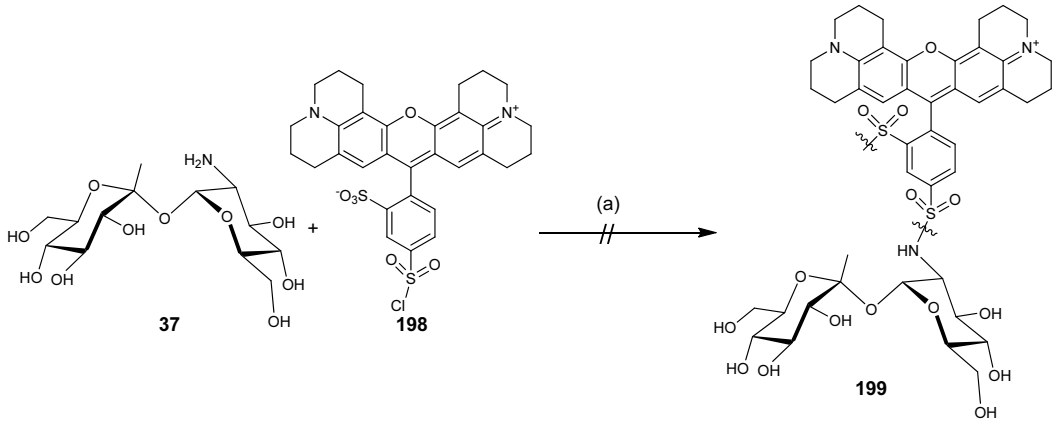


Scheme 2.57. Synthesis of biotin-trehalose 54. (a) Anhydrous DMF, Na_2CO_3 1 day 40 °C, 4 days , RT.

2.2.13.7 Synthesis of rhodamine-trehalose 56

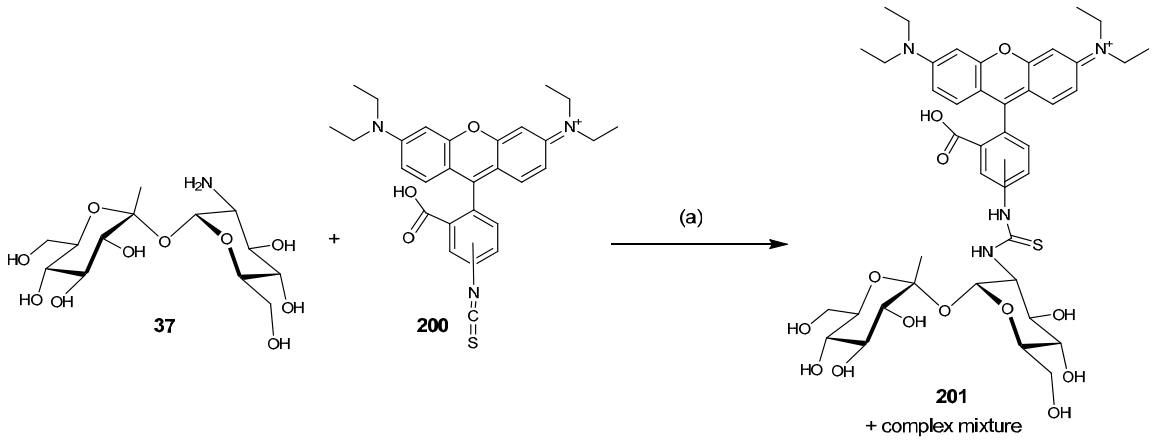
A red-fluorescent trehalose was hypothesized to be of use for labeling in their primary respiratory epithelial culture model, as the epithelial cells autofluoresce green. Additionally, red-trehalose might be useful in staining of tissue, as the tissue, particularly the lung tissue, has a green autofluorescence that makes distinguishing between label and background difficult. Because of these applications, carboxyrhodamine labeled trehalose analogue **56** was synthesized, utilizing **37** as a starting point.

Couplings of **37** were first attempted with sulfonyl chlorides of rhodamine and Texas red dyes. These reagents were a mixture of the two chlorides, due to rapid exchange between the sulfonates. The sulfonyl- chloride is quite susceptible to aqueous hydrolysis. Without base in anhydrous DMF, no reactivity between the sulfonyl chloride Texas red and the trehalose was observed, and instead, complete hydrolysis of the sulfonyl-chloride was detected (**Scheme 2.58**).



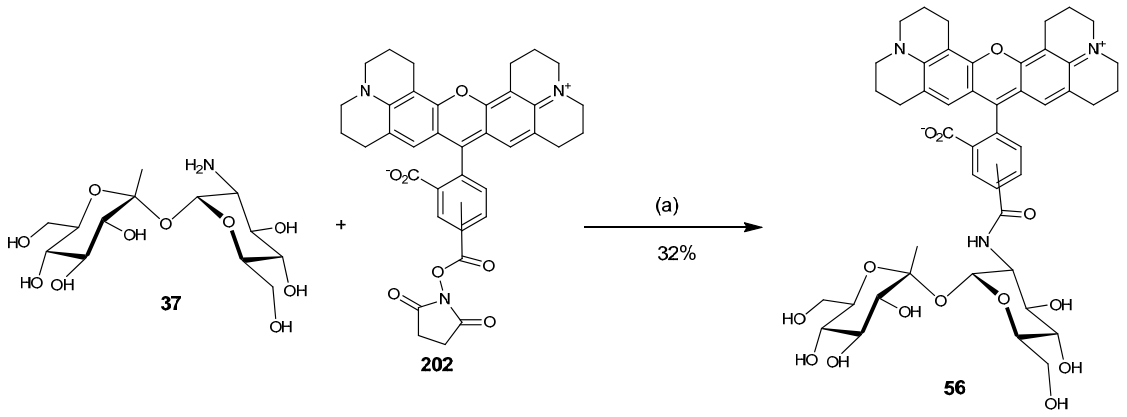
Scheme 2.58. Attempted coupling between 37 and carboxyrhodamine sulfonyl chloride 198 (a) anhydrous DMF, RT, 2 days.

Addition of base might lead to acylation of the free hydroxyls in addition to the free amine, so the synthesis was altered and couplings with the NHS ester of carboxyrhodamine **202** or the isothiocyanate of rhodamine-B **200** were attempted. Rhodamine derivatives are very expensive as the pure 5-derivative and instead are most commonly obtained as a mixture of the 5- (and-6)-carboxyrhodamine or sulfonyl-rhodamine or rhodamine B isothiocyanate (RBITC). We obtained both RBITC and the NHS carboxyrhodamine and attempted couplings with both. For the RBITC coupling reaction was conducted in the FITC optimized sodium bicarbonate pH 9 conditions (**Scheme 2.59**). Upon HPLC purification, the desired product was obtained as a complex mixture of the desired product as a mixture of the 5 and 6 isomers as well as other components that coeluted on the C18 column (See **Appendix 1** for trace). Alterations in the purification failed to afford pure rhodamine-trehalose compound **201**.



Scheme 2.59. Coupling between 37 and rhodamine isothiocyanate (RBITC) 200. (a) 75 mM sodium bicarbonate buffer, pH 9, acetonitrile, RT, 3 days.

The coupling of the NHS ester of carboxyrhodamine **202** to **37** (**Scheme 2.60**) was then attempted. This reaction afforded a much cleaner HPLC trace (See chemistry **Appendix 1**). However, upon lyophilization, a very small amount of product was isolated. The low yield may be due to the low reactivity of the sterically hindered amine and ester. This reaction was repeated with the additions of DIPEA and longer reaction time and with mild heating. Yield of **56** remained low, but was sufficient for preliminary biological testing



Scheme 2.60. Synthesis of carboxyrhodamine-trehalose 56 (a) Anhydrous DMF, Na_2CO_3, 50 °C, 24 h.

2.3 Conclusions

A survey of many trehalose modification strategies, including radical bromination, enzymatic synthesis, intramolecular aglycone delivery, direct glycoslyations and the use of anomeric directing groups, led us to the development of the methyl ketoside glycosylation methodology as a robust means to rabidly synthesize a library of trehalose analogues. The glycosylation and deprotection reactions have been optimized for yield and stereoselectivity and scaled to allow for gram scale synthesis of target compounds. These compounds include modifications at nearly every position of the sugar scaffold as well as charged, bulky and radioactive groups. Ketoside glycosylations facilitated the late stage installation of isotopic labels, particularly ^{18}F tags for PET imaging and the development of a panel of trehalose probes (fluorescein, biotin, rhodamine, quantum dots) that in future chapters will be shown to have the capacity to label *Mtb.* In summary, we have evaluated the scope of the ketoside glycosylation methodology and used this chemistry to stereoselectively synthesize the challenging α,α-1,1-trehalose linkage, affording a library of trehalose compounds that have biological relevance to *Mtb.*

2.4 References

1. Ronning, D.R. et al. Crystal structure of the secreted form of antigen 85C reveals potential targets for mycobacterial drugs and vaccines. *Nat. Struct. Biol.* **7**, 141-6 (2000).
2. Ronning, D.R., Vissa, V., Besra, G.S., Belisle, J.T. & Sacchettini, J.C. *Mycobacterium tuberculosis* antigen 85A and 85C structures confirm binding orientation and conserved substrate specificity. *J. Biol. Chem.* **279**, 36771-7 (2004).
3. Anderson, D.H., Harth, G., Horwitz, M.A. & Eisenberg, D. An interfacial mechanism and a class of inhibitors inferred from two crystal structures of the *Mycobacterium tuberculosis* 30 kDa major secretory protein (Antigen 85B), a mycolyl transferase. *J. Mol. Biol.* **307**, 671-81 (2001).
4. DeLano, W.L. *The PyMOL Molecular Graphics System,* (Delano Scientific, Palo Alto, CA, USA, 2002).
5. Backus, K.M. et al. Uptake of unnatural trehalose analogs as a reporter for *Mycobacterium tuberculosis*. *Nat. Chem. Biol.* **7**, 228-35 (2011).
6. Klutts, S. et al. Purification, cloning, expression, and properties of mycobacterial trehalose-phosphate phosphatase. *J. Biol. Chem.* **278**, 2093-100 (2003).
7. Ferrier, R.J. & Furneaux, R.H. Photobromination of carbohydrate derivatives. VI. Functionalization at C6 of 2,3,4-Tri-O-acetyl-1,6-anhydro-β-D-glucopyranose. *Aust. J. Chem.* **33**, 1025-1036 (1980).
8. Skelton, B.W., Robert V. Stick, R.V., Stubbs, K.A., Watts, A.G. & H., W.A. The Fluorination (at C5) of Some Derivatives of D-Glucose. *Aust. J. Chem.* **4**, 345-353 (2004).
9. Hartman, M.C.T. & Coward, J.K. Synthesis of 5-Fluoro-N-Acetylglucosamine Glycosides and Pyrophosphates via Epoxide Fluoridolysis: Versatile Reagents for the Study of Glycoconjugate Biochemistry. *J. Am. Chem. Soc.* **124**, 10036-10053 (2002).
10. Doores, K. The University of Oxford (2007).
11. Blattner, R. & Ferrier, R.J. Photo-bromination of carbohydrate derivatives. Part 2. Penta-O-acetyl-β-D-glucopyranose; the 5-bromo-derivative and products of further bromination. *Journal of the Chemical Society, Perkin Transactions 1*, 1523-1527 (1980).
12. Nunes, S.C.C., Jesus, A.J.L., Moreno, M.J. & Eusébio, M.E.S. Conformational preferences of α,α-trehalose in gas phase and aqueous solution. *Carbohydr. Res.* **345**, 2048-2059.
13. Namme, R., Hideyo, T.M. & Ikegami, T.S. Development of Ketoside-Type Analogues of Trehalose by Using α-Stereoselective O-Glycosidation of Ketose. *Eur. J. Org. Chem.* **2007**, 3758-3764 (2007).
14. Gomez, A.M., Slawomir, C.U., Valverde, J.S. & Lopez, J.C. Stereoselective Synthesis of C- and N-Ketosides by Lewis Acid-Catalyzed C- and N-Glycosidation of Alkynyl, Phenyl, and Methyl Ketoses. *Eur. J. Org. Chem.* **2003**, 4830-4837 (2003).
15. Griffin, Frank K., Paterson, Duncan E., Murphy, Paul V. & Taylor, Richard J.K. A New Route to exo-Glycals Using the Ramberg-Bäcklund Rearrangement. Vol. 2002 1305-1322 (2002).
16. Gomez, A.M., Pedregosa, A., Valverde, S. & Lopez, J.C. Stereoselective synthesis of substituted exo-glycals from 1-exo-methylene pyranoses. *Tetrahedron Lett.* **44**, 6111-6116 (2003).
17. Yamanoi, T., Inoue, R., Matsuda, S., Katsuraya, K. & Hamasaki, K. Synthesis of trehalose mimics by bismuth(III) triflate or bis(trifluoromethane)sulfonimide-catalyzed 1-C-methyl-d-hexopyranosylation. *Tetrahedron: Asymmetry* **17**, 2914-2918 (2006).
18. Li, X., Ohtake, H., Takahashi, H. & Ikegami, S. A facile synthesis of 1'-C-alkyl-α-disaccharides from 1-C-alkyl-hexopyranoses and methyl 1-C-methyl-hexopyranosides. *Tetrahedron* **57**, 4297-4309 (2001).
19. Li, X., Ohtake, H., Takahashi, H. & Ikegami, S. An efficient synthesis of new 1'-C-methyl-α-O-disaccharides using 1-methylenesugars as the glycosyl donors. *Tetrahedron* **57**, 4283-4295 (2001).
20. Yamanoi, T. et al. Synthesis of 1-C-alkyl-α-d-glucopyranosides by Lewis acid- or Bronsted acid-catalyzed O-glycosidation. *Tetrahedron* **62**, 10383-10392 (2006).
21. Tebbe, F.N., Parshall, G.W. & Reddy, G.S. Olefin homologation with titanium methylene compounds. *J. Am. Chem. Soc.* **100**, 3611-3613 (1978).
22. Petasis, N.A. & Bzowej, E.I. Titanium-mediated carbonyl olefinations. 1. Methylenations of carbonyl compounds with dimethyltitanocene. *J. Am. Chem. Soc.* **112**, 6392-6394 (1990).
23. Jansson, K. et al. 2-(Trimethylsilyl)ethyl glycosides. 3. Synthesis, anomeric deblocking, and transformation into 1,2-trans 1-O-acyl sugars. *J. Org. Chem.* **53**, 5629-5647 (1988).
24. Yang, W.B. et al. Stereochemistry in the synthesis and reaction of exo-glycals. *J. Org. Chem.* **67**, 3773-82 (2002).
25. Pratt, M.R., Leigh, C.D. & Bertozzi, C.R. Formation of 1,1-α,α-glycosidic bonds by intramolecular aglycone delivery. A convergent synthesis of trehalose. *Org. Lett.* **5**, 3185-8 (2003).
26. Tiwari, P. & Misra, A.K. An Efficient Stereoselective Dihydroxylation of Glycals using a Bimetallic System, RuCl3/CeCl3/NaIO4. Vol. 71 2911-2913 (2006).

27. Hadd, M.J. & Gervay, J. Glycosyl iodides are highly efficient donors under neutral conditions. *Carbohydr. Res.* **320**, 61-69 (1999).
28. Danishefsky, S.J., Hu, S., Cirillo, P.F., Eckhardt, M. & Seeberger, P.H. A Highly Convergent Total Synthetic Route to Glycopeptides Carrying a High-Mannose Core Pentasaccharide Domain *N*-linked to a Natural Peptide Motif. *Chemistry - A European Journal* **3**, 1617-1628 (1997).
29. Khan, R., Konowicz, P.A., Gardossi, L., Matulova, M. & Degennaro, S. Regioselective Deacetylation of Fully Acetylated Mono- and Di-Saccharides With Hydrazine Hydrate. *Aust. J. Chem.* **49**, 293-298 (1996).
30. Nyffeler, P., Gonzalez, S., Burkart, M., Vincent, S. & Wong, C. Selectfluor: Mechanistic Insight and Applications. *Angewandte Chemie International Edition* **44**, 192-212 (2005).
31. Vincent, S.P., Burkart, M.D., Tsai, C.Y., Zhang, Z. & Wong, C.H. Electrophilic Fluorination-Nucleophilic Addition Reaction Mediated by Selectfluor: Mechanistic Studies and New Applications. *J. Org. Chem.* **64**, 5264-5279 (1999).
32. Burkart, M.D., Zhang, Z., Hung, S.C. & Wong, C.H. A New Method for the Synthesis of Fluoro-Carbohydrates and Glycosides Using Selectfluor. *J. Am. Chem. Soc.* **119**, 11743-11746 (1997).
33. Diksic, M. & Jolly, D. Synthesis of 2-deoxy-2-fluorohexoses by fluorination of glycals in aqueous media. *Carbohydr. Res.* **153**, 17-24 (1986).
34. Horlacher, T. & Seeberger, P.H. The utility of carbohydrate microarrays in glycomics. *Omics-a Journal of Integrative Biology* **10**, 490-498 (2006).
35. Feizi, T., Fazio, F., Chai, W. & Wong, C.-H. Carbohydrate microarrays - a new set of technologies at the frontiers of glycomics. *Curr. Opin. Struct. Biol.* **13**, 637-645 (2003).
36. Likhosherstov, L.M., Novikova, O.S., Derevitskaja, V.A. & Kochetkov, N.K. A new simple synthesis of amino sugar β-D-glycosylamines. *Carbohydr. Res.* **146**, C1-C5 (1986).
37. Braun, M.P., Dean, D.C. & Melillo, D.G. Suzuki cross-coupling for the incorporation of labeled methyl groups onto aryl halides. A synthesis of ^{14}C-tosyl chloride and its use in the synthesis of ^{14}C-L-738,167. Vol. 42 469-476 (1999).
38. Reynolds, K., O'Hagan, D., Gani, D. & Robinson, J. Butyrate Metabolism in Streptomycetes. Characterization of an Intramolecular Vicinal Interchange Rearrangement Linking isobutyrate and Butyrate in *Streptomyces cinnamonensis*. *J. Chem. Soc. Perkin Trans. I* **1**, 10 (1988).
39. Rosenblum, S.B. & Bihovsky, R. Synthesis of the papulacandin C-arylglucosyl spiroketal nucleus. *J. Am. Chem. Soc.* **112**, 2746-2748 (1990).
40. Tang, C.-J., Shen, X. & Wu, Y. Synthesis of An Advanced Precursor to the Amino Epimer of (+)-Pyloricidins. *Monatshefte für Chemie / Chemical Monthly* **134**, 1617-1622 (2003).
41. Schneider, E. et al. Magnetic resonance spectroscopy for measuring the biodistribution and *in situ in vivo* pharmacokinetics of fluorinated compounds: validation using an investigation of liver and heart disposition of tecastemizole. *J. Clin. Pharm. Ther.* **31**, 261-273 (2006).
42. Weiss, R.G., Chacko, V.P. & Gerstenblith, G. Fatty acid regulation of glucose metabolism in the intact beating rat heart assessed by carbon-13 NMR spectroscopy: the critical role of pyruvate dehydrogenase. *Journal of Molecular and Cellular Cardiology* **21**, 469-478 (1989).
43. Myszka, H., Bednarczyk, D., Najder, M. & Kaca, W. Synthesis and induction of apoptosis in B cell chronic leukemia by diosgenyl 2-amino-2-deoxy-β-glucopyranoside hydrochloride and its derivatives. *Carbohydr. Res.* **338**, 133-141 (2003).
44. Bailey, P.D. et al. Debenzylation using catalytic hydrogenolysis in trifluoroethanol, and the total synthesis of (-)-raumacline. *Tetrahedron Lett.* **49**, 2150-2153 (2008).
45. Vaidyanathan, G. & Zalutsky, M.R. Labeling proteins with fluorine-18 using N-succinimidyl 4-^{18}F-fluorobenzoate. *Int. J. Rad. Appl. Instrum. B* **19**, 275-81 (1992).
46. Vaidyanathan, G. & Zalutsky, M.R. Synthesis of N-succinimidyl 4-^{18}F-fluorobenzoate, an agent for labeling proteins and peptides with ^{18}F. *Nat. Protocols* **1**, 1655-1661 (2006).
47. Leatherdale, C.A., Woo, W.K., Mikulec, F.V. & Bawendi, M.G. On the Absorption Cross Section of CdSe Nanocrystal Quantum Dots. *The Journal of Physical Chemistry B* **106**, 7619-7622 (2002).
48. Howarth, M. et al. Monovalent, reduced-size quantum dots for imaging receptors on living cells. *Nat Meth* **5**, 397-399 (2008).
49. Lewis, J. et al. Copper-64-diacetyl-bis(N4-methylthiosemicarbazone): An agent for radiotherapy. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 1206-11 (2001).

3 Chapter 3: Ag85 enzyme assays

3.1 Introduction

Chapter 3 reports the implementation of an Ag85 assay that allows full kinetic analysis as well as high-throughput screening for both substrates and inhibitors of the Ag85 enzymes. The kinetic parameters and catalytic mechanism of Ag85C were determined, and the trehalose library as well as a panel of monosaccharides were screened to determine the structure-activity relationships of the varied configurations of the sugar rings. These data reveal previously unknown substrate preferences and guided the design of a selective mechanism-based inhibitor of Ag85C.

The Ag85 enzymes have been proposed as potential targets for novel anti-mycobacterial drugs[1]. Although crystal structures of Ag85A, B and C are available and K_M values have been obtained, no detailed kinetic analysis or structure-activity relationship (SAR) studies have been reported and inhibitors of only millimolar efficiency have been identified. The development of inhibitors would benefit from a rapid and accurate assay of Ag85 activity and a thorough understanding of the enzyme mechanism.

The substrate specificity and a full kinetic analysis of the Ag85 enzymes has not previously been fully attempted, presumably because of the difficult nature of the radiochemical assay using natural substrates. Heterogeneity in the mycolic fatty acid chain length (C30-C90), as well as low solubility of both substrates and products in aqueous media, has complicated prior assays. The widely employed ^{14}C-Tre radioassay[2,3] monitors ^{14}C-Tre radiolabel incorporation into TDM and TMM yielding, at best, approximate $k_{cat(apparent)}$ values; such assays have suggested that Ag85B has a lower (~20%) activity than Ag85A and Ag85C[2]. Current assays also include coupled enzymatic assays for Ag85C[4,5]. Such methods preclude the possibility of automated screening and do not allow complete dissection of the kinetic steps of Ag85 acyltransfer, respectively.

3.2 Results

3.2.1 Antigen 85 Expression.

Purification of Ag85A,B and C was required for the development of any activity assay. Plasmids were obtained from Colorado State TBVTRM contract as Rv3804/Ag85A Rv1886c/Ag85B Rv0129c/Ag85C in pET23b vector containing a hexahistidine tag (His-tag) and were expressed as has been described previously[4]. Expression proceeded smoothly (See **Appendix 2**, for full details). Pure protein was obtained for all three Ag85 isoforms and mass spectrometry (MS) and sequencing analysis confirmed the correct identity of the proteins. Previously, the Ag85B and A proteins were reported to express in lower yields than Ag85C[6]. We found that protein expression levels were comparable and the lower yield was most likely due to a lower affinity for nickel resin of the Ag85A and B hexa his constructs. A decrease in the concentration of imidazole in the wash buffer increased the yield of the Ag85A and B proteins, allowing purification of all proteins in good yield.

During MS analysis of purified Ag85A (**Figure 3.1**) it was found that both a protein of the expected mass as well as a protein of mass + 178 Da were obtained in the purified protein. Analysis indicates that this mass is caused by His-tag phosphonogluconoylation[7] of a GSS–His sequence present in the N-terminal protein His-tag. In separate expressions a mixture of the gluconoylated and non-gluconoylated proteins and purely gluconoylated protein were obtained. No significant difference in activity for the gluconoylated and non-gluconoylated protein was observed. Furthermore, when the His-tag mixture of Ag85A was cleaved with thrombin, a single protein of the correct mass was obtained, confirming that the modification was in the His-tag and not in a more critical portion of the protein.

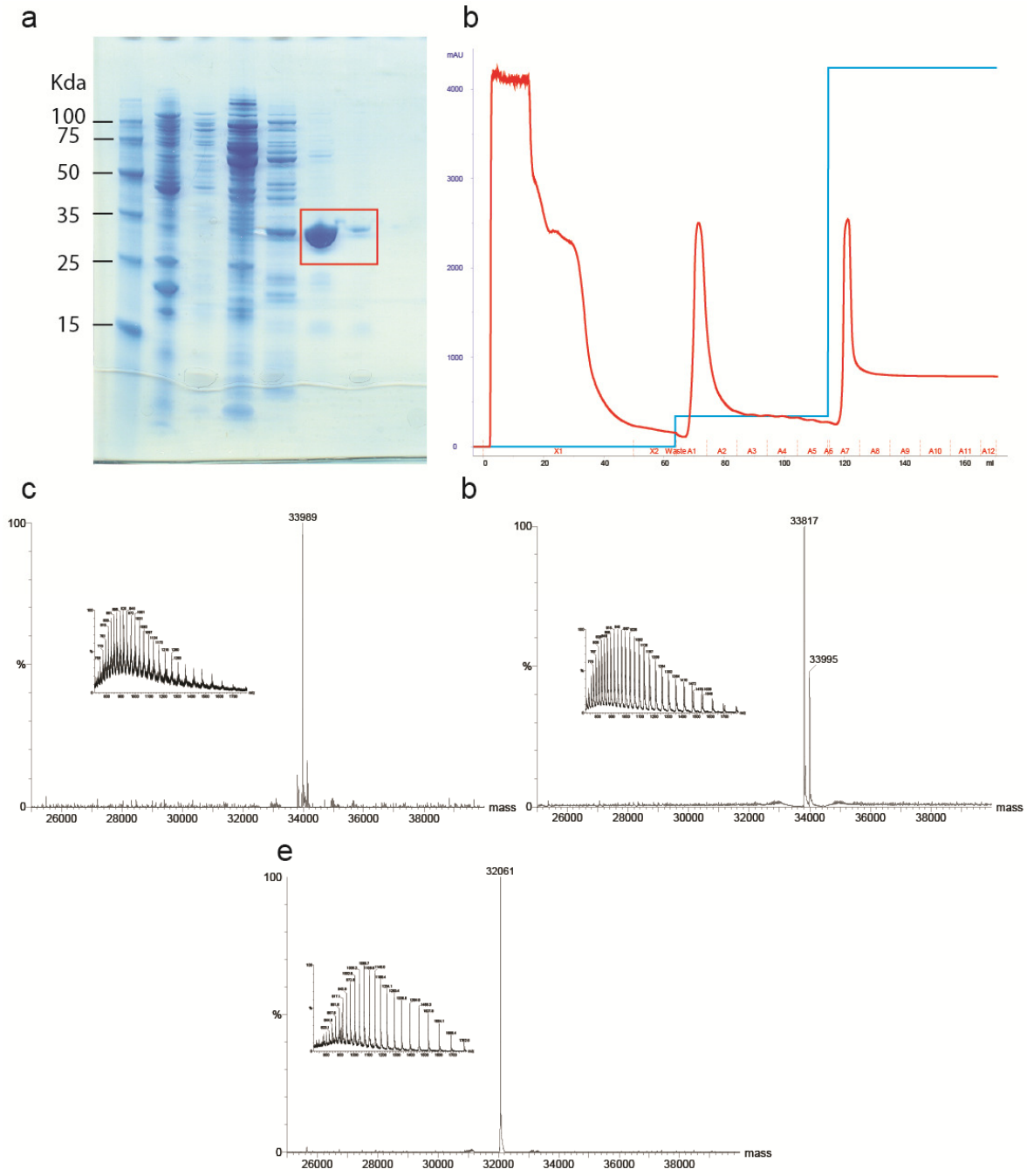


Figure 3.1. Purification and analysis of Ag85A. (a) SDS page gel of Ag85A, lane 1- novagen perfect protein makers 15-150 Kda, box indicates fractions collected (b) FPLC trace of protein expression red- protein UV absorbance, blue- gradient of imidazole. (5 mM to 1M) (c) Mass spectrum of Ag85A with gluconoylation. Insets show protein spectrum before deconvolution. Calculated mass Ag85A-glc:33993, Found: 33989; (d) Mass spectrum of Ag85A as a mixture with and without gluconoylation . Calculated mass Ag85A: 33814 Found: 33817; Calculated mass Ag85A-glc:33993, Found: 33995; (e) Mass spectrum of Ag85A after thrombin cleavage. Calculated mass Ag85A: 32058, Found: 32061.

3.2.2 Radiometric assays

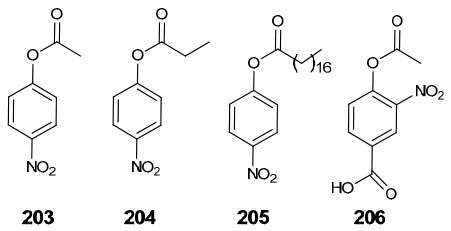
To design a novel, high throughput enzyme assay that allowed for the dissection of the complete Ag85 kinetics, the strengths and weakness of prior reported enzyme assays were first considered. At the inception of this project, only a radiometric assay TLC based assay had been reported in the literature. The radiometric assay employs ^{14}C-Trehalose and purified Trehalose monomycolate (TMM), which in the presence of Ag85, with sonication to solubilize to glycolipids, will generate ^{14}C-Trehalose dimycolate (^{14}C-TDM) and ^{14}C-Trehalose monomycolate (^{14}C-TMM). The product mixture is then extracted from the aqueous reaction mixture and separated by TLC, imaged by exposure to a phosphor-storage screen and then quantified by phosphorimager analysis. The same process can also be implemented by HPLC, equipped with a radioactivity detector. The advantage of this system is that it employs the native Ag85 substrates. However, the use of radioactivity makes the assay relatively cumbersome, expensive and difficult to implement in all laboratories. The native substrates are very poorly soluble, requiring extensive sonication to generate a homogenous reaction mixture. TMM is difficult to purify completely from mycobacterial lipid extracts and is often contaminated with other glycolipids, further complicating experimental results. This radioassay has been implemented successfully to assay several putative Ag85 inhibitors, including 6-azido-6-deoxy-D-trehalose[2] and trehalose phosphonates and sulfonates[8,9]. Preliminary results (see Section **3.4**) suggested that the radiographic TLC was possible to implement, but could not be transformed into a high throughput screening format.

3.2.3 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) has previously been employed to calculate K_M values for esterases[10] and was considered as an alternative method for Ag85 screening. ITC allows for the accurate measurement of binding affinities of small molecules with proteins, without the addition of any chromophors or use of coupled assays[11]. ITC was attempted with the Ag85 enzymes in order to calculate K_M and K_d. All ITC efforts were implemented with trehalose and Ag85C and no apparent binding was observed under all conditions employed. ITC only allows for one reaction at a time and is limited by its requirement of large amounts of both enzyme and substrate. These factors, in combination with the preliminary lack of measurable binding, led us to consider alternate assays.

3.2.4 Coupled assays

Because TLC and ITC assays seemed untenable for the screening of a library of compounds as substrates, we surveyed the previously reported serine esterase assay conditions. Nitrophenyl ester substrates (nitrophenyl acetate **203** and nitrophenyl benzoate) have previously been implemented in a high throughput fashion as substrates of serine esterases and the catalytic parameters of these enzymes can be measured indirectly from the release of colored nitrophenol[12].



Scheme 3.1. Nitrophenyl esters 203-206 synthesized from the corresponding alcohols as potential Ag85 substrates.

Several nitrophenyl esters compounds **203-206** (**Scheme 3.1**) were synthesized to investigate the hydrolysis of these compounds when treated with Ag85 in a 96 well plate, monitored by the plate reader at 360 and 410 nm. Solubility proved problematic and the formation of precipitate confounded the results. *p*-nitrophenyl acetate (PNPA) **203** was soluble at a 10mM concentration in a 1% DMSO phosphate buffer pH 6.5 mixture. However, *p*-nitrophenyl propionate (PNPP) **204** require the addition of 10% DMSO, and DMSO is thought to inhibit Ag85C[13]. Elevation of pH to above 7.5 led to extensive buffer-catalyzed hydrolysis of the substrates. *p*-Nitrophenyl stearate **205** (**Scheme 3.1**) was synthesized in good yield, but was completely insoluble in aqueous buffer. It was thought that the addition of a carboxylic acid in **206** might enhance solubility. Unfortunately, to solubilize **206** required an increase of the pH to above 7.5. These basic conditions led to increased baseline hydrolysis of the substrate in buffer. Due to insolubility and low reactivity towards the Ag85 enzymes, none of these esters seemed amenable to the development of a high throughput Ag85 enzyme assay.

While this work was in progress, two additional reports of Ag85 assays[4,5] were published. Boucau et al. implemented a coupled assay, similar to those previously reported for esterases. In this assay, Ag85 catalyzes the removal of an octanoyl ester from an 6-octanoyl *p*-nitrophenyl-β-d-glucoside, freeing the *p*-nitrophenyl-β-d-glucoside. In the absence of the 6-

hydroxyl modification, this glucoside is readily hydrolyzed by β-glucosidase to release the *p*-nitrophenolate chromophore. The chromogenic nature of this reaction allows for multi-well screening of compounds as inhibitors, but is limited by the fact that it implements glucose, rather than the native trehalose, and is a coupled assay that does not directly read out the catalytic information. This group reported that the basal rate of hydrolysis of the acyl glucose in the absence of enzyme is negligible and, therefore, they did not account for background hydrolysis in their screening. In contrast, we have found that comparable sugar esters are somewhat labile in buffer, suggesting that the baseline hydrolysis could confound their results. To date the best inhibitor reported by this group has an IC_{50} of 25 mM and has been suggested to form a covalent adduct, although only X-ray density and no mass spectrometry was used to confirm the presence of an acyl enzyme species.

Also while our work was ongoing, an additional coupled assay was published[5], reporting that the native trehalose monomycolate could be employed in a coupled assay that measured the degradation of trehalose (released by Ag85A) into glucose by a trehalase enzyme included in the reaction mixture. This assay has the advantage of employing the native substrate, but, as with Boucau et al., does not differentiate between hydrolysis and mycolyl transfer. Also both of these assays report k_{cat} and K_M for only the first step of the transesterification reaction, missing critical reaction information from later steps in the enzyme cycle.

3.2.5 Design of the Ag85 enzyme assay

Mass spectrometry (MS) was selected as our Ag85 assay method of choice. Since unnatural substrate features, such as chromo/fluorophores, are not necessary, MS can be readily automated, and multiple natural substrates and products can be monitored simultaneously to give an unimpeded view of kinetic processes[14-17]. Due to their heterogeneity and large size, the native TDM and TMM have only been found to be compatible with matrix-assisted laser desorption/ionization (MALDI) mass spectrometry[18]. MALDI did not suit the development of a high throughput, quantitative MS assay, so we turned our focus to the development of water soluble, TDM and TMM-like compounds that could be implemented in an automated liquid chromatography- mass spectrometry (LCMS) assay.

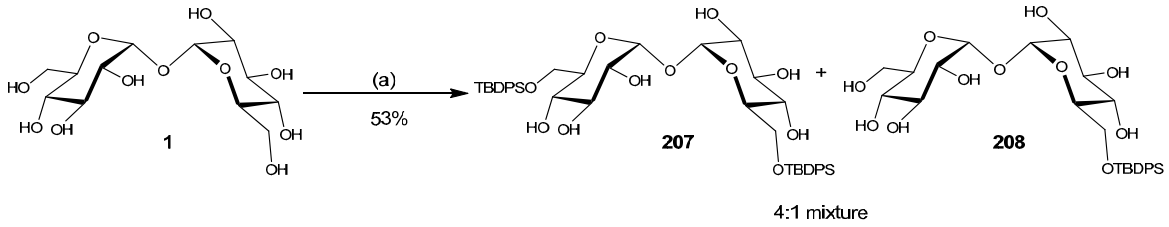
Based on prior reports that the Ag85 enzymes processed acyltrehalose appended with shorter lipids relatively well[3], it was hypothesized that trehalose modified with either one or two short acyl chains could serve as the basis for a new Ag85 enzyme assay, in which acyl

trehalose was mixed with compounds from the trehalose library (**Chapter 2**) and the transesterification reaction was monitored by LCMS. To validate this hypothesis, both trehalose mono and dihexanoate (TMH **212** and TDH **215**) were synthesized, starting from trehalose.

3.2.5.1 Synthesis of trehalose mono and dihexanoate (TMH, 212, TDH, 215)

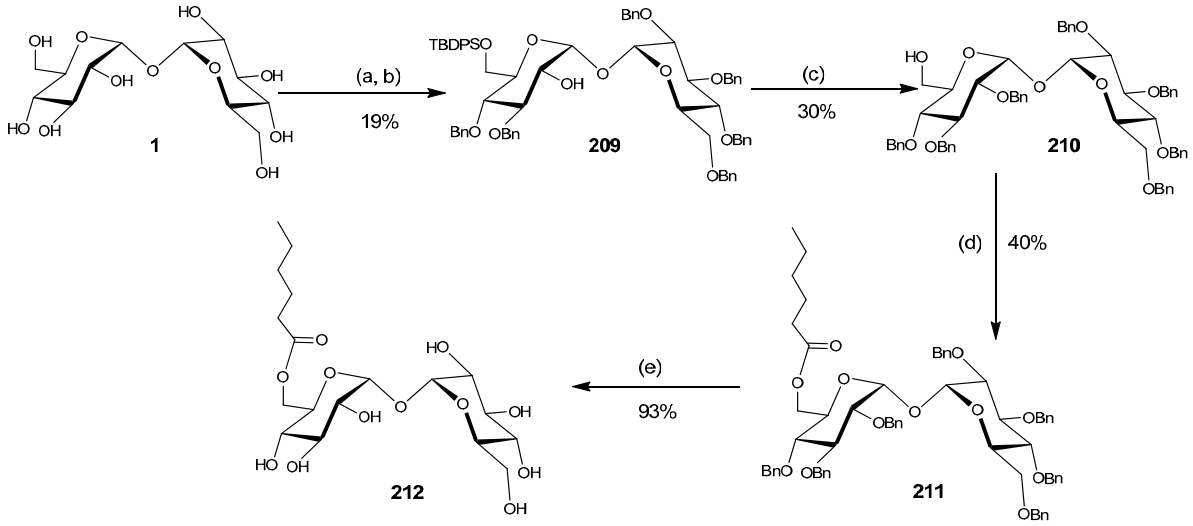
Orthogonal protection of the trehalose scaffold at the 6-positions was required for both the installation of the hexanoyl lipid and the selective deprotection of the rest of the sugar scaffold. Syntheses of both hepta[19] and hexa[19] benzyl trehalose analogues, **4** and **210,** have both been reported previously. Because the benzyl group can be removed under mild hydrogenation conditions without affecting the 6 position ester, we thought that these benzyl protected trehalose scaffolds could serve as useful intermediates in our synthesis.

Beginning with deprotected trehalose, the sugar can be selectively modified on the 6- and 6'-hydroxyls with silyl protecting groups to afford **207** and **208** in moderate yield. In prior 6-hydroxyl modifications of trehalose[20], *t*-butyldimethylsilyl(TBDMS) protection was employed. For our purposes, we selected the more hindered, less labile *t*-butyldiphenylsilyl (TBDPS) protection. This group offers the advantage of truly selective modification of only the 6-hydroxyls, whereas less hindered silyl groups (TBDMS) may also afford some 4-hydroxyl modification. The silylation reaction proceeded smoothly in the presence of imidazole in 53% yield (**Scheme 3.2**). The resulting mono and diTBDPS products **207** and **208** are extremely insoluble in most organic solvents and, although they have very different R_f values, separate poorly on column chromatography. Multiple rounds of chromatography afforded some pure disilyl **208**, allowing for full characterization. As has been reported previously, the stochiometry of the silylation reaction did not result in the expected statistical distribution of 1:2:1 unprotected:mono-silyl:di-silyl and, instead, strongly favored the formation of the di-TBDPS product, potentially due to the solubilizing affect of the silylation[21]. With a slight excess of TBDPSCl a 4:1 ratio of di-silyl:mono-silyl products was observed. However, we found that addition of less than one equivalent of TBDPSCl can afford mainly the monoTBDPS modified trehalose analogue **207**, with the diTBDPS compound **208** not purified from the reaction mixture in a low (19%) yield (**Scheme 3.3**).



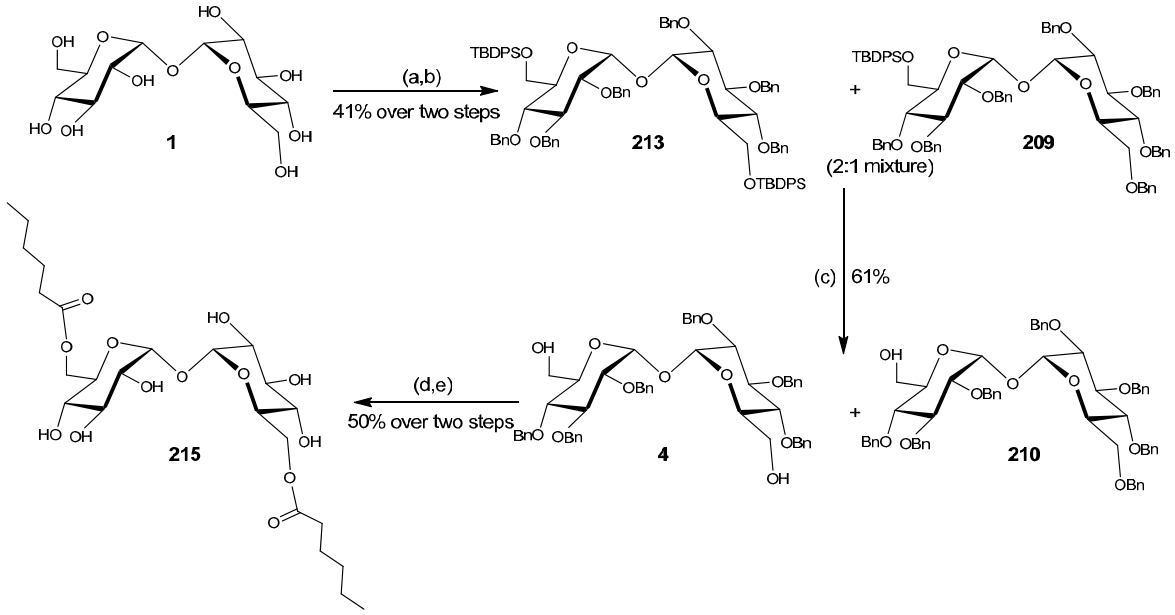
Scheme 3.2. (a) TBDPSCl, imidazole, DMF, RT, 24 h.

207 was benzylated with benzyl bromide and sodium hydride, following which, the TBDPS groups were selectively removed under tetra-*n*-butylammonium fluoride (TBAF), or acetyl chloride in anhydrous methanol (**Scheme 3.4** and **Scheme 3.3**) to afford **210**. Removal of the silyl group under acidic conditions afforded the desired compounds as well as significant production of the 6-OH acetylated compound, presumably because the free hydroxyl reacted with the acetyl chloride *in situ*. The free hydroxyl was then acylated with hexanoyl chloride in moderate yield and deprotected to afford the target 6-monohexanoyltrehalose **212**. The low yield of the acylation step was due to the addition of only one equivalent of hexanoyl chloride of which presumably some was hydrolyzed, prior to reaction with the sugar. Excess hexanoyl chloride increased the rate of reactivity, but the hydrolyzed hexanoic acid proved difficult to remove completely by column chomatography.



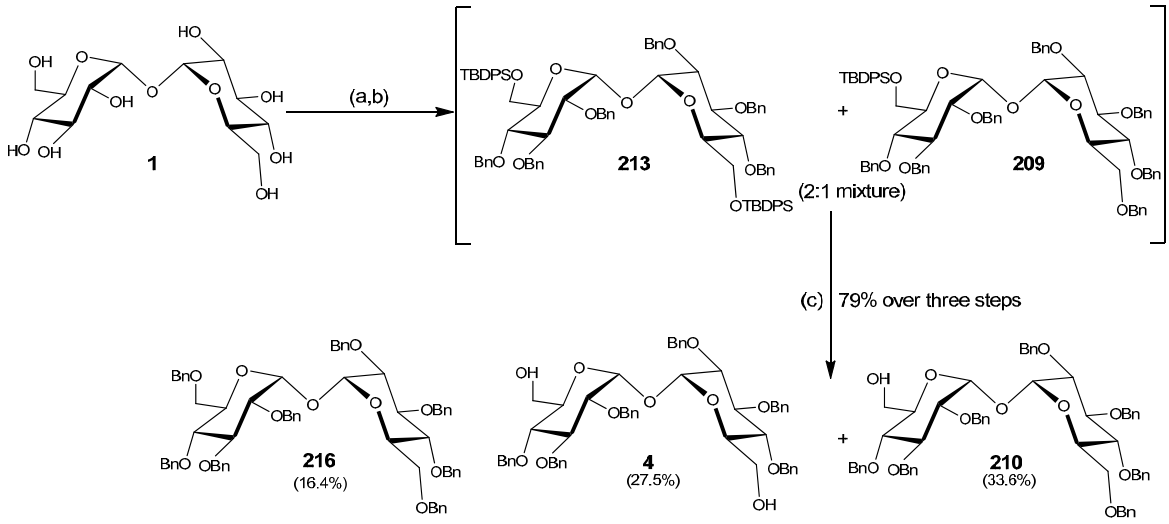
Scheme 3.3. Synthesis of 6-monohexanoyltrehalose (TMH) 212. (a) TBDPSCl, imidazole, RT, 36 h (b) BnBr, NaH, RT, 24 h (c) acetyl chloride, anhydrous MeOH, RT, 48 h (d) DCM, pyridine, hexanoyl chloride, RT, 16 h (e) 10% Pd/C cartridge, H-Cube®, 80 bar, 45 ˚C, 3 h.

6,6'-Dihexanoyltrehalose **215** (TDH) (**Scheme 3.4**) was synthesized from trehalose in a similar manner to the synthesis of TMH. The mixture of mono and diTBDPS benzylated **213** and **209** compounds, obtained through silylation and benzyl protection of trehalose, separated poorly by column chromatography. It was found that **213** and **209** could be deprotected as a mixture with TBAF, affording the corresponding mono and dihydroxyl compounds, **210** and **4**, which were easily separable by column chromatography. Acylation and deprotection of the dihexanoyl compound proceeded smoothly, affording the final compound **215** in good yield.



Scheme 3.4. Synthesis of 6,6'-dihexanoyltrehalose 215 (a) TBDPSCl, imidazole, RT, 24 h, 55% (b) BnBr, NaH, RT, 18 h. (c) TBAF, 60 ˚C, 3 h (d) Pyridine, hexanoyl chloride, RT, 30 mins, 91% (e) H2, Pd/C, ethanol, RT, 16 h, 55%.

To synthesize larger quantities of TMH and TDH and to investigate the synthesis of other 6 modified compounds, gram scale quantities of **4** and **210** were resynthesized. Without purifying the mono- and diTBDPS compounds, we found that **4** and **210** were accessible in reasonable yield over three contiguous steps (**Scheme 3.5**). The advantage of this approach is that, unlike the deprotected and protected mono and diTBDPS compounds, the hepta and hexabenzyl-trehaloses, **4** and **210**, separated easily by column chromatography. This route afforded multigram scale quantities of the desired products with a quicker synthesis and higher yields when compared with purification at each step.



Scheme 3.5. (a) TBDPSCl, imidazole, DMF, RT, 24 h. (b) BnBr, NaH, RT, 72 h (c) TBAF, 70 ˚C, 72 h. Yields for individual compounds based on (2:1) ratios of starting materials.

3.2.5.2 Mass Spectrometry[i,ii]

TMH and TDH substrates proved to be water soluble and readily turned over by Ag85. When trehalose and dihexanoyl trehalose were mixed with Ag85C, rapid formation of monohexanoyl trehalose was observed (**Figure 3.2).** Likewise, when TMH and Ag85C were mixed, TDH and trehalose formation could be observed by mass spectrometry. TMH exhibited the strongest ionization efficiency at low concentration of the three compounds, so it was decided to monitor the formation of TMH produced from the transesterification reaction between trehalose and TDH as the foundation of Ag85 enzyme assay. This reaction formed the basis of the substrate and inhibitor screening efforts against Ag85 as well as full kinetic analysis of the Ag85 transesterification reaction.

The TDH, trehalose Ag85 enzyme assay was implemented to yield full kinetic analysis of the Ag85 acyl transfer reaction, affording the previously unattainable complete dissection of all kinetic steps of the reaction. This analysis was a result of the author's development of the TDH and TMH compounds as substrates of Ag85; all kinetic analysis was conducted by Dr.

[i] All MS work was conducted in collaboration with Dr. Conor Barry. Dr. Barry implemented the MS assay(designed and synthesized above) and analyzed most of the data reported. Permission from Dr. Barry was granted to include the kinetic analysis data in this thesis.

[ii] Parts of this section have been reprinted with permission from Barry, C.S., Backus, K.M., Barry, C.E. & Davis, B.G. ESI-MS Assay of M. tuberculosis Cell Wall Antigen 85 Enzymes Permits Substrate Profiling and Design of a Mechanism-Based Inhibitor. *Journal of the American Chemical Society*, ASAP (2011). Copyright 2011 American Chemical Society

Conor Barry. As the full kinetic analysis sheds light on the later screening work, we will briefly discuss the results from the Ag85 kinetics.

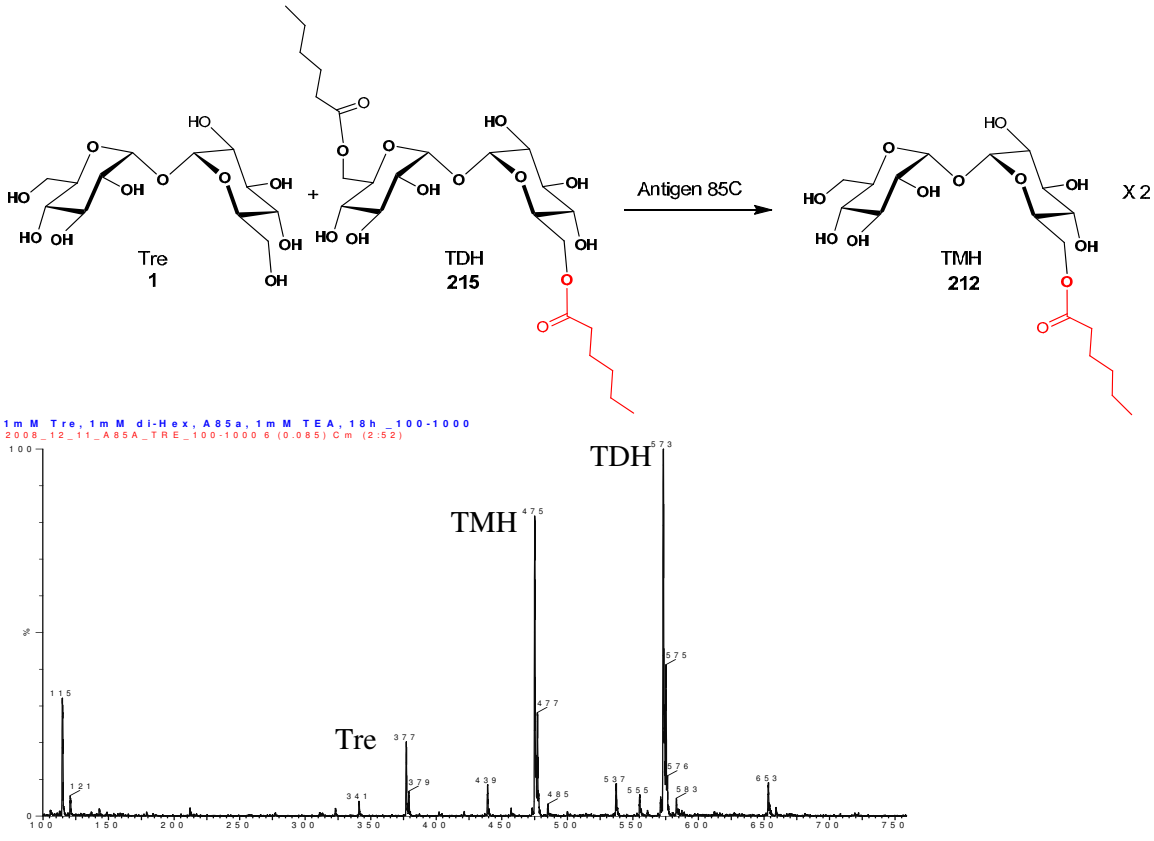


Figure 3.2. Example Ag85 enzyme assay. MS trace shows the formation of 6-monohexanoyltrehalose (TMH, 212) from 6,6'-dihexanoyltrehalose (TDH, 215) and trehalose (Tre, 1).

Kinetic analysis was performed in an automated LCMS format, where the total ion counts (TICs) of ions corresponding to trehalose, 6,6'-dihexanoyltrehalose **215** (TDH) and 6-hexanoyltrehalose **212** (TMH) and were monitored over time (**Figure 3.3**). To normalize for ionization efficiency, an internal standard (*N*-acetylglucosamine) was injected with the sample aliquot; calibration plots for each monitored mass were also generated, allowing the TIC value to be converted to a quantitative measure of concentration.

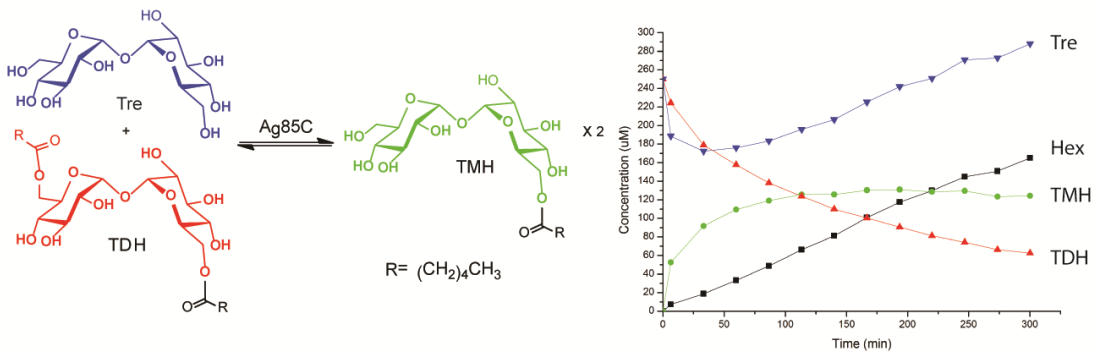


Figure 3.3. The reaction catalysed by Ag85, and a typical timecourse at high enzyme concentration. Timecourse conditions: 250 µM 1, 250 µM 215, 2 µM Ag85C, 1 mM TEA buffer (pH 7.2), 37 °C. Tre-trehalose, Hex-hexanoate, TMH-6-monohexanoyltrehalose and TDH-6,6'-dihexanoyltrehalose.

Previous reports suggest that the Ag85 enzymes should fit a bi-bi ping-pong kinetic model[22]. A modified version is shown in **Figure 3.4**. The Ag85 enzymes employ a typical Ser-His-Glu catalytic triad at the trehalose binding site[23] and it has been reported that the catalytic cycle occurs through an acyl-enzyme intermediate[3]. The catalytic serine can be covalently modified with a known α/β hydrolase inhibitor[24]. These previous reports suggested that the Ag85 enzymes would incorporate an acyl enzyme intermediate into the reaction cycle and behave in a manner similar to other lipases and proteases, which often fit the bi-bi ping pong model.

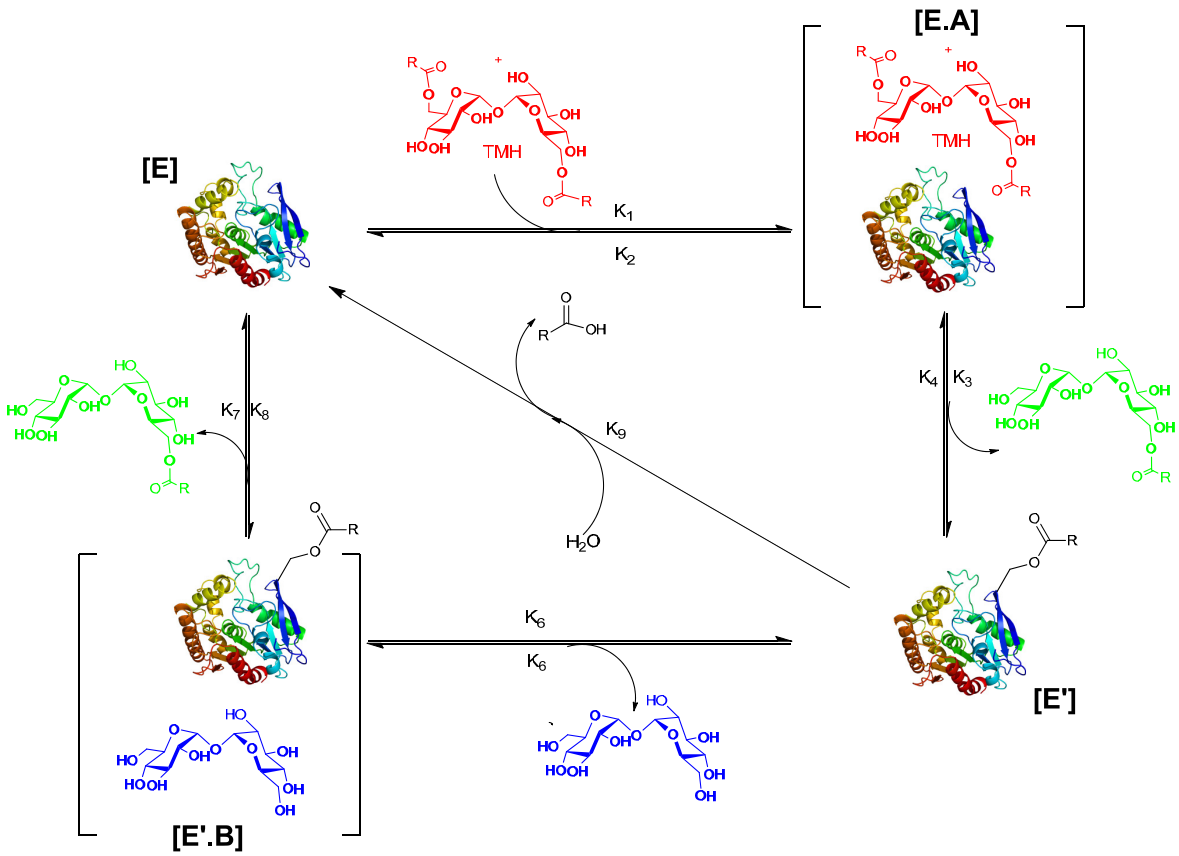


Figure 3.4. Kinetic model for Ag85 activity that includes acyl hydrolase activity.

When the Ag85C catalyzed reaction was conducted at high enzyme concentration (2 μM), previously unreported acylhydrolase activity was observed, with an accumulation of hexanoate over time, which was not apparent in the absence of enzyme **(Figure 3.3**). This hydrolysis was attributed to enzyme-mediated acylhydrolysis[25]. The absence of prior reports of acylhydrolase activity by Ag85 may be due to a greater stability of the mycolyl acyl enzyme, as compared with the hexanoyl acyl enzyme, with the shorter lipid being more labile. Alternatively, the hydrolysis may have been below the detection limit of previous assays. Dual acyltransferase/hydrolase activities have been observed for other enzymes[26,27].

To incorporate the hydrolysis into the kinetic model, an enzymatic mechanism for Ag85 was postulated that featured a ping-pong mechanism with an acyl-enzyme intermediate that is hydrolytically unstable[25] (**Figure 3.4**). Bisubstrate kinetic analysis was performed to determine the mechanism of the Ag85 reaction. Initial rates of TMH **212** formation were measured at varied TDH **215** and trehalose 1 concentrations, using *N*-acetyl-glucosamine as a

pseudo-internal standard[14,15,28]. The initial rate data were fitted by non-linear regression methods to standard kinetic models and also to a modified rate equation, which accounted for both acyltransfer and acylhydrolysis (see **Appendix 2** for derivation). The bisubstrate and double-reciprocal analysis ruled out a bi-bi ping-pong mechanism[29], as the rate of hydrolysis was too great to be accounted for in the bi-bi mechanism[30] (see **Appendix 2** for more details).

Full kinetic parameters (R^2>0.99) for Ag85C (**Table 3.1**), including K_M values for both substrates and separate maximal reaction velocities for acyl-transfer (V_t) and hydrolysis (V_h) were determined for Ag85C. The value for $K_{M(Tre)}$ found through the TDH enzyme assay correlated well with the only previous value of 8.3 μM, which was also determined under pseudo-single-substrate conditions[3]. The much higher K_M value determined for TDH, compared to that reported for TDM, is likely a reflection of the shorter lipid chain length of **215,** in comparison to the natural mycolic acid substrates. **Table 3.1** reveals considerable acyl hydrolase activity (k_{cat} 0.49 s^{-1}), which again, contrasts with prior reports[3,4].

Table 3.1. Kinetic constants for Ag85C determined by quantitative mass spectrometry analysis of bisubstrate kinetics.

V_t (μM/min)	0.70 ± 0.05	$k_{cat\ (acyl\text{-}transfer)}$ (s^{-1})	1.16
V_h (μM/min)	0.29 ±0.08	$k_{cat\ (hydrolysis)}$ (s^{-1})	0.49
$K_{M(TDH)}$ (μM)	266.8 ±41.2	$k_{cat}/K_{M(TDH)}$ ($M^{-1}s^{-1}$)	4.35x10
$K_{M(Tre)}$ (μM)	16.9 ±5.3	$k_{cat}/K_{M(Tre)}$ ($M^{-1}s^{-1}$)	6.88x10
$K_{i(TDH)}$ (μM)	355.3 ±105.6		

Pseudo single substrate conditions were employed to compare the relative catalytic efficiencies of the three Ag85 isoforms. The production of both TMH and hexanoate was monitored concurrently(**Table 3.2**)[iii]. This analysis reveals similar K_M values of all three isoforms for trehalose. Also the previously reported[2] higher activity of Ag85C compared to Ag85A or Ag85B was supported here, with an order of activity (K_{cat}) of Ag85C>>Ag85A>Ag85B. In Chapter 6 this large difference in activity between enzymes will be discussed.

[iii] vtransfer = (vTMH – vhydrolysis)/2

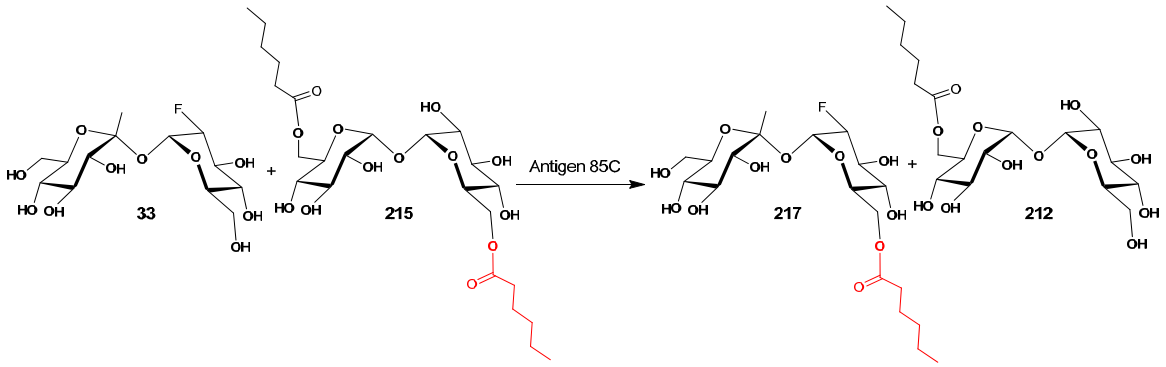
Table 3.2. Kinetic parameters for Ag85 isoforms recorded under pseudo-single substrate conditions.

Ag85 Isoform	Varied Substrate[a]	K_M(app) (uM)	k_{cat}(app) (s^{-1})	k_{cat}(app)/K_M(app) ($M^{-1}s^{-1}$)
a	**1**	175	0.014	82
b	**1**	112	0.003	30
c	**1**	62	0.182	2952

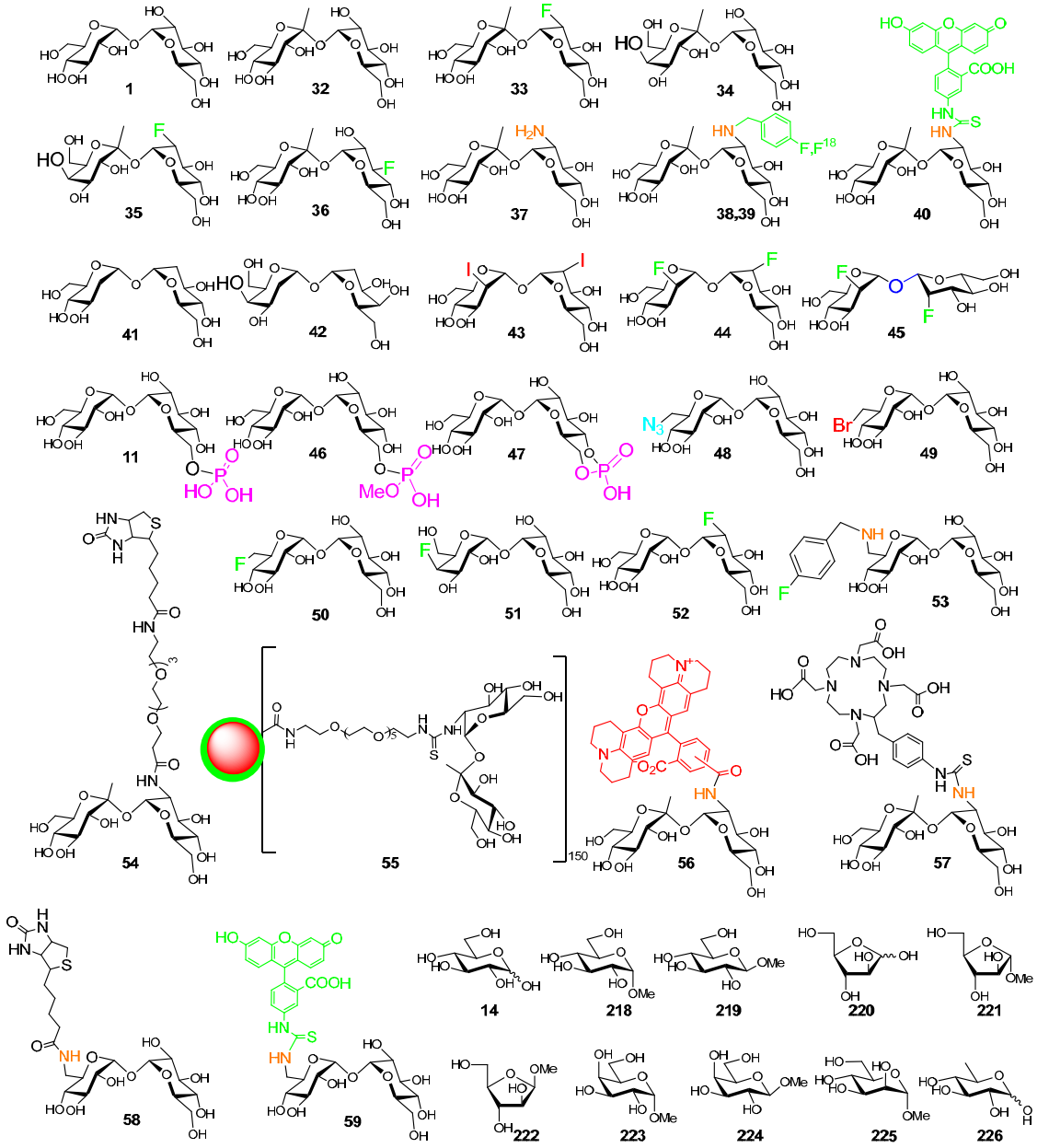
[a] Assay conditions: [TDH]=500 μM, [trehalose]=10-250 μM, [Ag85A/b]=500 nM, [Ag85C]=50 nM.

3.2.6 Green Amber Red (GAR) screen

With an accurate automated assay of Ag85 activity in place we wished to probe the selectivity of the enzyme active-site for the trehalose library (synthesis reported in **Chapter 2** and **Appendix 1**) as well as for a selection of monosaccharide substrates to evaluate the importance of the configuration of the sugar ring for substrate selectivity. Others in our lab have previously shown the utility of a qualitative green-amber-red (GAR) high-throughput screen to discover novel substrates for other bisubstrate (glycosyltransferase) enzymes[14,15,28].. The feasibility of a GAR screen with the Ag85 enzymes was investigated with non-natural 2-F-methyl-trehalose **33**. Gratifyingly, when compound **33** was mixed with TDH and Ag85C, formation of **33**-monohexanoate compound **217** was observed by MS (**Scheme 3.6**). This result confirmed, as had been suggested by the open nature of the crystal structure, that the Ag85 enzymes would tolerate modifications to the sugar scaffold.



Scheme 3.6. Example Ag85 substrate screen.



Scheme 3.7. Library of trehalose analogues and related monosaccharides screened in the Green Amber Red format (GAR) against the Ag85 enzymes. GAR screen results can be found in Tables 3.3 and 3.4.

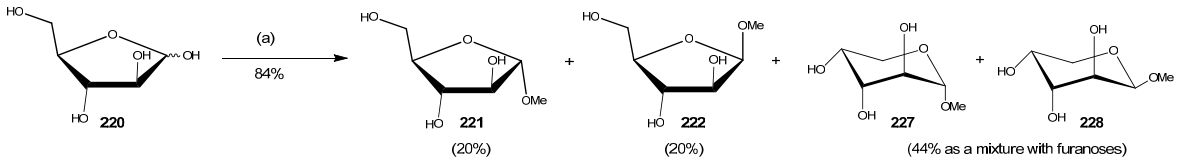
Table 3.3. GAR screen of panel of compounds presented in (Scheme 3.7) reveals a broad tolerance by the Ag85 enzymes towards modifications of the trehalose scaffold. A ratio of product formation above 75% was denoted bright green, 30-75% dark green, 5-30% amber and below 5% red as a poor substrate. MIC-minimum inhibitory concentration (concentration at which compound completely inhibits mycobacterial growth). MIC_{50}-minimum concentration that causes 50% growth inhibition. †-partial inhibition of growth.

Compound	Ag85A	Ag85B	Ag85C	MIC	MIC_{50}
1	100	100	100	>200 µg/mL	>200 µg/mL
32	32	13	111	>200 µg/mL	>200 µg/mL
33	73	1415	102	>200 µg/mL	>200 µg/mL
34	12	7	57	>200 µg/mL	>200 µg/mL
35	19	13	64	>200 µg/mL	>200 µg/mL
36	7	8	76	>200 µg/mL	>200 µg/mL
37	47	56	91	>200 µg/mL	>200 µg/mL
38	25	14	81	>200 µg/mL	>200 µg/mL
40	31	38	44	>200 µg/mL	>200 µg/mL
41	50	16	129	>200 µg/mL	>200 µg/mL
42	1	4	1	100 µg/mL	ND
43	148	54	166	>200 µg/mL	>200 µg/mL
44	66	24	98	>200 µg/mL	>200 µg/mL
45	38	35	82	>200 µg/mL	>200 µg/mL
11	6	4	28	>200 µg/mL	>200 µg/mL
46	37	24	113	†>200µg/mL	ND
47	18	17	94	>200µg/mL	>200µg/mL
48	37	13	49	†>200µg/mL	135 µg/mL
49	102	92	99	200µg/ml	134 µg/mL
50	149	412	91	200µg/ml	75 µg/mL
51	229	162	191	100µg/ml	50 µg/mL
52	115	156	274	>200 µg/mL	>200 µg/mL
54	16	6	49	ND	ND
59	9	10	15	ND	ND
58	16	7	50	ND	ND
53	25	59	24	ND	ND
55	ND	ND	ND	>200 µg/mL	>200 µg/mL
56	ND	ND	ND	>200 µg/mL	>200 µg/mL
57	ND	ND	ND	>200 µg/mL	>200 µg/mL

Mass spectrometric analysis provided a rapid means to screen Tre analogs and a selection of *Mtb*-related monosaccharides, as substrates of all Ag85 isoforms (**Scheme 3.7**, **Table 3.3** and **Table 3.4**). A calibrated green-amber-red (GAR)[17,28] screen for catalysis of transfer of the acyl chain from natural Tre to unnatural Tre-analog (Tre*), using these pure synthetic substrates allowed determination of relative reactivity ratios Tre*:Tre (**Table 3.3** and **Table 3.4**), based on the ratio of total ion counts (TICs). For this assay, The trehalose analogues (500 µM) and TDH (500 µM) were incubated with Ag85 (100 nM) in 1 mM TEA at 37 ^{o}C.

Negative controls omitting the Ag85 were performed in all cases. After incubation for 160 min a sample of the reaction mixture was analyzed by MS, scanning the mass range 150-900Da and recorded the ratio of ratio of product signal measured for mono-hexanoyl[Tre*] peaks/mono-hexanoyl-Tre. It was found that all substrates were processed by all Ag85 enzymes but to varying degrees. A ratio above 75% was denoted bright green, 30-75% dark green, 5-30% amber and below 5% red as a poor substrate. These ratios were considered to be more consistent than S/N, as S/N varies substantially depending on where the noise is selected in the spectrum. This semi-quantitative method afforded relative reaction velocity (v_{rel}) and allowed for rapid screening of the trehalose library and identified a broad plasticity of the Ag85 enzymes towards a wide variety of modifications to the sugar scaffold.

This screen revealed a strong selectivity for Tre-like disaccharides (**Table 3.3**) over monosaccharides (**Table 3.4**) such as D-glucose (**14**), but a striking plasticity for all tested Tre analogs. Modifications on every position of the sugar scaffold were tolerated including even C-1 methyl groups at the crowded anomeric linkages in **32-40,** positive charges, such as in the 2-amino-trehalose **37,** and even stereochemically 'incorrect' 2,2'-di-fluoro-αβ-manno-trehalose **45**. Even a previously reported inhibitor[2], 6-azido-Tre **48** and putative tetrahedral transition state analogs, such as trehalose-6-phosphate **11**, were also processed. Unexpectedly, as the 6 position participates in the mycolation reaction, all three enzymes tolerate 6 hydroxyl modifications, including bulky groups. Bulky appendages off of the 2 amino trehalose compound **37** (fluorobenzyl **38,** fluorescein **40**, biotin, **54**) are processed. Notably, galactose configuration in **34** and **42** reduces activity. Likewise, 2,2'-dideoxy in **41** and **42** reduces activity. The combination of 2,2'-dideoxy-galacto-trehalose in **42** abolishes enzyme activity. As the assay was conducted in the negative ionization mode, assay of positively charged compounds, including rhodamine-trehalose **56** and DOTA-trehalose **57**, was attempted but not successful. In summary the GAR assay reveals that the Ag85 enzymes tolerate a wide range of modifications to the disaccharide trehalose scaffold.



Scheme 3.8. Synthesis of arabanofuranosides (a) Acetyl chloride, anhydrous methanol, 0 °C, 3 days.

Having screened all trehalose analogues, we turned our attention towards monosaccharides and thought to explore a GAR with some anomeric free monosaccharides as well as capped anomeric methyl sugars. Methyl glucosides **218** and **219**, galactosides **223** and **224** and mannoside **226** are all commercially available. However, the methyl arabinosides **221** and **222**, which mimic of the outermost portion of the arabinogalactan, are not commerically available. Arabinose **220**, as a free sugar, sits as an open chain, the furanose and the pyranose. The poor turnover of arabinose observed in the GAR screen may be due to the mixtures of sugar configurations in solution. Following a procedure by Sanki et al.[31,32], the α and β arabinofuranosides **221** and **222** were synthesized. Arabinose was treated to acid under anhydrous conditions at 0 °C for three days and was allowed to warm briefly to RT before reaction was quenched. Column chromatography afforded primarily the α and β furanosides in 20% yield for each as well as a 44% yield of the undesired pyranosides **227** and **228** as a mixture with trace amounts of the desired furanosides and (**Scheme 3.8**).

Table 3.4. GAR of a panel of monosaccharides reveals a strong preference for disaccharide sugars (Table 3.3), compared with monosaccharides. As with disaccharides, a strong preference for gluco rather than galacto conformation is revealed. A ratio of product formation above 75% was denoted bright green, 30-75% dark green, 5-30% amber and below 5% red as a poor substrate.

Compound	Ag85A	Ag85B	Ag85C
14	18	19	45
218	16	14	45
219	5	6	44
220	6	28	8
221	12	11	21
222	4	2	14
223	1	0.05	0.01
224	0.03	0.04	0.01
225	2	2	16
226	4	0.7	3

In contrast with the disaccharides in, monosaccharides (**Table 3.4**) are transesterified with lower efficiency. It was found that free glucose **14** and methyl glucosides **218** and **219** were utilized less efficiently than the disaccharides, indicating that the enzyme makes contacts with the second pyranose ring of trehalose to facilitate its turnover. Nonetheless, both glucose and the arabinofuranosides **220** and **221** were processed relatively efficiently compared with other monosaccharides tested. This is consistent with a possible role of Ag85 in generating glucose monomycolate (GMM)[33] and in mycolation of arabinose found in the *Mtb* arabinogalactan[31,34]. For both glucose and arabinose anomeric selectivity was observed whereby the α-anomer was preferred by Ag85A and B, with Ag85C showing little or no anomeric discrimination. The nearly zero v_{rel} value determined for galactosides **223** and **224** for all three Ag85 isoforms is further evidence that the epimeric configuration at C-4 is crucial to enzyme activity. Conversely, α-methyl mannoside **225** was utilized by Ag85C, suggesting that some degree of flexibility with the C-2 pyranose substituent is possible, as was evidenced with the screen of C-2 modified trehalose analogues. Xyloside **226** was not processed by the enzymes, confirming that acyltransfer is to the C-6 hydroxyl group. These data suggest that sugar stereochemistry may be used as a recognition motif for inhibitor design.

In summary, the GAR screen revealed a differential selectivity of the Ag85 isoforms towards both mono and disaccharides. Ag85C exhibits a broad tolerance towards modifications of the sugar scaffold, including relatively efficient esterification of monosaccharides. In contrast, Ag85A and B process monosaccharides inefficiently compared to Ag85C. **Table 3.3** suggests that Ag85B has a specificity for Methyl-2-F-trehalose **33**, whereas Ag85A and C seem to be more selective for the 4-F-epi-trehalose **51** and Ag85C appears to process a variety of C-6 modified compounds greater proficiency than Ag85A or B. In future work, this differential selectivity may be co-opted to differentiate between subpopulations of mycobacteria that express more of one particular isoform of Ag85.

3.3 Resistance of modified trehalose to trehalase degradation

The broad substrate specificity of the Ag85 enzymes (**Table 3.3** and **Table 3.4**) suggested that trehalose derivatives such as FITC-Tre could potentially be employed as probes of Ag85 activity in cell culture and animal studies. However, both mammals and bacteria encode trehalase enzymes, glycosidases that hydrolyse the glycosidic linkage of trehalose to afford to units of glucose. Mammalian trehalase is a transmembrane protein and has been shown to

exist as a brush border enzyme mostly in kidney and liver in rabbits, a model animal of TB[35]. Mycobacterial trehalase is a cell wall bound protein that is known to be highly selective for trehalose[36]. Activity of the host trehalases towards the trehalose probes could non-productively degrade these compounds, prior to the probe's arrival at the mycobacteria. A trehalase assay was conducted towards several representative compounds and this assay conclusively ruled out enzymatic degradation of trehalose analogues by host and pathogen's endogenous trehalases.

Crystal structures of mycobacterial, murine and rabbit trehalases have not been reported, but the X-Ray structure of the *E. coli* trehalase has been solved[37] (PDBs 2WYN, 2JJB, 2JF4, 2J60) and sequence homology suggests that the *E. coli* structure should show structural homology towards mammalian trehalases. A comparison of the sequences of mouse and *E. coli* trehalases are 37% conserved (**Figure 3.5**). More importantly, nearly all residues within 5 angstroms of the trehalose are conserved in both species (residues in bold in **Figure 3.5**).

```
37.4% identity in 404 residues overlap; Score: 623.0; Gap frequency: 5.0%

ecoli        108 SLLPLPEPY**VVPGGRFREVYYWD**SYFTMLGLAESGHWDKVADMVANFAHEIDTYGHI**PNG**
mouse        151 SLIYSKHPFIVPGGRFVEFYYWDSYWVMEGLLLSEMASTVKGMLQNFLDLVKTYGHIPNG
                 **     *  ****** * ******  * **  *     *  *  **     ********

ecoli        168 NR**SYY**LS**RSQPP**FFALMVELLAQHEGDAA-LKQYLPQMQKEYAYWMDGVENLQAGQQEKR
mouse        211 GRIYYLQRSQPPLLTLMMDRYVAHTKDVAFLQENIGTLASELDFW--------TVNRTVS
                  * *** *****   **      *  * * *         *   *

ecoli        227 VVKLQDGTLLNRYWDDRDT**PRPES**WV**ED**IATAKSNPNRPATEIYRD**L**RS**AAASGWDF**SSR
mouse        263 VVSGGQSYVLNRYYVPYGGPRPESYRKDAELANSVPEGDRETLWAELKAGAESGWDFSSR
                 **       ****      *****   *   * * *          *   * ********

ecoli        287 WM---DNPQQLNTLRTTSIVPVDLNSLMFKMEKILARASKAAGDNAMANQYETLANARQK
mouse        323 WLVGGPDPDLLSSIRTSKMVPADLNAFLCQAEELMSNFYSRLGNDTEATKYRNLRAQRLA
                 *      *  *   **   ** ***      *          *    *  *  *   *

ecoli        344 GIEKYLWNDQQGWYADYDLKSHKVRNQLTAAALFPLYVNA----AAKDRANKMATATKTH
mouse        383 AMEAVLWDEQKGAWFDYDLEKGKKNLEFYPSNLSPLWAGCFSDPSVADKALKYLEDSKI-
                   *  **  * *   ****   *         * **          * * *     *

ecoli        400 LLQPGGLNTTSVKS**GQQWD**APNG**WAPL**QWVATEGLQNYGQ---**KEVA**MDISWHFLTNVQH
mouse        442 LTYQYGIPTSLRNTGQQWDFPNAWAPLQDLVIRGLAKSASPRTQEVAFQLAQNWIKTNFK
                 *    *  *     ***** ** *****     **         ***
```

Figure 3.5. Sequence homology between murine and *E. coli* trehalases. Stars indicate conserved residues. Bolded residues indicate residues that contact the substrate. Alignment generated by T-Coffee.

The crystal structure of the *E. coli* trehalase pdb 2JG0[37] revealed a deeply embedded, seemingly selective binding pocket (**Figure 3.5**) with binding interactions between the protein and nearly all components of the sugar scaffold. In fact, when protein was viewed entirely in a surface filling format, trehalose mimic thiatetrazolin was completely engulfed by

the protein and is not visible from the exterior of the protein. **Figure 3.6** shows a cutaway view to reveal the buried active site within the protein.

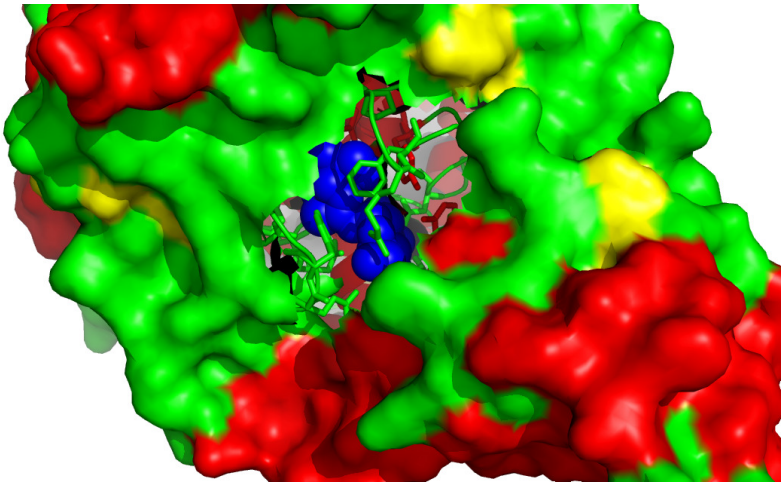


Figure 3.6. ***E. coli*** **trehalase (Pdb 2JG0) has a very embedded binding pocket. Blue spheres- 1-thiatetrazolin (trehalase resistant trehalose mimic). Figure generated in Pymol.**

Although the crystal structure of the trehalase was suggestive that the trehalase enzymes would not process modified trehalose, we employed a commercially available enzyme assay (Sigma Aldrich) for trehalase activity in order to rule out degradation by trehalases as a source of complication for probe development[iv]. Commercially available porcine kidney trehalase was used. Five of our probes, 2-fluoro-trehalose (2F-Tre, **50**), 4-fluoro-epi-trehalose (4F-Tre, **51**), 6-fluoro-trehalose (6F-Tre, **50**), fluorobenzyl-trehalose (2FB-Tre, **38**) and FITC-trehalose (FITC-Tre, **40**) were assayed along with the natural substrate trehalose as the control. 2F-Tre, 4F-Tre, 2FB-Tre and FITC-Tre displayed no activity with trehalase. 2FB-Tre, 4F-Tre and FITC-Tre likely fail to bind to the active site properly due to the bulky substituent or wrong stereochemistry blocking enzyme binding. 6-fluoro-trehalose **50** showed diminished activity, possibly due to the loss of binding energy. In the case of 2F-Tre, the powerful electron-withdrawing nature of fluorine at 2-position has been reported to destabilize the positively charged transition state of the reaction[38,39]. These results suggest that most trehalose analogues synthesized should serve as poor or negligible substrates of the mammalian trehalases.

[iv] The Trehalase assay work was done in collaboration with Dr. Seung S Lee.

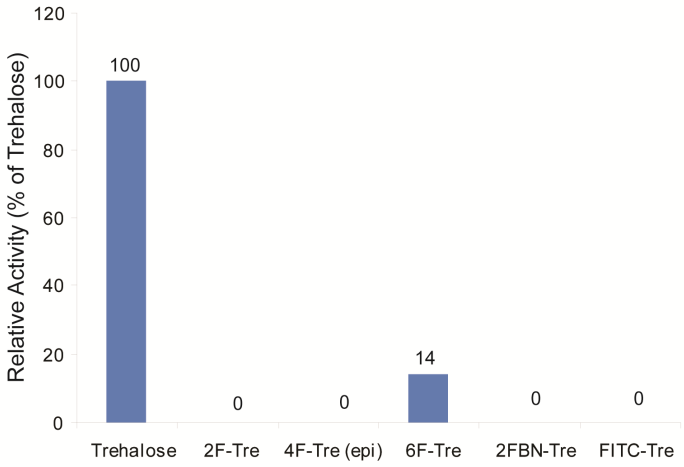


Figure 3.7. Relative activity of porcine trehalase towards representative trehalose analogues shows resistance to degradation for all compounds. Fluorobenzyl (FB), Fluorescein isothiocynate (FITC).

3.4 Design of an Ag85 inhibitor

The high throughput Ag85 enzyme assay was employed in the discovery of an Ag85 inhibitor. An efficient inhibitor of the Ag85 enzymes could serve as a chemical genetics-type probe, allowing for a knock down of Ag85 activity and dissection of the roles of Ag85 in cell wall biosynthesis without the need to produce the *Ag85* knockout strains of *Mtb.* As the triple *ag85Abc* knockout of *Mtb* is thought to be lethal, Ag85 inhibitors could also serve as potential anti-mycobacterial drug candidates.

3.4.1 The inhibitory properties of trehalose-6-azide 48

Prior reports of trehalose-6-azide **48** as an inhibitor of both the Ag85 enzymes[2] *in vitro* and of mycobacterial growth (albeit a weak one with a reported MIC against a related mycobacterial species of 200 µg/mL)[2] prompted us to examine the possibility that some of the trehalose analogs also induced growth inhibition through inhibition of Ag85 or other mechanisms. Minimum inhibitory concentration (MIC) analysis was carried out on key representative compounds **(Table 3.3).** 2,2'-dideoxy-lyxo-Tre **42**, which was not a substrate for any of the three Ag85 isoforms, showed the most significant inhibitory effect (MIC 100 µg/mL) against mycobacterial growth. C-6 modified compounds that were processed well by Ag85 inhibited growth at higher concentrations (MIC 100-200 µg/mL). An explanation of the poor MICs of these compounds is that many of them serve as competitive substrates for all Ag85 isoforms,

suggesting that some (or all) of the observed antibacterial activity may arise from downstream events after Ag85-mediated mycolyltransfer.

Given that the previously reported Ag85 inhibitor, **48**, actually served as a substrate of the Ag85 enzymes in our assay, with approximately 50% efficiency of native trehalose, the inhibitor nature of this compound was probed further. A survey of publications on the inhibition of Ag85 enzymes by trehalose-6-azide[2], revealed that Belisle et al. reported an unknown compound X that was found in the lipid extract from mycobacteria treated with trehalose-6-azide and concurrently radiolabeled with ^{14}C-acetate. This unknown compound has the approximate R_f value that would correspond to monomycolyl trehalose-6-azide, further supporting that the apparent inhibitory affects are a result of trehalose-6-azide serving as a competitive substrate.

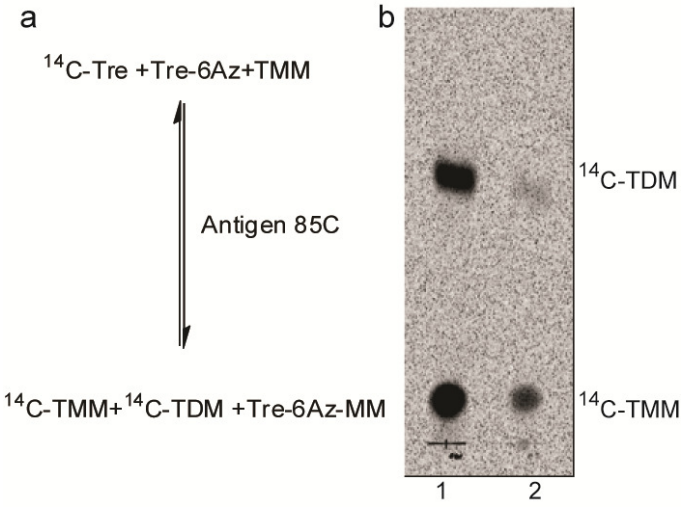


Figure 3.8. (a) Scheme of radioTLC reaction run in the presence of trehalose-6-azide 48 (Tre-6Az) illustrates the potential to generate unlabeled trehalose-6-azide-monomycolate (Tre-6az-MM) (b)Radiographic TLC assay of Ag85 activity. Lane 1 reflects transesterification of ^{14}C-trehalose and TMM to generate ^{14}C-TMM and TDM. Lane 2 reflects the diminished transesterification that occurs in the presence of 48[v].

Consistent with previous reports, we also found that in *in vitro,* radioTLC assays trehalose-6-azide does inhibit the formation of ^{14}C-TDM and ^{14}C-TMM (**Figure 3.8).** ^{14}C-trehalose, TMM and Antigen 85C were mixed in aqueous buffer and transesterication is allowed to proceed, generating ^{14}C-TDM and ^{14}C-TMM. Reaction was then quenched and extracted into chloroform:methanol to remove all remaining ^{14}C-trehalose. In the presence of trehalose-6-

[v] Trehalose-6-azide inhibition assays were conducted in collaboration with Dr. Kapil Tahlan.

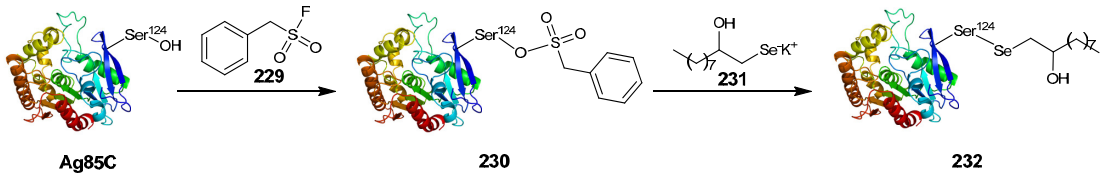
azide **48**, transesterification of the ^{14}C-trehalose is dramatically reduced. However, this assay fails to account for the trehalose-6-azide transesterification, as the non-radiolabeled monomycolyl-trehalose-6-azide product will not be detected by the phosphorimager scan.

3.4.2 Protein modification- covalent inhibitors

Serine proteases[40] and lipases are known to react with phenylmethylsulfonyl fluoride (PMSF, **229**), forming irreversible, covalent acyl enzyme complexes. Ag85C has previously been shown to react with a diethylphosphonate inhibitor[24]. Based on these results, Ag85 was reacted with PMSF to determine whether PMSF reactivity could inform the design of additional, covalent Ag85 inhibitors, based on sulfonyl or phosphoryl reactive groups to serve as active site-directed irreversible inhibitors of the Ag85 enzymes.

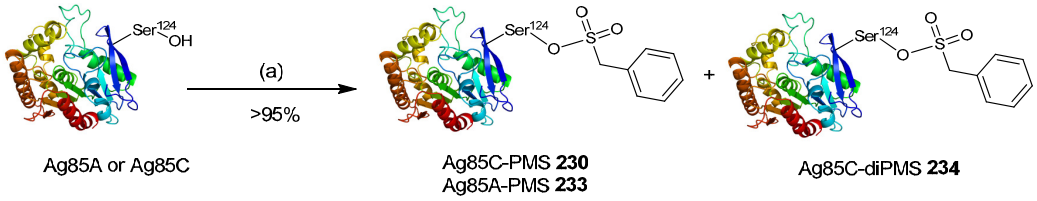
Inhibition of the Ag85 enzymes with PMSF also offered a route towards novel crystal structures, which may further inform the development of inhibitors and our understanding of the mycolyl transferase catalytic cycle. Literature precedent[41] and experience within the Davis group (unpublished) suggested that the PMSF acyl enzyme could be displaced with a variety of nucleophiles (thiolates, selenides, etc), affording the seleno or cysteine Ag85 mutant, through 'chemical mutagenesis.' PMSF modification of the catalytic serine would allow the formation of covalent thio or seleno-enzyme complexes, which could then be crystallized. No crystal structures of Antigen 85, covalently bound to either trehalose or a lipid are known and it was hypothesized that crystallization with such ligands could force the Ag85 enzymes to assume previously unreported conformations.

Analysis of the Ag85 crystal structures revealed that in the structure bound to diethylphosphonate analogue[24], compared with all other reported structures[24,42-44], a dramatic conformational shift occurs, involving a loop (helix eight) of antigen 85 adjacent to the active site, potentially opening up a channel through which TMM might be able to travel to the active site (see **Chapter 6** for more details). To confirm the location of the mycolate in the crystal structure and to evaluate the helix eight movement, it was thought that the Ag85 enzymes could be modified and crystallized with a mimic of a mycolate, installed through a selenide displacement of a PMS ester (**Scheme 3.9**).



Scheme 3.9. Proposed seleno modification to covalently link a mycolyl mimic, featuring the characteristic mycolate β-hydroxyl, to the active site in order to trigger helix 8 movement (putative transition state).

PMSF modification of Ag85 proceeded slowly, relative to reports of PMSF modifications of proteases, and required the addition of nearly 300 equivalents of PMSF in acetonitrile. This result corroborates with our kinetic data that suggest that Antigen 85A,B and C are a relatively slow enzymes. Initially, modification of both Ag85A and C to afford **230** and **233** proceeded to only about 20% completion (**Figure 3.10** and **Figure 3.9**). Notably, both Ag85A and gluconylated Ag85A reacted with PMSF, which further confirms that both are active serine esterases (**Scheme 3.10**).



Scheme 3.10. Reaction of Ag85A/C with PMSF forms covalent acyl enzyme intermediate (a) PMSF (275 eq) in MeCN, TEA 20 mM, pH 7.3, 4 °C, 16 h

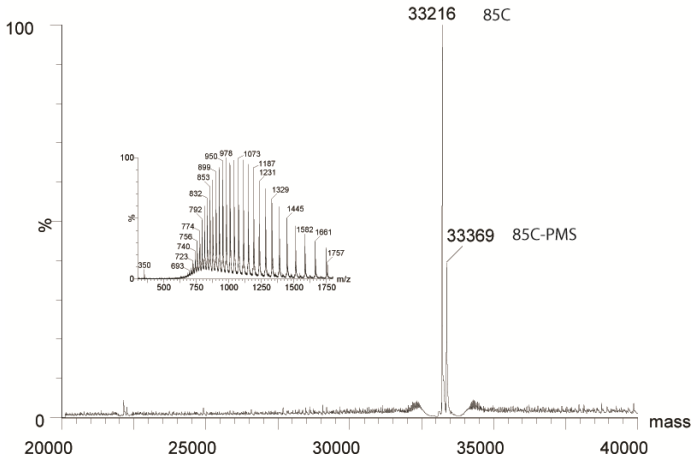


Figure 3.9. Modification of Ag85C with PMSF affords approximately 30% modification. Calculated mass Ag85C: 33217 Found: 33216; Calculated mass Ag85C-PMS 230:33370, Found: 33369.

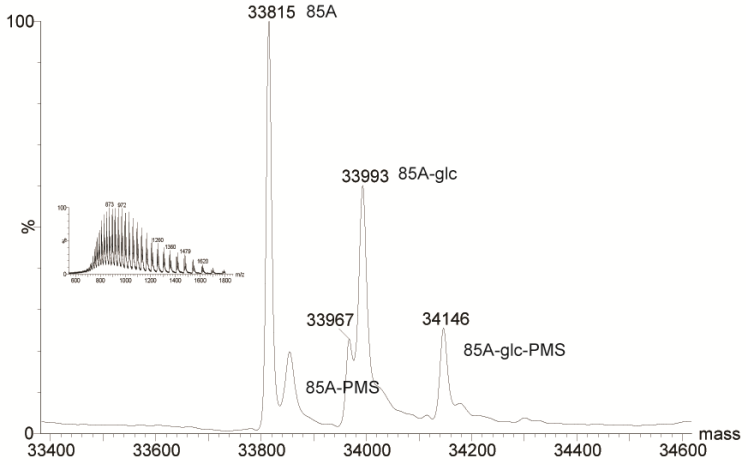


Figure 3.10. Modification of Ag85A and gluconylated Ag85A with PMSF affords roughly 30% modification after 1 hours at RT. Calculated mass Ag85A: 33814 Found: 33815; Calculated mass Ag85A-glc:33993, Found: 33993; Calculated mass Ag85A-PMS: 33814, Found: 33815 Calculated mass Ag85A-PMS: 33968, Found: 33967; Calculated mass Ag85A-glc-PMS:34145, Found: 34146;

Incubation of Ag85C overnight at 4 °C with further equivalents of PMSF forced the reaction to go to completion, affording PMS modified Ag85C and a small amount of doubly modified enzyme (**Figure 3.11**). Further attempts at reaction optimization failed to eliminate the problem with double modification of the Ag85C and it was thought that one of the numerous surface exposed lysine may react with the PMSF.

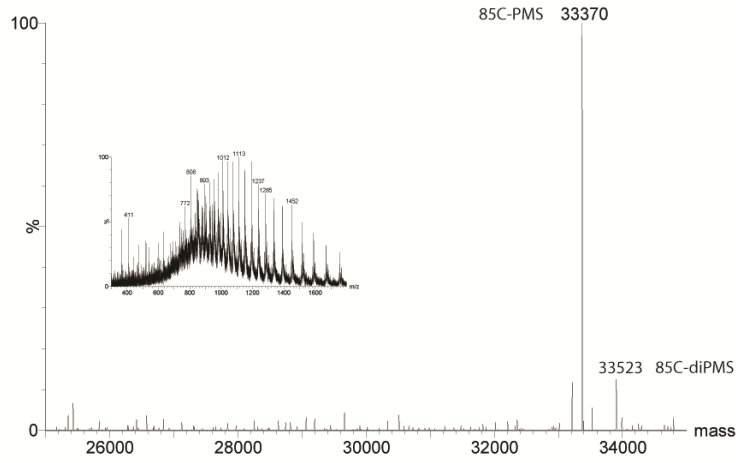
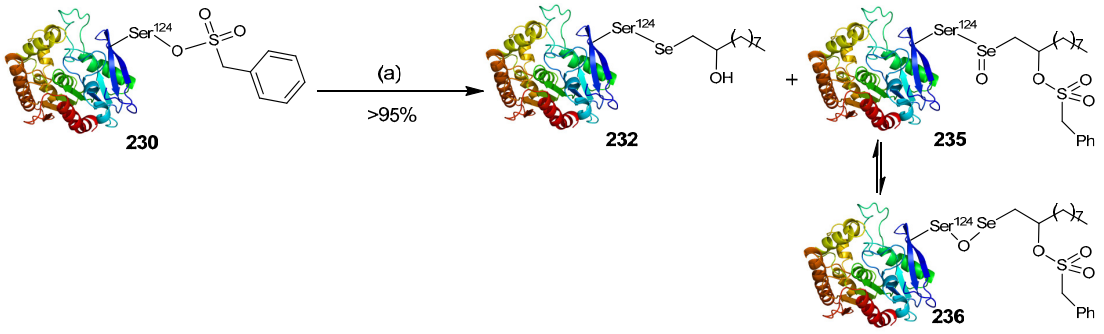


Figure 3.11. Fully modified Ag85C after additional equivalents of PMSF. Calculated mass Ag85C-PMSF:33371, Found: 33370. Calculated Ag85C-diPMS:33524 Found:33523.

With the PMSF modified protein in hand, displacement of the PMS with a bis(2-hydroxydecyl)-1,1'-diselenide reagent was attempted[vi]. The reaction completely consumed the starting protein and yielding the desired product by mass and a mixture of alternative seleno-modified protein species. The major second species in the mixture was identified as a the oxidized selenide species with additional modification of the β-hydroxyl by PMS, presumably by a rearrangement that occurs during the displacement.



Scheme 3.11. Seleno-ester synthesis with Antigen 85C-PMS (a) bis(2-hydroxydecyl)-1,1'-diselenide 231, MeOH, 12 h, 22 °C.

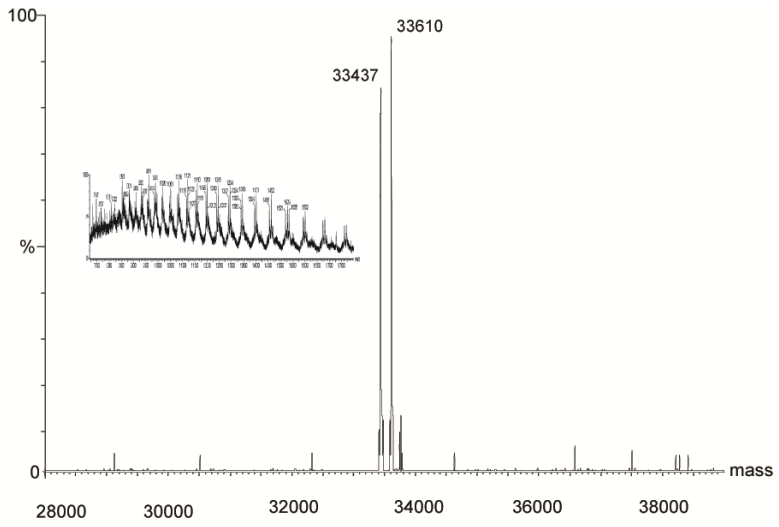


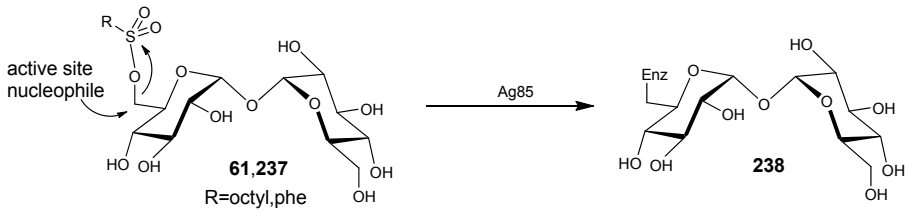
Figure 3.12. Fully modified Ag85C after reaction with potassium selenide reagent 231. Calculated mass Ag85C-Seleno 232:33436, Found: 33437. Calculated mass 235,236: 33606, Found: 33610.

[vi] The PMSF and selenide modification experiments were conducted in collaboration with Dr Omar Boutureira

Additional efforts to try to force this reaction towards one product proved unsuccessful. The complex product mixture, and the potential for the formation of the non-productive selenoxide product (hydrolysis of which would afford the starting catalytic serine, rather than seleno-serine), forced a reevaluation of the feasibility of the synthesis of seleno-Ag85C.

3.4.3 Alternative inhibitor strategies

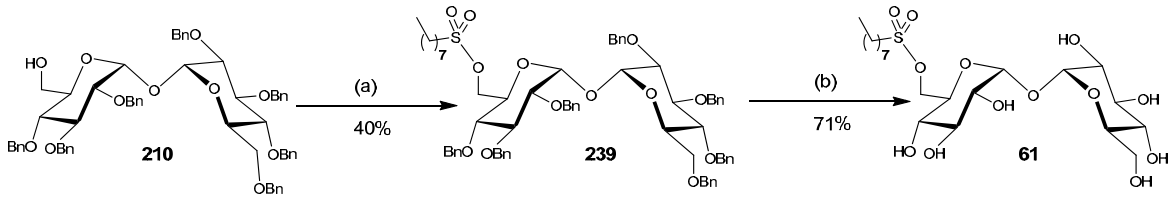
Two alternative strategies for covalent inhibitors, based on the trehalose scaffold were then investigated. First we considered the utility of sulfonyl based leaving groups, whereby any nucleophile in the active site could displace the sulfate and become covalently attached to the trehalose (**Scheme 3.12**). This approach has previously been successful with active site cysteine, serine as well as aspartate and glutamate and substrates[45,46]. Tosylates have been used *in vivo* to inhibit specific enzymes and inhibition has been monitored by fluorine nmr[46].



Scheme 3.12. Proposed reaction of Ag85 with sulfonyl trehalose

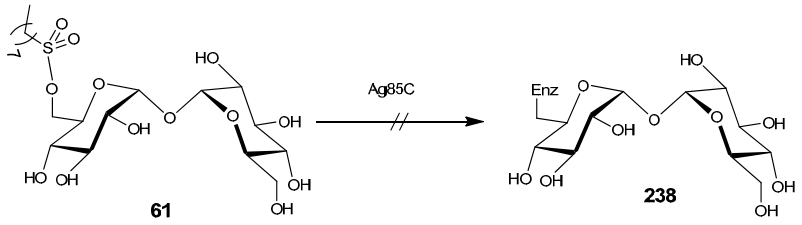
Phosphoryl fluorides were chosen as a second inhibitor strategy. These are known to be reasonably stable, react covalently with esterases and have been used by Cravatt et al. in esterase pulldowns[47-49]. Fluorophosphoryl-trehalose could serve as a reactive transition-state mimic, covalently modifying the active site serine and killing enzyme activity.

Octane-sulfonyl trehalose compound **61** was synthesized from the corresponding sulfonyl chloride and **210** in two steps and 28% yield (**Scheme 3.13**). When mixed with Ag85C, even in vast excess, we observed no covalent-enzyme species by liquid chromotography time-of-flight (LCT) mass spectrometry (**Scheme 3.14**). Enzyme assays with sulfonyl trehalose revealed a low substrate turnover, but this inhibition was likely due to competitive inhibition rather than covalent active site inhibition.



Scheme 3.13. Synthesis of Octanesulfonyl-trehalose (a) octanesulfonyl chloride, pyridine, RT, 2 h (b) H_2, Pd/C, RT, 16 h.

It was postulated that the octylsulfonyl substituent may not be a good enough leaving group to react with Ag85, due, perhaps, to the absence of a conjugated group in the side chain. However, the equivalent tosyl-trehalose was synthesized on a small scale as a mixture of the 4- and 6-tosyl trehalose compounds and likewise proved unreactive with Ag85C.

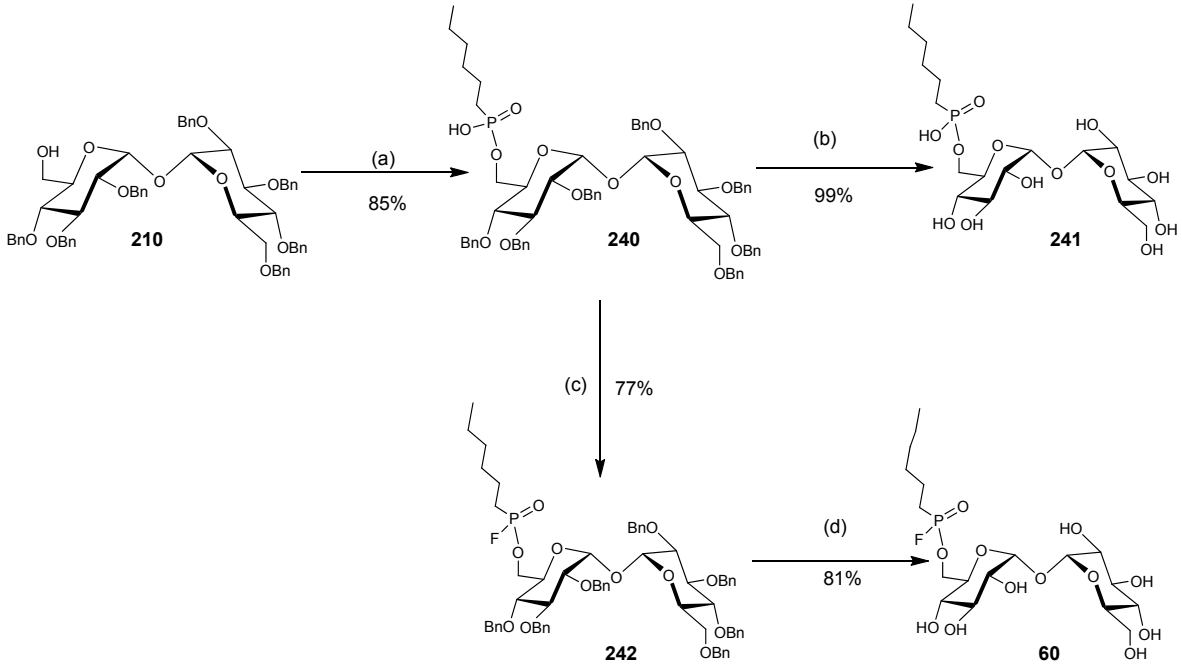


Scheme 3.14. Attempted reaction between sulfonyl trehalose 61 and Ag85C. 238 was not observed by LCT MS.

3.4.3.1 Synthesis of Ag85 inhibitor trehalose fluorophosphonate 60

Phosphoryl fluorides, known inhibitors of serine hydrolases[50] lipases[51] and proteases[52], were chosen as an alternative strategy to covalently knock out Ag85 activity. Ag85C has been shown to be covalently modified by diethylfluorophosphonate[24]. However, diethylfluorophosphonate is not selective for Ag85 over other esterases. It was hypothesized that, through the introduction of suitable substituents to a diethylphosphonate or PMSF-like template, selectivity for Ag85C over other serine hydrolases could be achieved. The survey of Ag85 substrate preferences presented in **section 3.2.6** revealed that trehalose is a superior substrate to any of the monosaccharides. To further mimic the natural mycolyl-trehalose substrates, we designed phosphonate **241** and fluorophosphonate **60**, which bear hydrophobic alkane chains (**Scheme 3.15**).

The target trehalose fluorophosphonate **60** was synthesized in three steps from heptabenzyl-trehalose analogue **210** (**Scheme 3.15**). Protected trehalose fluorophosphonate **242** was synthesized in one step from and hexylphosphonic dichloride. It was initially throught that fluoride would react directly with the trehalose hexylphosphonate chloride generated *in situ*, during the reaction between **210** and hexylphosphonic dichloride. The acylation reaction proceeded slowly in anhydrous pyridine and, after 48 hours, only **240** was detected by TLC and MS, presumably due to hydrolysis of the labile chlorophosphonate by trace water in the reaction.



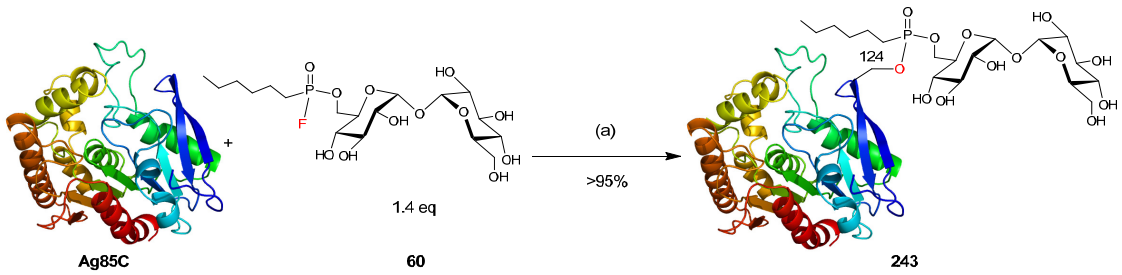
Scheme 3.15. Synthesis of trehalose fluorophosphonate 60. (a) hexyl-phosphonic dichloride, pyridine, RT, 48 h. (b) H_2 Pd/C, RT, 12 h. (c) BTFFH, DIPEA, DMF, RT, 18 h (b) H_2 Pd/C, RT, 12 h.

Installation of the phosphoryl fluoride was then attempted through fluorination of the resulting trehalose phosphonate **240** by a selection of fluorination reagents. Following literature precedent, fluorination of the free phosphate was first attempted with diethylaminosulfur trifluoride (DAST). This reaction afforded a complex mixture of products that could not be fully separated. Fortunately, it was found that treatment with fluoro-N,N,N',N'-bis(tetramethylene)formamidinium hexafluorophosphate, BTFFH, a peptide coupling reagent. and DIPEA afforded the desired trehalose fluorophosphonate **242** in good yield. **242** could then be deprotected by standard hydrogenation with very minimal hydrolysis

of the phosphoryl fluoride occurring during deprotection. This route (**Scheme 3.15**) also afforded the free trehalose hexylphosphonate **240**, which served as a good compound with which to compare the fluorophosphonate **242** in terms of inhibition of Ag85 and MIC.

3.5 Reaction of Ag85 with potential inhibitors

Ag85 was found to react with phosphoryl fluorophosphonate **242**, affording covalent, irreversible enzyme inhibition. Ag85C was incubated with stochiometric amounts of inhibitor **242** and conversion to modified protein was monitored by LCT mass spectrometry. 1.4 equivalents of **242** completely modified the Ag85C within 20 minutes (**Scheme 3.16**), affording conversion to the acyl-Ag85 complex in nearly quantitative yield. Reaction was easily scalable and several repeats generated 10 milligrams of modified protein. This was sent to Dr. Carl Christianson at the Max Planc institute with Peter Seeberger for attempts at crystallization. Some crystal hits were achieved, but none were of diffraction quality. Tryptic digest of the modified protein and MS MS analysis confirmed that the trehalose modification occurred exclusively at the catalytic serine 124[30].



Scheme 3.16. Covalent inhibition of Ag85C (a) 20 mM TEA, pH 7.2, 20 min, RT.

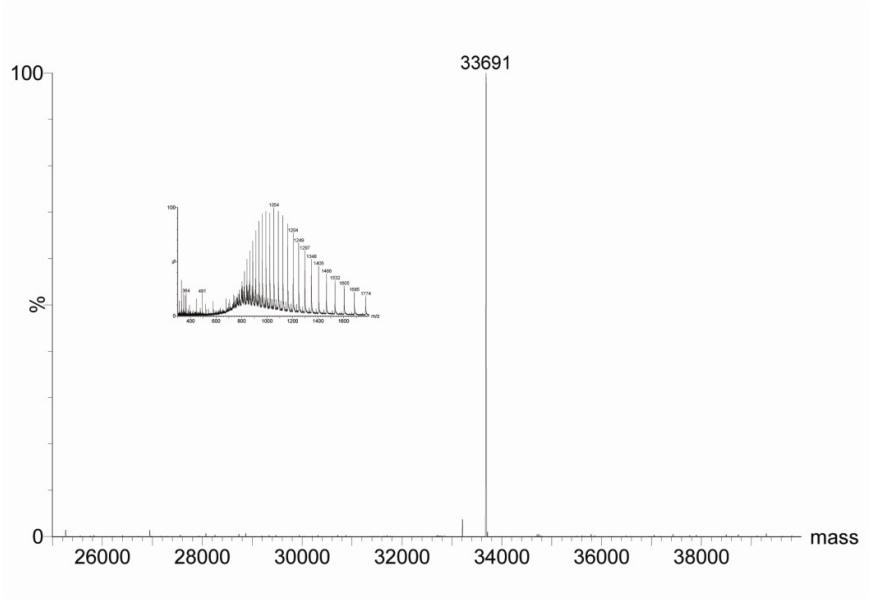
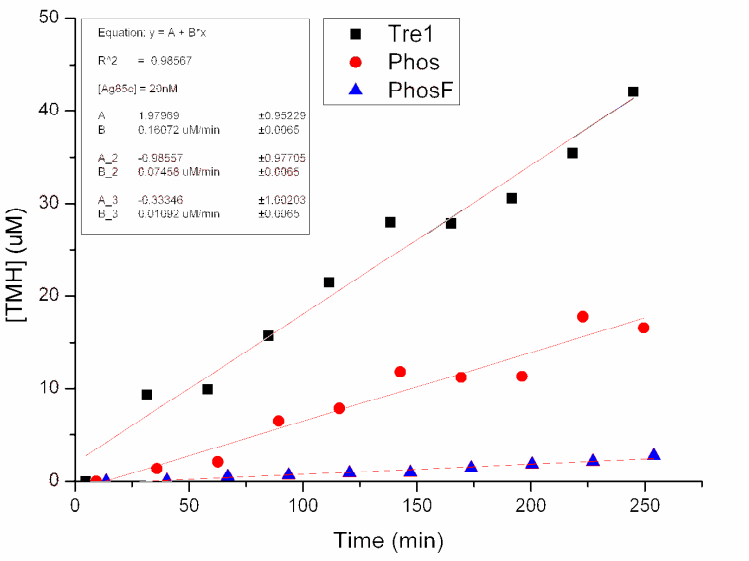


Figure 3.13. Completely modified Ag85C after reaction with 60. Calculated mass Ag85C-60: 33691, Found: 33691;

To investigate the inhibitory potential of **241** and **60** further, a percentage-inhibition assay of Ag85C was conducted (**Figure 3.14**). A saturating concentration of trehalose (> 25× K_M) was employed to ensure maximal turnover rate. Under these conditions, it was found that phosphonate **241** resulted in inhibition of Ag85C to just 46% of the maximal rate. In contrast, fluorophosphonate **60** almost entirely abolished enzyme activity, with just 7% residual activity detected. These results further substantiate that compound 60 is a covalent, irreversible inhibitor of the Ag85 enzymes, in which the phosphoryl fluoride reacts to covalently modify the active site serine, abolishing enzyme activity.

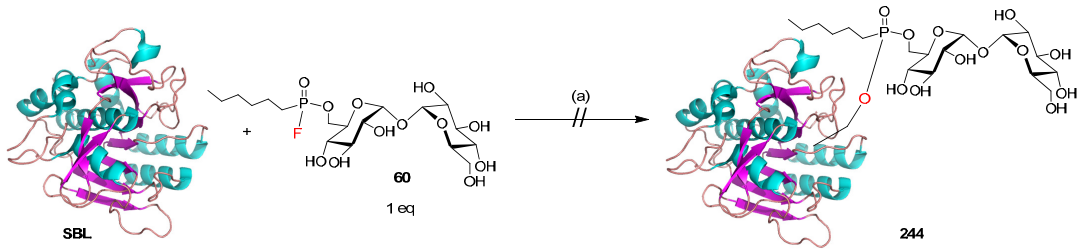
In whole cell MIC assays, trehalose fluorophosphonate and phosphonate **60** and **242** have comparable MICs, with only partial growth inhibition at 200 ug/mL. These results may be explained by degradation of the labile fluorophosphonate in mycobacterial growth medium. Future work with this compound could consider a dosage schedule to sustain the concentration of inhibitor over the length of the growth inhibition assay.



Inhibitor	Initial Rate (uM/min)	% Activity	MIC
none	0.16	100%	NA
Tre Phos **241**	0.075	46%	Partial at 200ug/mL
Tre PhosF **60**	0.011	7%	Partial at 200ug/mL

Figure 3.14. Inhibition of Ag85C with trehalose phosphonate 241 and fluorophosphonate 60 shows nearly complete inhibition of Ag85C by 60 and only 50% abolition of catalytic activity by 241.

With inhibitory potential confirmed, the selectivity of **60** for Ag85 was investigated. Due to the highly reactive and, potentially, non-selective nature of some phosphoryl fluorides (ex: O-isopropyl methylphosphonyl fluoridate or Sarin gas), compound **60** was screened against several other proteases and esterases and confirmed that it is selective for the Ag85 enzymes. Under identical reaction conditions to those employed with Ag85C, subtilisin *Bacillus lentus* (SBL) protease and *Arctica candida* B lipase were incubated with **60** and then analysed by mass spectrometry. No modification of either SBL or *Arctica candida* B was detected, suggesting that **60** offers selectivity for the Ag85 enzymes over other, related enzyme classes (**Scheme 3.17**, **Scheme 3.18**, **Figure 3.15** and **Figure 3.16**). The development of inhibitor **60** may allow the efficient and selective *in vitro* knock-down of Ag85 activity in short term experiments or when dosed to compensate for hydrolysis.



Scheme 3.17. SBL shows no reactivity towards fluorophosphonate 60. (a) 20 mM TEA, pH 7.2, RT, 20 min.

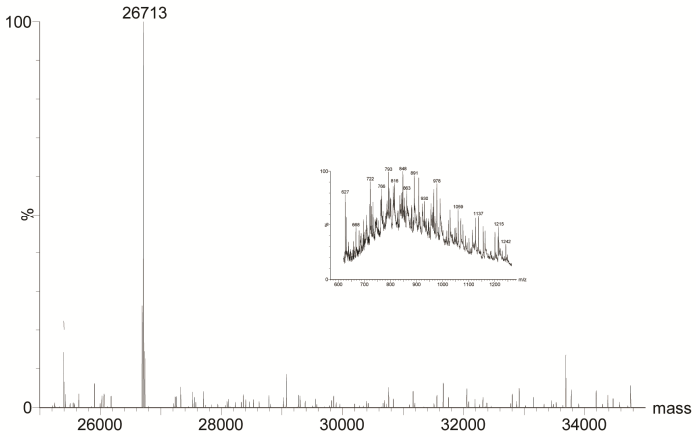
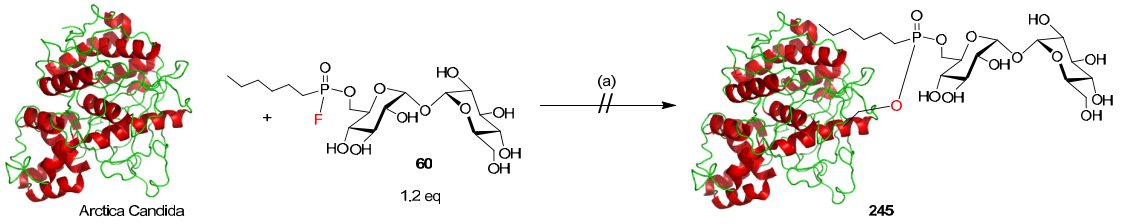


Figure 3.15. Completely unreacted SBL after reaction with fluorophosphonate 60. Calculated mass unreacted-SBL: 26714, Found: 26713;



Scheme 3.18. Completely unreacted Arctica Candida after reaction with fluorophosphonate 60 (a) 20 mM TEA, pH 7.2, 20 min.

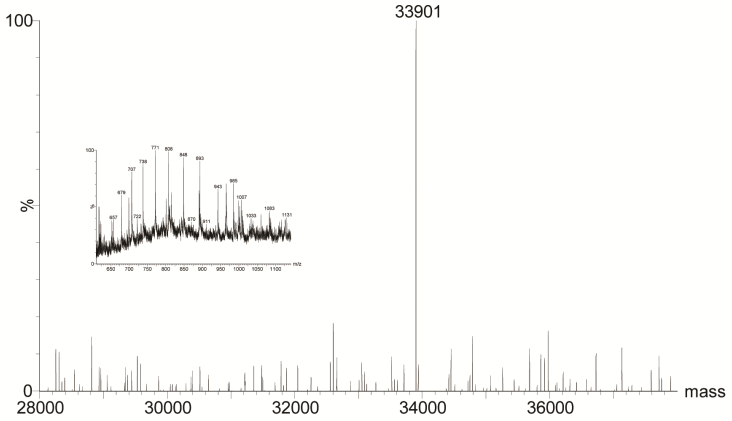


Figure 3.16. Completely unreacted Arctica Candida lipase after reaction with fluorophosphonate 60. Calculated mass unreacted-SBL: 33902, Found: 33901;

3.6 Conclusions

In this chapter, Ag85-based assay has been developed and used to determine the scope of the Ag85 enzymes tolerance towards dissacharide and monosaccharide substrates. This assay revealed a broad substrate tolerance of the Ag85 enzymes and has informed the design of an Ag85-selective inhibitor. A variety of trehalose analogues, modified with probe-like handles (biotin, fluorescein, quantum dots) were found to be substrates of the Ag85 enzymes, suggesting that these compounds may be employed to label mycobacteria within cells and animals harboring tuberculosis. We have demonstrated that these compounds are not processed by host trehalases and have minimal inhibitory activity towards *Mtb.* In the upcoming chapters (**Chapters 4** and **5**), we will discuss the progress towards implementation of these compounds *in vivo.*

3.7 References

1. Chatterjee, D. The mycobacterial cell wall: structure, biosynthesis and sites of drug action. *Curr. Opin. Chem. Biol.* **1**, 579-588 (1997).
2. Belisle, J.T. et al. Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. *Science* **276**, 1420-2 (1997).
3. Sathyamoorthy, N. & Takayama, K. Purification and characterization of a novel mycolic acid exchange enzyme from *Mycobacterium smegmatis*. *J. Biol. Chem.* **262**, 13417-23 (1987).
4. Boucau, J., Sanki, A.K., Voss, B.J., Sucheck, S.J. & Ronning, D.R. A coupled assay measuring *Mycobacterium tuberculosis* antigen 85C enzymatic activity. *Anal. Biochem.* **385**, 120-127 (2009).
5. Elamin, A.A., Stehr, M., Oehlmann, W. & Singh, M. The mycolyltransferase 85A, a putative drug target of *Mycobacterium tuberculosis*: development of a novel assay and quantification of glycolipid-status of the mycobacterial cell wall. *J. Microbiol. Methods* **79**, 358-63 (2009).

6. Lucas, M. TB Vaccine Testing and Research Materials Vol. 2011.
7. Geoghegan, K.F. et al. Spontaneous α-N-6-Phosphogluconoylation of a "His Tag" inEscherichia coli:The Cause of Extra Mass of 258 or 178 Da in Fusion Proteins. *Anal. Biochem.* **267**, 169-184 (1999).
8. Gobec, S. et al. Design, synthesis, biochemical evaluation and antimycobacterial action of phosphonate inhibitors of antigen 85C, a crucial enzyme involved in biosynthesis of the mycobacterial cell wall. *Eur. J. Med. Chem.* **42**, 54-63 (2007).
9. Gobec, S. et al. Phosphonate inhibitors of antigen 85C, a crucial enzyme involved in the biosynthesis of the *Mycobacterium tuberculosis* cell wall. *Bioorg. Med. Chem. Lett.* **14**, 3559-62 (2004).
10. Antonczak, A.K., Simova, Z. & Tippmann, E.M. A Critical Examination of *Escherichia coli* Esterase Activity. *J. Biol. Chem.* **284**, 28795-28800 (2009).
11. Jelesarov, I. & Bosshard, H.R. Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J. Mol. Recognit.* **12**, 3-18 (1999).
12. Schmidt, M. & Bornscheuer, U.T. High-throughput assays for lipases and esterases. *Biomol. Eng.* **22**, 51-56 (2005).
13. Stahl, C. et al. MspA provides the main hydrophilic pathway through the cell wall of *Mycobacterium smegmatis*. *Mol. Microbiol.* **57**, 1509-1509 (2005).
14. Flint, J. et al. Structural dissection and high-throughput screening of mannosylglycerate synthase. *Nat. Struct. Mol. Biol.* **12**, 608-14 (2005).
15. Yang, M. et al. Probing the breadth of macrolide glycosyltransferases: in vitro remodeling of a polyketide antibiotic creates active bacterial uptake and enhances potency. *J. Am. Chem. Soc.* **127**, 9336-7 (2005).
16. Lee, K.S. et al. Diacyltrehalose of *Mycobacterium tuberculosis* inhibits lipopolysaccharide- and mycobacteria-induced proinflammatory cytokine production in human monocytic cells. *FEMS Microbiol. Lett.* **267**, 121-8 (2007).
17. Yang, M., Brazier, M., Edwards, R. & Davis, B.G. High-throughput mass-spectroscopy monitoring for multisubstrate enzymes: Determining the kinetic parameters and catalytic activities of glycosyltransferases. *ChemBioChem* **6**, 346-357 (2005).
18. Fujita, Y., Naka, T., McNeil, M.R. & Yano, I. Intact molecular characterization of cord factor (trehalose 6,6'-dimycolate) from nine species of mycobacteria by MALDI-TOF mass spectrometry. *Microbiology* **151**, 3403-3416 (2005).
19. Baer, H.H., Breton, R.L. & Shen, Y. Synthesis of a trehalose homolog, 6-deoxy-α-gluco-heptopyranosyl 6-deoxy-α-gluco-heptopyranoside, and the corresponding bis(heptosiduronic acid). *Carbohydr. Res.* **200**, 377-389 (1990).
20. Lin, F.L., van Halbeek, H. & Bertozzi, C.R. Synthesis of mono- and dideoxygenated α,α-trehalose analogs. *Carbohydr. Res.* **342**, 2014-2030 (2007).
21. Patel, M.K. & Davis, B.G. Flow chemistry kinetic studies reveal reaction conditions for ready access to unsymmetrical trehalose analogues. *Org. Biomol. Chem.* **8**, 4232-4235.
22. Cleland, W.W. & Begley, T.P. Enzyme Kinetics. in *Wiley Encyclopedia of Chemical Biology* (John Wiley & Sons, Inc., 2007).
23. Anderson, D.H., Harth, G., Horwitz, M.A. & Eisenberg, D. An interfacial mechanism and a class of inhibitors inferred from two crystal structures of the *Mycobacterium tuberculosis* 30 kDa major secretory protein (Antigen 85B), a mycolyl transferase. *J. Mol. Biol.* **307**, 671-81 (2001).
24. Ronning, D.R. et al. Crystal structure of the secreted form of antigen 85C reveals potential targets for mycobacterial drugs and vaccines. *Nat. Struct. Biol.* **7**, 141-6 (2000).
25. Cook, P.F. et al. Substitution of pyridoxal 5'-phosphate in the O-acetylserine sulfhydrylase from Salmonella typhimurium by cofactor analogs provides a test of the mechanism proposed for formation of the alpha-aminoacrylate intermediate. *J. Biol. Chem.* **271**, 25842-25849 (1996).
26. Brown, A.J. & Snyder, F. Alkyldihydroxyacetone-P Synthase - Solubilization, Partial-Purification, New Assay-Method, and Evidence for a Ping-Pong Mechanism. *J. Biol. Chem.* **257**, 8835-8839 (1982).
27. Casals, C., Acebal, C. & Arche, R. Substrate Selectivity of Lysophosphatidylcholine - Lysophosphatidylcholine Acyltransferase from Rabbit Lung. *Int. J. Biochem.* **16**, 773-778 (1984).
28. Yang, M., Davies, G.J. & Davis, B.G. A Glycosynthase Catalyst for the Synthesis of Flavonoid Glycosides. *Angew. Chem. Int. Ed. Engl.* **46**, 3885-3888 (2007).
29. Cook, P.F. & Cleland, W.W. *Enzyme kinetics and mechanism*, xxii, 404 p. (Garland Science, London ; Abingdon, 2007).
30. Barry, C.S., Backus, K.M., Barry, C.E. & Davis, B.G. ESI-MS Assay of *M. tuberculosis* Cell Wall Antigen 85 Enzymes Permits Substrate Profiling and Design of a Mechanism-Based Inhibitor. *J. Am. Chem. Soc.*, ASAP (2011).

31. Sanki, A.K., Boucau, J., Ronning, D.R. & Sucheck, S.J. Antigen 85C-mediated acyl-transfer between synthetic acyl donors and fragments of the arabinan. *Glycoconj. J.* (2008).
32. Sanki, A.K. et al. Synthesis of methyl 5-S-alkyl-5-thio-D-arabinofuranosides and evaluation of their antimycobacterial activity. *Bioorg. Med. Chem.* **16**, 5672-82 (2008).
33. Matsunaga, I. et al. Mycolyltransferase-mediated Glycolipid Exchange in Mycobacteria. *J. Biol. Chem.* **283**, 28835-28841 (2008).
34. Jackson, M. et al. 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 (1999).
35. Sato, S., Okamoto, K., Minami, R., Kohri, H. & Yamamoto, S. Trehalose Can Be Used as a Parenteral Saccharide Source in Rabbits. *J. Nutr.* **129**, 158-164 (1999).
36. Carroll, J.D., Pastuszak, I., Edavana, V.K., Pan, Y.T. & Elbein, A.D. A novel trehalase from *Mycobacterium smegmatis* - purification, properties, requirements. *FEBS J.* **274**, 1701-14 (2007).
37. Gibson, R.P. et al. Molecular basis for trehalase inhibition revealed by the structure of trehalase in complex with potent inhibitors. *Angew. Chem. Int. Ed. Engl.* **46**, 4115-9 (2007).
38. Withers, S.G., MacLennan, D.J. & Street, I.P. The synthesis and hydrolysis of a series of deoxyfluoro-glucopyranosyl phosphates. *Carbohydr. Res.* **154**, 127-144 (1986).
39. Street, I.P., Kempton, J.B. & Withers, S.G. Inactivation of a β-glucosidase through the accumulation of a stable 2-deoxy-2-fluoro-α-D-glucopyranosyl-enzyme intermediate: a detailed investigation. *Biochemistry (Mosc.)* **31**, 9970-9978 (1992).
40. James, G.T. Inactivation of the protease inhibitor phenylmethylsulfonyl fluoride in buffers. *Anal. Biochem.* **86**, 574-579 (1978).
41. Neet, K.E. & Koshland, D.E., Jr. The conversion of serine at the active site of subtilisin to cysteine: a "chemical mutation". *Proc. Natl. Acad. Sci. U. S. A.* **56**, 1606-11 (1966).
42. Ronning, D.R., Vissa, V., Besra, G.S., Belisle, J.T. & Sacchettini, J.C. *Mycobacterium tuberculosis* antigen 85A and 85C structures confirm binding orientation and conserved substrate specificity. *J. Biol. Chem.* **279**, 36771-7 (2004).
43. Sanki, A.K., Boucau, J., Umesiri, F.E., Ronning, D.R. & Sucheck, S.J. Design, synthesis and biological evaluation of sugar-derived esters, α-ketoesters and α-ketoamides as inhibitors for *Mycobacterium tuberculosis* antigen 85C. *Mol. BioSyst.* **5**, 945-956 (2009).
44. Wilson, R.A., Maughan, W.N., Kremer, L., Besra, G.S. & Futterer, K. The structure of *Mycobacterium tuberculosis* MPT51 (FbpC1) defines a new family of non-catalytic alpha/beta hydrolases. *J. Mol. Biol.* **335**, 519-30 (2004).
45. Werder, M., Hauser, H. & Carreira, E.M. Carbohydrate Sulfonyl Chlorides for Simple, Convenient Access to Glycoconjugates. *Org. Lett.* **7**, 1145-1148 (2005).
46. Tsukiji, S., Miyagawa, M., Takaoka, Y., Tamura, T. & Hamachi, I. Ligand-directed tosyl chemistry for protein labeling in vivo. *Nat. Chem. Biol.* **5**, 341-343 (2009).
47. Cravatt, B.F., Wright, A.T. & Kozarich, J.W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **77**, 383-414 (2008).
48. Kidd, D., Liu, Y. & Cravatt, B.F. Profiling Serine Hydrolase Activities in Complex Proteomes. *Biochemistry (Mosc.)* **40**, 4005-4015 (2001).
49. Liu, Y., Patricelli, M.P. & Cravatt, B.F. Activity-based protein profiling: The serine hydrolases. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 14694-14699 (1999).
50. Bachovchin, D.A. et al. Superfamily-wide portrait of serine hydrolase inhibition achieved by library-versus-library screening. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 20941-20946 (2010).
51. Björkling, F., Dahl, A., Patkar, S. & Zundel, M. Inhibition of lipases by phosphonates. *Bioorg. Med. Chem.* **2**, 697-705 (1994).
52. Fischer, G. Trends in protease inhibition. *Nat. Prod. Rep.* **5**, 465-495 (1988).

4 Chapter 4: Labeling mycobacteria *in vitro*

4.1 Introduction

In Chapter 3 a novel Antigen 85 (Ag85) enzyme assay revealed that the antigen 85 enzymes will tolerate a wide variety of modifications to the trehalose scaffold. These results support an Ag85 mediated labeling of whole mycobacteria, but it remains to be seen whether trehalose and analogues will label *Mtb in vivo.* It has been well documented that drug, and presumably probe, uptake by mycobacteria can be confounded by numerous variables in addition to enzyme activity (some of which include permeability, growth inhibition, stability and off target effects)[1]. Penetration of drugs into the lung is known to be poor across many drug classes and subtherapeutic levels of antibiotic at the site of the mycobacterial infection is thought to be one of the causes of the emergence of resistant strains of *Mtb*[2].

To fully establish that trehalose could label whole cell mycobacteria via the Ag85 enzymes, an evaluation of the trehalose incorporated into mycobacteria was required. Chapter 4 demonstrates that trehalose and the analogues described in Chapter 2 are readily processed by the Ag85 enzymes and taken up into the *Mtb* cell wall in batch cultures, providing radioactive and fluorescent labeling of bacteria. These results support the hypothesis that trehalose can serve as a novel small molecule, carbohydrate probe to label mycobacteria and will provide the foundation of our work describing trehalose uptake *in vivo*, detailed in Chapter 5.

4.2 Results

4.2.1 Trehalose uptake by whole cells of *Mtb*

To understand the processing of trehalose and trehalose analogues by tuberculosis, *in vitro* uptake assays of trehalose into whole cell bacteria were employed to demonstrate that trehalose is taken up by the bacteria. Genomic studies of *Mtb* sugar transport suggested a precedent that extracellular trehalose could be taken up by mycobacteria[3,4]. Using comparative radioprobes ^{14}C-glycerol, ^{14}C-trehalose, ^{14}C-galactose, ^{14}C-arabinose and ^{14}C-glucose, trehalose uptake in *Mtb* culture was observed over 2 and 24 hour time periods (Figure 4.1 a-c)[5]. In contrast, galactose and arabinose showed no time-dependent uptake (Figure 4.1b). Modest glucose uptake was observed (**Figure 4.1c**), presumably because the culture medium (7H9 supplemented with ADC) contains 10mM unlabeled glucose.

Both glycerol and glucose are well documented to have mycobacterial uptake[4-6]. As sugars found throughout the eukaryotic and prokaryotic glycome, these were not of particular interest for selective mycobacterial probe development and instead served as positive controls. Galactose and arabinose are known components of the arabinogalactan[7] and were considered alternative mycobacterial labeling compounds if trehalose failed to label bacteria. Galactose appeared to either be taken up rapidly before the first sampling, or the galactose may have adhered to the membrane used in the uptake experiment. Arabinose had low levels of uptake. These results suggest that trehalose is readily taken up by mycobacteria and is a potential target for selectively labeling tuberculosis *in vivo*.

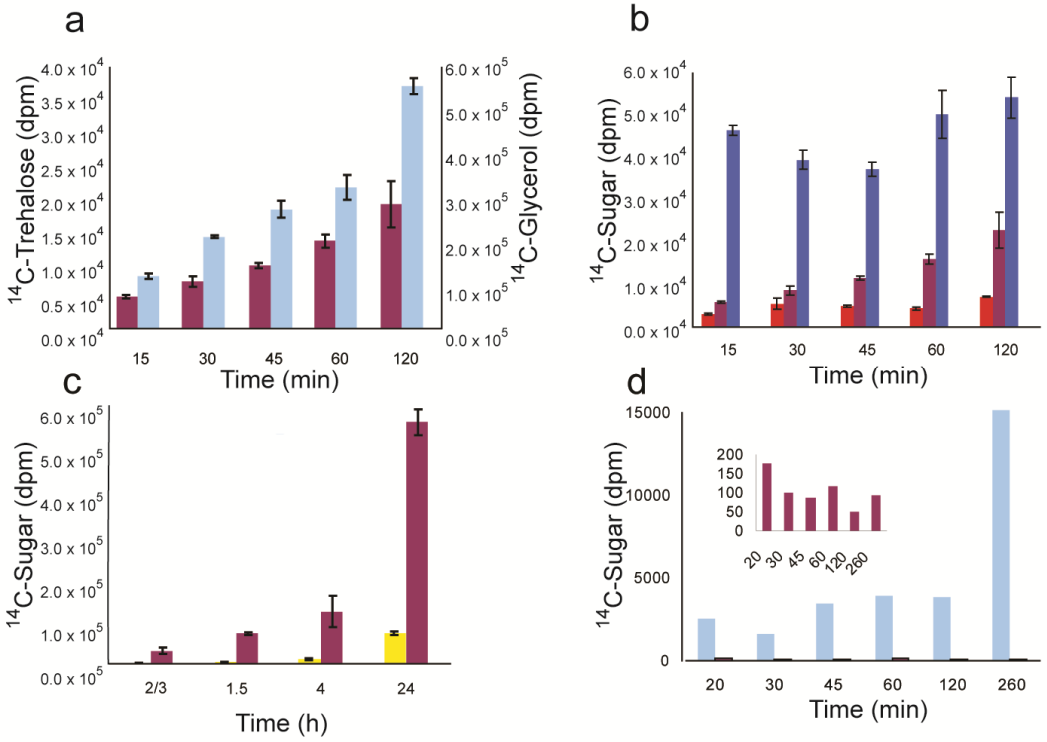


Figure 4.1. Trehalose is incorporated into *Mtb in vitro* (a) ^{14}C-trehalose is taken up by *Mtb* culture over 2 hours, at a rate that is only slightly slower than glycerol. purple-^{14}C-trehalose, blue-^{14}C–glycerol. (b) In contrast ^{14}C-arabinose and ^{14}C-galactose show no time dependent uptake into mycobacteria. Dark blue-galactose, purple-trehalose, red-arabinose. Trehalose was used at a tenfold lower concentration of radioactivity relative to the other sugars (0.1 µCi/mL), rather than (1 µCi/mL). (c) Uptake of trehalose continues over a 24 hour period. purple-^{14}C-trehalose, yellow-^{14}C–glucose. (d) An attempt to evaluate the uptake of putative, synthetic ^{14}C methyl galacto-gluco-trehalose analogue 170 shows no uptake of the compound synthesized. Blue-glycerol, purple (inset)-trehalose analogue 170. DPM-Disintegrations per minute.

To confirm that the trehalose analogues synthesized (**Chapter 2**) would label mycobacteria in comparable fashion to unmodified trehalose, the synthesis of ^{14}C versions of the trehalose derivatives was attempted. The original synthesis involved the introduction of a ^{14}C label at

the anomeric position, utilizing ^{14}C-methyl iodide (described in **Chapter 2**). This synthesis failed, due presumably, to the difficulty of such small scale lithium halogen exchange reactions and/or the questionable purity of the ^{14}C-methyl iodide. In test reactions, synthesizing ^{13}C-methyl-glucose, the optimized reaction conditions afforded good yield of the target compound.

In an alternative synthetic strategy non-radioactive, 'cold,' galactose was doped with a small amount of ^{14}C-galactose, remaining from our uptake assays. At the time, glucose was not available, so galactose had to suffice. From galactose, a synthesis of ^{14}C-methyl-galacto trehalose (reported in **Chapter 2**) was attempted. With access to insufficient chemical characterization, due to the radioactive nature of the compound, this synthesis proved challenging and it was nearly impossible to conclusively characterize the final product by TLC alone. Uptake of the crude product, after deprotection, was assayed in batch *Mtb* culture (**Figure 4.1d**).

Uptake studies of the ^{14}C-methyl-gluco-galacto trehalose analogue **170** were conduct in a manner comparable to previous experiments with the exception that less ^{14}C-glycerol was available for uptake. However, over 4 hours, uptake was observed. In contrast, no perceivable uptake was observed with **170**. It is tempting to conclude that *Mtb* does not tolerate the galacto configuration, several alternative explanations. However, due to the slow nature of the hydrogenation step, the product may not have been completely deprotected and remaining benzyl groups could block any incorporation.

This preliminary foray in radiochemistry revealed the challenge of synthesizing ^{14}C-trehalose derivatives. These difficulties suggested that synthesis of ^{14}C-trehalose analogues may not be the most straightforward approach towards evaluation of trehalose labeling. To circumvent some of these difficulties with the use of ^{14}C, fluorescent trehalose (FITC-Tre, synthesis described in **Chapter 2**) was explored as an alternative trehalose probe.

4.2.2 ^{14}C-trehalose labels TDM and TMM of mycobacteria

Prior to investigating the FITC-Tre labeling, incorporation of trehalose is into TDM and TMM was quantified, disproving a model that proposed non-specific sugar uptake and metabolism by endogenous trehalases. While this research was concluding, a trehalose transporter was reported in the literature[8]. Kalscheuer *et al* suggest that trehalose, released by the transmycolation reactions is trapped by the transporter and recycled back into the

mycobacteria. The implications of this publication were that some of the trehalose uptake that we observed (**Figure 4.1**) was not due to direct incorporation in the TDM and TMM, but rather due to uptake by the trehalose transporter.

The selective accumulation uptake of ^{14}C-Tre into TDM and TMM was confirmed by autoradiographic lipid analysis (**Figure 4.2**)[9]. Trehalose uptake was significant, and the strongest band labeled by ^{14}C-trehalose comigrated with a band labeled by ^{14}C-acetate and comigrated with cold TMM (not shown), which was visualized by staining in 5% sulfuric acid [**Figure 4.2** (lane ii)]. ^{14}C-trehalose also labeled a band that comigrated with TDM, although not labeled to the same extent as TMM. To confirm that these bands were TMM and TDM, cells were also labeled after treatment with isoniazid (INH), which inhibits mycolic acid biosynthesis and therefore blocks production of both TMM and TDM[10]. INH effectively blocked biosynthesis of these two bands in a concentration dependent manner [see **Figure 4.2** (lane iii,iv)]. Together these results demonstrated that trehalose is taken up by the bacteria and incorporated into TDM and TMM by *M. tuberculosis in vitro*.

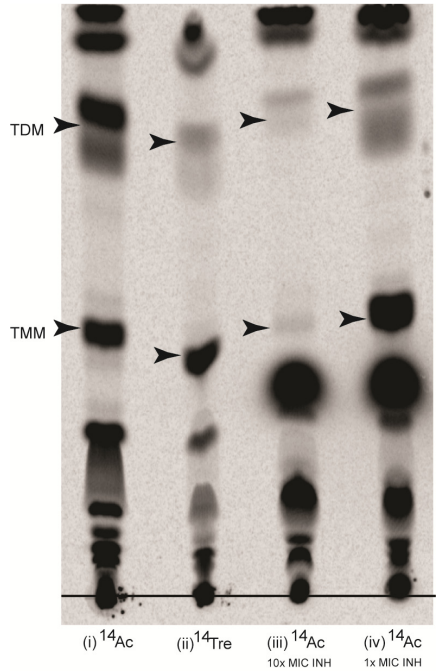


Figure 4.2. ^{14}C-trehalose is incorporated into TMM and TDM in *Mtb*. Arrowheads show the migration position of authentic samples of TMM and TDM. In (i) the culture was labeled for 24 hr with ^{14}C-acetate to label fatty acids. In (ii) the culture was labeled for 24 h with ^{14}C-trehalose. In (iii) culture was treated with isoniazid (INH) at 10X MIC to inhibit formation of TMM and TDM. (iv) 1X MIC treatment with isoniazid (INH) shows concentration dependent, partial depletion of TDM and TMM. ^{14}C-Trehalose labels spots that co-migrate with TMM and TDM.

4.2.3 FITC-trehalose specifically labels *M. tuberculosis* in batch culture

Having established that ^{14}C-trehalose labels mycobacteria and is incorporated into the TDM and TMM, we turned our attention to the mycobacterial uptake of other trehalose derivatives. For the purpose of developing a trehalose-based probe to label mycobacteria, it was necessary to ascertain that not simply native trehalose, but also analogues were readily taken up by mycobacteria. Because the initial attempts at synthesizing ^{14}C-analogues of the library of trehalose derivatives proved futile (**Chapter 2**), a fluorescent approach (FITC-Tre) was selected as an alternate strategy for the evaluation of trehalose uptake.

Whole-cell uptake assays and microscopy revealed that FITC-Tre labels *Mtb* in a time dependent, Ag85 dependent and mycobacteria selective fashion. *Mtb* strain H37Rv was grown in the presence of FITC-trehalose and at various time points aliquots were removed and washed to reduce non-specific binding. A significant increase in fluorescence over time was observed in bacteria exposed to FITC-trehalose (**Figure 4.3a**). To demonstrate that this incorporation was dependent upon Ag85, we obtained a mutant in Ag85C that has been previously reported to have 40% less mycolic acid incorporation into the mycobacterial cell wall[11]. This mutant incorporated approximately 30% less FITC-trehalose than did wild type (**Figure 4.3a**), supporting the proposed mechanism of anchoring of FITC-trehalose through Ag85-mediated incorporation of mycolic acids.

To demonstrate that incorporation was specific, uptake into live *M. tuberculosis* was compared with the autofluorescence of unlabeled control bacteria and to the FITC-Tre labeling of heat killed bacteria. Significantly greater uptake into the FITC-Tre treated live bacteria was observed relative to controls [**Figure 4.3b** (lanes i, ii and iv)]. Heat-killed *M. tuberculosis* form large aggregates that are particularly difficult to wash adequately, perhaps contributing to the relatively high background observed in these organisms (lane ii). Compared with FITC-Tre, FITC-glucose[i] was found to label mycobacteria with much reduced efficiency [**Figure 4.3b** (lane iii)], consistent with the poor efficiency of glucose as a substrate for the Ag85 enzymes (and the unlikely uptake of this probe into the cell by glucose-specific transporters). Finally, to establish specificity of labeling, three other organisms commonly found in the human lung were exposed to FITC-trehalose: *Staphylococcus aureus*, *Haemophilus influenzae* and *Pseudomonas aeruginosa*. None of

[i] FITC-glucose was synthesized by Dr. Francois D'Hooge

these organisms were found to show appreciable labeling with FITC-trehalose [**Figure 4.3b** (lane v-viii)]. These results demonstrated that FITC-Tre is a selective means to label mycobacteria.

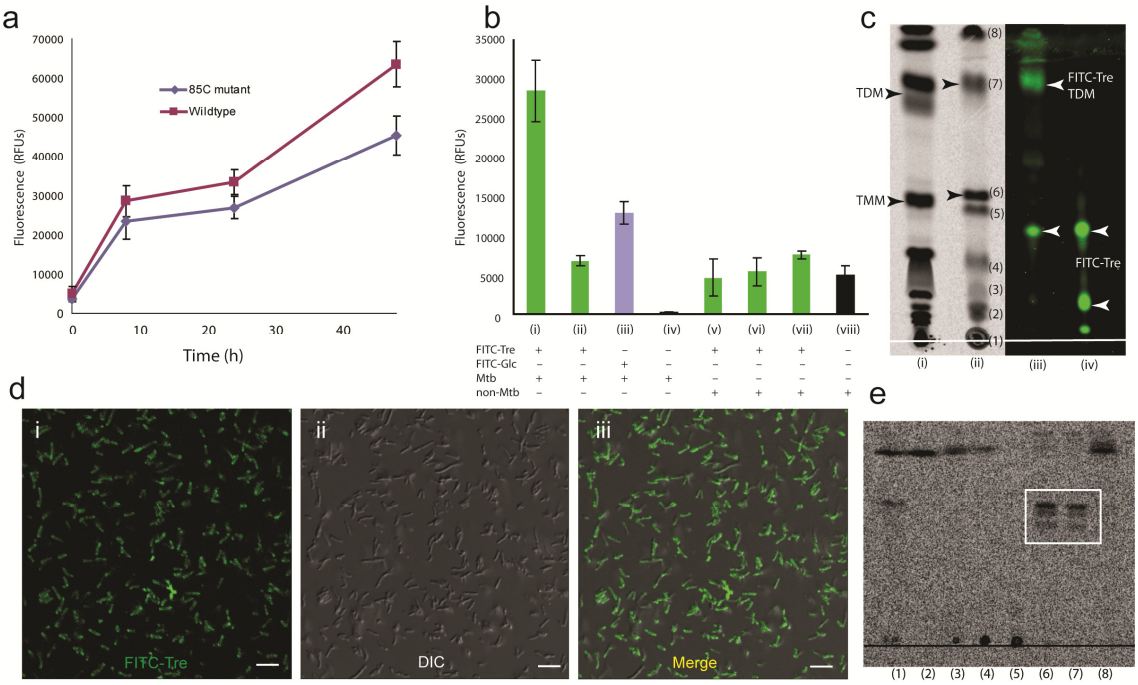


Figure 4.3. FITC-Tre is incorporated by *Mtb* into FITC-TDM through the action of the Ag85 enzymes. (a) Uptake of FITC-Tre into wild-type CDC 1551 (purple squares) or *ag85C* mutant, known to have a 40% reduction in mycolic acid incorporation into the cell wall (blue diamonds). (b) Uptake of FITC-Tre or FITC-Glc into (i , iii, and iv) live *Mtb,* (ii) heat killed *Mtb,* (v) *Staphylococcus aureus,* (vi) *Pseudomonas aeruginosa,* (vii) *Haemophilus influenza,* (viii) mean autofluorescence of *H. influenzae as* representative of autofluorescence of non-mycobacterial species. (c) FITC-Tre and ^{14}C-acetate colabel glycolipids; (i) the radioTLC from ^{14}C-acetate labeled *Mtb.* (ii) the radioTLC from dual labeling with FITC-Tre and ^{14}C-acetate. (iii) the fluorescence intensity in lane "ii" (iv) untreated FITC-Tre in the same TLC system. Multiple spots are the charged, uncharged and cyclized versus uncyclized species of fluorescein (arrowheads) (d) FITC-Tre labeled *Mtb.* (i) FITC-Tre, (ii) Transmitted Light Differential Interference Contrast (DIC) image, (iii) Merge of 'i-ii,' Scale bars 5 µm. (e) The glycolipids extracted from "C-ii" that co-migrate with TDM and TMM contain mycolic acids. (1-8) correspond to individual spots from lane "ii" of "c" that were excised, saponified and run on a second TLC. The glycolipids in spots 6 and 7 contain lipids that co-migrate with authentic mycolic acids and give the characteristic banding pattern of the three major mycolates. Error bars reflect standard deviation (SD) for four experiments. Scale bars 5 µm.

To establish that FITC-trehalose was localizing to cells through esterification with mycolic acids, we simultaneously labeled cells with both FITC-trehalose and ^{14}C-acetate and isolated individual fluorescent spots by preparative TLC [**Figure 4.3c** (lanes ii, iii)]. Fluorescent label co migrated with the spot identified in **Figure 4.2** as TDM. Saponification of all radiolabeled spots [**Figure 4.3c** (lane ii, 1-8)] and subsequent methylation and TLC (**Figure 4.3e),** revealed that the strongly fluorescent spot carrying FITC-trehalose (spot 7) was associated

with lipids comigrating with the characteristic triplet pattern of the three classes of mycolic acids (alpha, methoxy and keto) (**Figure 4.3e**).No further attempts were made to characterize the other less polar spots of low abundance that also apparently incorporated FITC-trehalose, although these might reasonably be supposed to be other trehalose-containing glycolipids found within the mycobacterial cell wall[12,13].

TLC analysis of glycolipid extracts confirmed that FITC-Tre was esterified into the TDM. It was not certain why accumulation of the probe was not observed in the TMM (spot 6). We and others have previously observed that TMM was less stable than TDM and tended to degrade over time (unpublished). *Mtb* contain less TMM than TDM and it is thought that the TMM is a more transient component of the outermembrane[14]. It was possible that the FITC-TMM was present at a level too low to be observed by TLC or that the product was unstable or not soluble in the extraction protocol utilized. Alternatively, the absence of a FITC-TMM could reflect an unknown aspect of the biosynthesis of TMM, TDM and the arabinogalactan linked mycolates.

4.2.4 Polar labeling of Mtb microscopy

The value of FITC-trehalose as a probe was demonstrated by the clear patterns of differential accumulation within live *M. tuberculosis* (**Figure 4.4**). Confocal fluorescence microscopy of H37Rv–*M. tuberculosis* expressing a red fluorescent protein (RFP, here mCherry, expressed by a pMV261 plasmid) revealed higher levels of green fluorescence from FITC-trehalose incorporation at outer membranes and, in particular, at the bacterial poles, and less in mid sections (where the RFP localizes, **Figure 4.4b**).

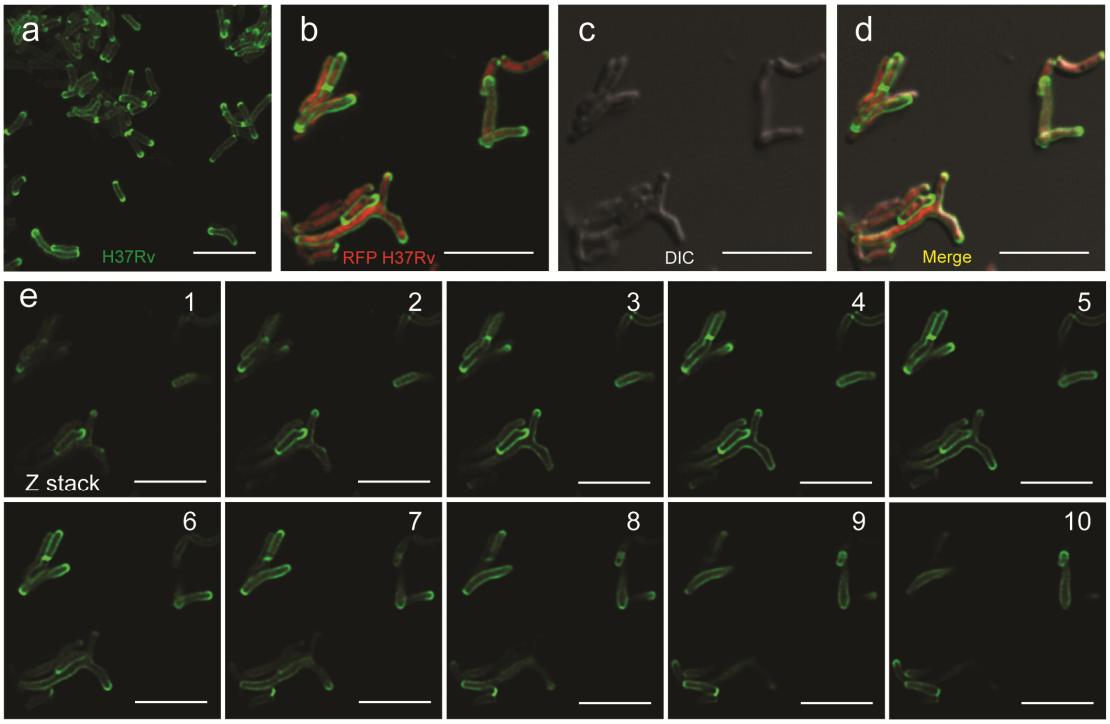


Figure 4.4. Polar labeling by FITC-Tre *in vitro* (a) shows FITC-Tre labeling of H37Rv (b) shows FITC-Tre labeling of the RFP expressing strain of H37Rv and (c) is the DIC image of (b), while (d) is an overlay of (b-c) and (d) depicts the Z stack (1-10) through the same image (step size of 0.13 µm) and demonstrates the FITC-Tre localization to the poles and membrane. Scale bars 5µm.

This specific polar localization was confirmed by microscopic sections along the Z-axis (Z-stacks, **Figure 4.4e**) and by statistical analysis of mean fluorescence ($P < 0.0001$**, Figure 4.5**). For the statistical analysis ROIs for the poles and midsections of the bacteria were generated using the marching cubes function in Imaris Bitplane image processing software (**Figure 4.5 b,c**). These observations were not only consistent with the mode of action proposed for these trehalose analogues but also suggested higher Ag85 activities at the poles that are consistent with a polar growth model.

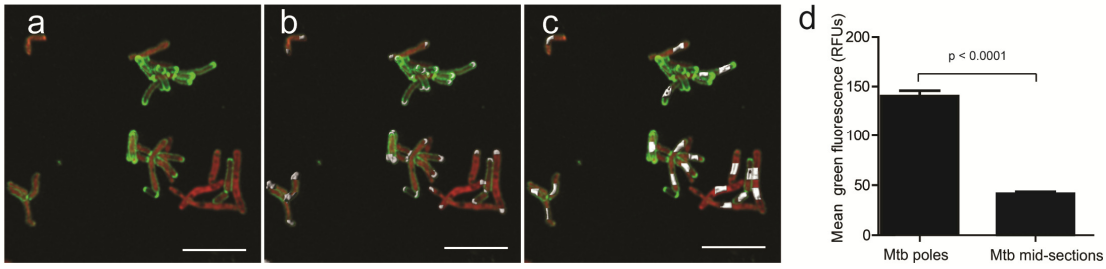


Figure 4.5. The quantification of the intensity of the poles and mid-sections of H37Rv *Mtb* (a) shows a representative image of H37Rv *Mtb,* labeled with FITC-Tre (green) and expressing RFP (red). (b) Indicates the regions of interest (ROIs) at the mycobacterial poles in white, while (c) shows the ROIs of the bacterial mid-sections in white. (d) Quantification of the relative fluorescence at the poles versus midsection of *Mtb* reveals a statistically significant accumulation of probe at the poles. Scale bars 5 µm.

Dosage with trehalose *in vitro* was found to modulate FITC-Tre labeling in bacterial membranes, substantiating that the probe accumulation was Tre-dependent. We found that high concentrations of trehalose abolished FITC-Tre labeling. Intermediate concentrations, show somewhat decreased accumulation of label at the poles (not shown). Notably, there is some background stickiness of the compound and slight labeling is still observed during treatment with high trehalose concentrations. However, the high trehalose abolishes all accumulation of probe at the mycobacterial poles (**Figure 4.6d**).

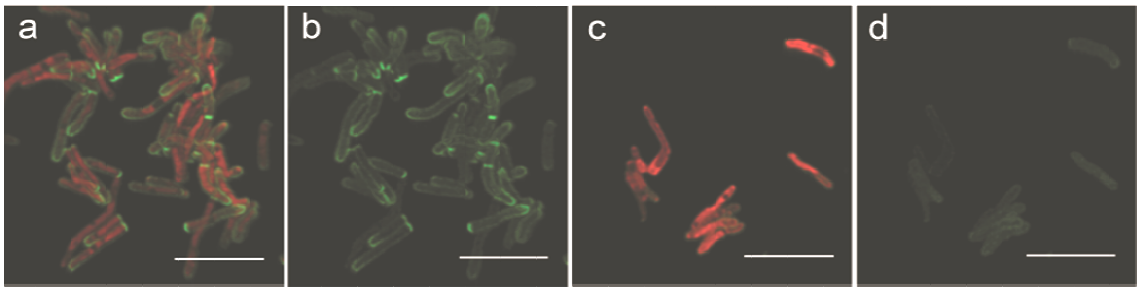


Figure 4.6. Trehalose saturation of FITC-Tre treated H37Rv *Mtb* expressing RFP shows no FITC-Tre labeling *in vitro*. (a,b) Depict FITC-Tre labeled H37Rv *Mtb* expressing RFP. (b) FITC-Tre channel of overlay shown in (a). (c,d) *Mtb* treated as in (a-b), except grown in the presence of saturating concentrations of trehalose. (c) Shows RFP channel, while (d) shows the disappearance of FITC-Tre labeling in the presence of trehalose (100 mM). Scale bars 5 µm.

4.2.5 MIC data

This apparent tolerance of mycobacteria for unnatural Tre analogues was confirmed by the effect of such substrates upon growth of *Mtb*; minimum inhibitory concentration (MIC) analysis was conducted for key compounds. 2,2'-dideoxy-lyxo-Tre **42**, which was not a substrate of any of the three Ag85 isoforms, showed a weak inhibition (MIC 100µg/mL); C-6 modified compounds were processed well by Ag85, and inhibited growth at 100-200 µg/mL. **Chapter 3, Table 3.4** details the MICs for all analogues. The growth inhibition, and Ag85 competitive substrate nature of the C-6 (Tre-C-6) modified compounds suggests that the synthesis of Tre-C-6-MM may have detrimental effects on mycobacterial growth. These inhibitory consequences may occur because the C-6 Tre cannot form TDM and instead shunts mycolates towards Tre-C-6-MM. In contrast, other Tre analogues (modified at C-2, C-4, etc) do not have any MIC, which further suggests that the decrease in TDM and subsequent increase in TMM levels may inhibit mycobacteria growth. An alternative explanation of this inhibition cannot be ruled out. The MIC of C-6 modified trehalose could be unrelated to

Ag85 and these compounds may block other trehalose synthetic pathways, such as the synthesis or degradation of trehalose.

4.3 Conclusions

The uptake of trehalose into mycobacteria has been established and is an efficient, selective method to label mycobacteria *in vitro*. This labeling occurs in the TMM and TDM glycolipids, and is, at least in part, dependent on the Ag85 enzymes. These compounds do not appreciably inhibit mycobacterial growth and are selective for mycobacteria over other Gram positive and negative pathogens. These compounds are attractive probes for future *in vitro* and *in vivo Mtb* labeling.

Consideration of the future directions of this section is merited. Both Chapters 3 and 4 have discussed the attempts at developing ^{14}C-trehalose analogues, containing the methyl group found in many of our trehalose probes. Although, this synthesis was unsuccessful, due to limitations in our synthetic resources for radiochemistry, we believe that the synthesis of additional ^{14}C-trehalose compounds could help to more fully dissect the mycobacterial lipid labeling process.

The lipid TLCs (**Figure 4.3**) reveal that trehalose is incorporated not, simply into TMM and TDM, but also into many other extractable lipid spots. This additional labeling could be due to the documented uptake and subsequent metabolism of trehalose, followed by incorporation into the FasI and FasII systems and, finally, lipidation of the cell wall. Because sugar transporters are selective, the transport of the bulky FITC-Tre seems unlikely and an alternate explanation could be that these spots contain additional trehalose glycolipids. The mycolate saponification experiments suggested that these spots do not contain mycolic acids. There are many other documented, trehalose containing glycolipids, including sulfolipid-1, diacyl and polyacyl trehalose. The biosynthesis of these important natural products has not been fully explored. Additional ^{14}C-trehalose feeding experiments could shed light on the nature of some of these compounds. The biotin-trehalose compound **54**, briefly described in Chapter 2 could prove a useful tool for pulling down trehalose containing glycolipids.

The full structure and growth model of the mycobacterial cell wall is not fully understood. The presence of additional antigen 85 enzymes has been discussed briefly in Chapter 4 and is precedented by *M. smegmatis* and *Corynebacteria*[15]. The putative trehalose monomycolate transporter has not been identified and it remains to be seen exactly how these large,

hydrophobic glycolipids are shuttled to the mycobacterial cell wall, orientated and directed to their appropriate destination. We believe that the process of synthesizing the mycobacterial cell wall is tightly regulated and that the trehalose compounds could help shed light on the precise mechanism by which the outer bilayer of the mycobacterial cell wall is synthesized.

4.4 References

1. Lambert, P.A. Cellular impermeability and uptake of biocides and antibiotics in Gram-positive bacteria and mycobacteria. *J. Appl. Microbiol.* **92**, 46S-54S (2002).
2. Elliott, A.M., Beming, S.E., Iseman, M.D. & Peloquin, C.A. Failure of drug penetration and acquisition of drug resistance in chronic tuberculous empyema. *Tuber. Lung Dis.* **76**, 463-467 (1995).
3. Deol, P. et al. Role of *Mycobacterium tuberculosis* Ser/Thr kinase PknF: implications in glucose transport and cell division. *J. Bacteriol.* **187**, 3415-20 (2005).
4. Titgemeyer, F. et al. A Genomic View of Sugar Transport in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. *J. Bacteriol.* **189**, 5903-5915 (2007).
5. Stahl, C. et al. MspA provides the main hydrophilic pathway through the cell wall of *Mycobacterium smegmatis*. *Mol. Microbiol.* **57**, 1509-1509 (2005).
6. Mailaender, C. et al. The MspA porin promotes growth and increases antibiotic susceptibility of both *Mycobacterium bovis* BCG and *Mycobacterium tuberculosis*. *Microbiology* **150**, 853-864 (2004).
7. Crick, D.C., Mahapatra, S. & Brennan, P.J. Biosynthesis of the arabinogalactan-peptidoglycan complex of *Mycobacterium tuberculosis*. *Glycobiology* **11**, 107R-118R (2001).
8. Kalscheuer, R., Weinrick, B., Veeraraghavan, U., Besra, G.S. & Jacobs, W.R., Jr. Trehalose-recycling ABC transporter LpqY-SugA-SugB-SugC is essential for virulence of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 21761-6 (2010).
9. Slayden, R.A. & Barry, C.E. Analysis of the Lipids of *Mycobacterium tuberculosis*. Vol. 54 229-245 (2001).
10. Slayden, R.A. & Barry, C.E., 3rd. The genetics and biochemistry of isoniazid resistance in mycobacterium tuberculosis. *Microbes Infect.* **2**, 659-69 (2000).
11. Jackson, M. et al. 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 (1999).
12. Hatzios, S.K. et al. PapA3 Is an Acyltransferase Required for Polyacyltrehalose Biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **284**, 12745-12751 (2009).
13. Converse, S.E. et al. MmpL8 is required for sulfolipid-1 biosynthesis and *Mycobacterium tuberculosis* virulence. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 6121-6126 (2003).
14. Takayama, K., Wang, C. & Besra, G.S. Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin. Microbiol. Rev.* **18**, 81-101 (2005).
15. De Sousa-D'Auria, C. et al. New insights into the biogenesis of the cell envelope of corynebacteria: identification and functional characterization of five new mycoloyltransferase genes in *Corynebacterium glutamicum*. *FEMS Microbiol. Lett.* **224**, 35-44 (2003).

5 Chapter 5: Labeling of mycobacteria *in vivo*

Chapter 5 reports mycobacteria selective labeling by trehalose *in vivo*. Chapters 3 and 4 demonstrated that the Ag85 enzymes tolerate a wide variety of modifications to the trehalose scaffold and that fluorescein modified trehalose (FITC-Tre) efficiently labels the glycolipids of mycobacteria in cell culture. Here we reveal that FITC-Tre is the first small molecule to selectively label bacilli within macrophages infected with *Mycobacterium tuberculosis*. Synthesis and implementation of an ^{18}F-trehalose derivative as a PET probe in rabbits infected with *Mtb*, suggests a utility for trehalose compounds in clinical applications. Trehalose probes offer uses in a variety of basic and clinical research applications in the field of *Mtb* diagnostics and disease monitoring.

5.1 Introduction

Prior imaging work with *Mtb* in infected macrophages has utilized several strategies[1,2], but none allowed for non-toxic imaging of bacteria *in vivo*. Bacteria can be labeled with Texas-Red fluorophore through fairly harsh, non-selective oxidation with sodium periodate prior to infection[1]. Although toxic towards *Mtb*[3], fluorescently-labeled vancomycin has been used to track the cell division patterns of *M. smegmatis* and *M. bovis* BCG[2], but its toxicity towards *Mtb*, may complicate experimental results. Antibodies can be utilized to label *Mtb* upon fixation and for enzyme linked immunosorbent assays (ELISA)[4], but are not useful for live imaging. PET/CT imaging of patients infected with mycobacteria has shown promise when conducted with the non-selective ^{18}F-2-fluoro-2-deoxy-glucose (FDG) probe, suggesting that more selective PET probes should have utility for monitoring tuberculosis response to treatment[5,6]. In contrast to all previous methods, FITC-Tre shows no apparent toxicity towards *Mtb* or macrophages and offers a mechanism-based method of labeling and imaging viable mycobacteria within cell culture.

5.2 Results

5.2.1 Labeling of *Mtb* in macrophages by ^{14}C-Tre

Prior to an investigation of FITC-Tre or ^{18}F-Tre *in vivo* labeling, ^{14}C-Tre was used to demonstrate the uptake of unmodified trehalose, not only in *Mtb* cell culture (**Chapter 4**), but also into macrophages infected with tuberculosis. Macrophage infection is an important feature of *Mtb* pathogenesis and persistence[7]. *Mtb* is thought to infect the alveolar

macrophages that line the lung epithelium; the macrophage uptake therefore represents an additional permeability hurdle that separates any potential probe from *Mtb*. Not only did ^{14}C-Tre successfully penetrate and internalize within murine macrophages, it was also subsequently taken up by *Mtb* growing within these cells, thereby validating the potential of Tre probes *in vivo* (**Figure 5.1**). While ^{14}C-Tre allows direct verification of uptake, these probes are not useful for non-invasive imaging and microscopy.

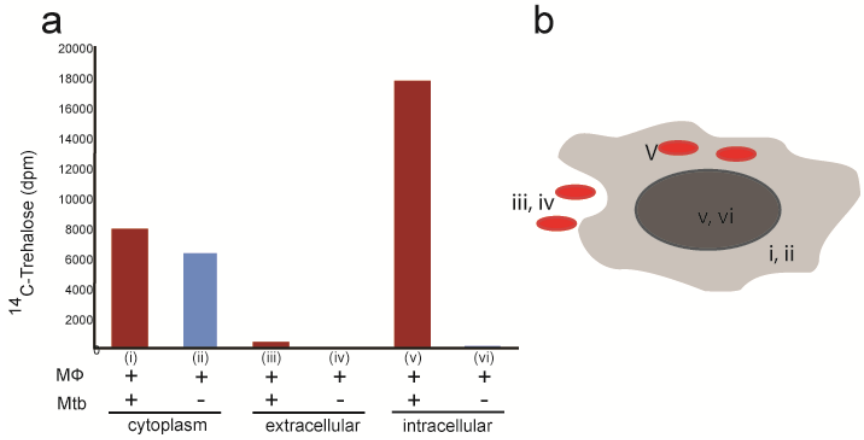


Figure 5.1. ^{14}C-trehalose is taken up into *Mtb*, residing within macrophages. (a) Uptake of ^{14}C-trehalose into TB infected macrophages (i) cytoplasm of infected macrophages (ii) cytoplasm of control macrophages (iii) floating *Mtb* from infected macrophages. Potentially, bacteria released after infection (iv) floating debris from control macrophages (v) *Mtb* from infected macrophages (vi) extracted pellet from control macrophages. Bars 'i' and 'ii' are 1 mL aliquots from 10 mL of the lysate of the macrophages. Therefore, total radioactivity for the macrophages should be about 10 fold higher, approximately 80,000 DPM. Accumulation of 20,000 DPMs of radioactivity into the *Mtb* within the macrophages suggests significant localization of the sugar. (b) Scheme of different cellular compartments evaluated for ^{14}C-trehalose radioactivity.

5.2.2 Labeling of *Mtb* in macrophages by FITC- Tre

Given that ^{14}C-trehalose readily labels *Mtb* residing within macrophages, we turned our attention towards the development of fluorescent probes that would be compatible with microscopy and allow for facile labeling of *Mtb in vivo*. Fluorescein labeled trehalose (for synthesis see **Chapter 2**) was added to macrophages (J774 cell line) infected with H37Rv *Mtb*. Small amounts of the compound (< 1 mg) were initially synthesized for proof of principle experiments. When macrophages infected with *Mtb* were treated with FITC-trehalose, fluorescent, rod-shaped putative bacteria were observed (**Figure 5.2**). Intensity of labeling was low and some background was observed, potentially due to low concentration of probe and impurities within the test sample of the FITC-Tre. However, these results

suggested that the FITC-trehalose penetrated the macrophages and labeled the *Mtb* residing within.

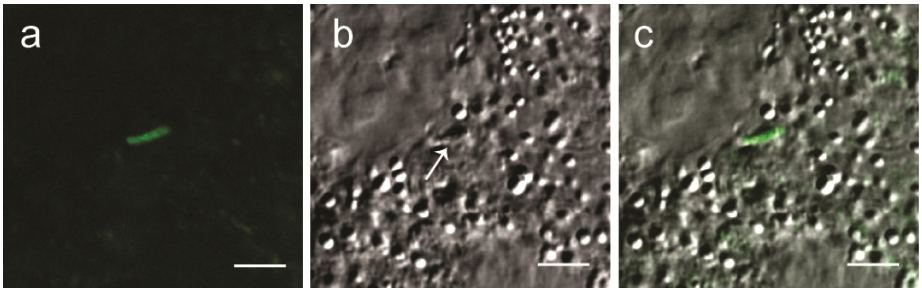


Figure 5.2. First examples of FITC-Tre treated macrophages infected with *Mtb* suggested that bacilli-selective labeling was occurring. (a) FITC-Tre labeled macrophages infected with H37Rv *Mtb*. (b) DIC image with arrow indicating putative mycobacteria (c) overlay of channels 'a' and 'b.'

With the synthesis of additional, high purity FITC-Tre, the labeling of *Mtb* by FITC-Tre within macrophages was demonstrated to be a highly selective, low background method to detect mycobacteria within cells (**Figure 5.3**, **Figure 5.4**). Initial labeling experiments were conducted with macrophages infected with H37Rv, first washed to remove any extracellular bacilli and then labeled with FITC-Tre. With the addition of greater concentrations of FITC-Tre, some background was observed when macrophages were fixed with buffered formalin. A permeabilization step (Triton X-100, 0.1% in PBS) was found to remove all background and afforded clear labeling of rod-like bacterial structures that were also visible in the DIC image as bacteria.

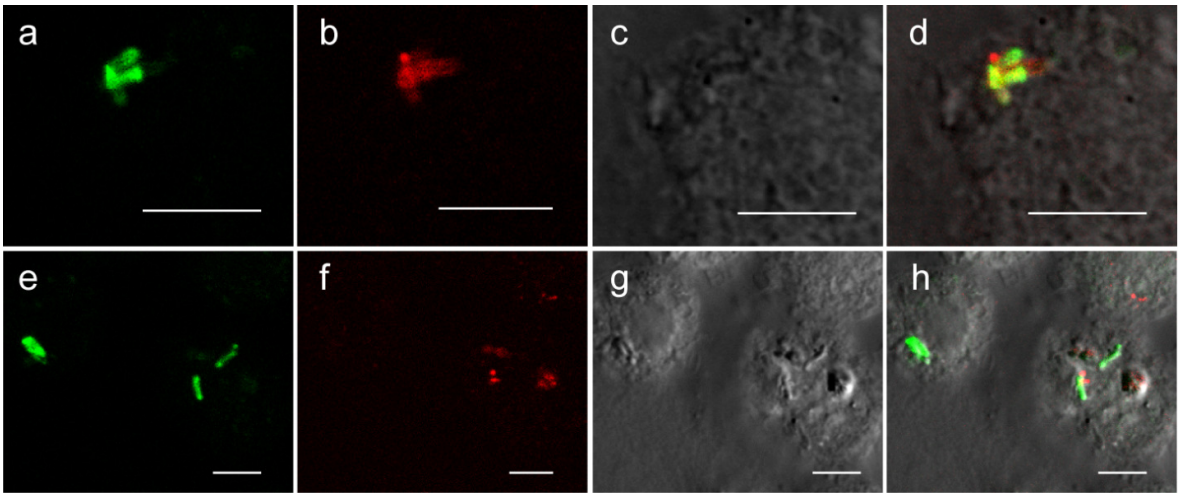


Figure 5.3. Antibody confirms that FITC-Tre probe labels *Mtb*. (a,e) H37Rv *Mtb* are labeled with FITC-Tre and (b,f) treated with anti-*Mtb* antibody (AbCam ab905) while (c,g) shows the differential interference contrast (DIC) and (d,h) is an overlay of all three channels and shows colocalization of FITC-Tre and antibody. Scale bars 5 µm.

Although rod-like, mycobacteria-shaped, fluorescently labeled structures were clearly visible in the DIC and fluorescent images, additional labeling was required to confirm that these structures were mycobacteria. Immunocytochemistry demonstrated that the FITC-Tre labeled rods were conclusively mycobacteria. The specificity of the FITC-Tre label for *Mtb* was demonstrated by colocalization of staining with an *Mtb*-specific antibody (**Figure 5.3**). Some, but not all, bacteria were labeled well by the antibody system employed, a fragment antigen-binding (Fab) based kit that eliminated the need for a secondary antibody[i]. Subsequent use of a primary and secondary antibody systems by others in the Barry group suggested that some of the poor labeling that was observed was in part due to the Fab system rather than the antibody itself (Dr. Laura Via and Dr Young-Soon Yoon, data not shown). Antibody colabeling experiments demonstrated that FITC-Tre labels mycobacteria within macrophages and was compatible with standard immunocytochemistry protocols. Well documented problems with mycobacterial acid fast stains[8] and antibody selectivity[9], combined with the sporadic labeling shown here, demonstrates that a major tuberculosis research problem is the absence of reproducible tools with which to label *Mtb*.

[i] Zenon ® Mouse IgG1 Labeling Kits, Molecular Probes

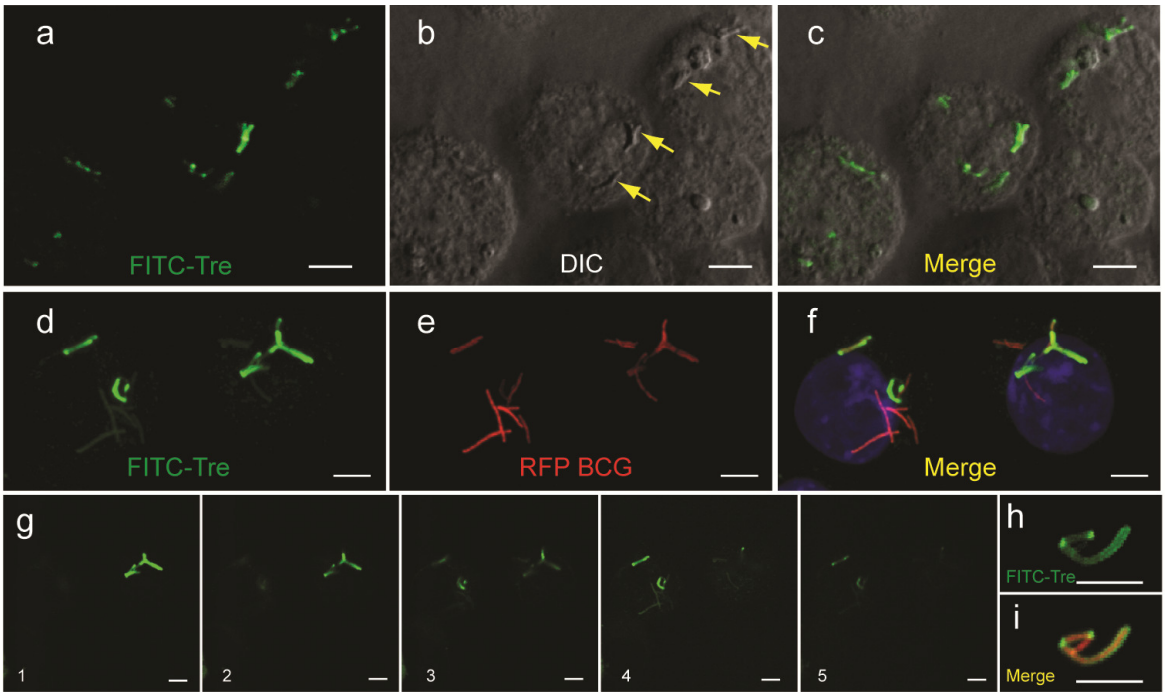


Figure 5.4. Macrophage infection experiments reveal differential accumulation of the FITC-Tre probe, particularly at the cell wall and poles. Macrophages were infected with RFP BCG or H37Rv Mtb where indicated. Red (Red Fluorescent Protein, RFP), green (FITC-Tre). DNA of macrophages is stained blue with 4',6-diamidino-2-phenylindole (DAPI). (a) Typical *Mtb* labeled with FITC-Tre are depicted. In (b) bacteria are clearly visible in DIC image, indicated by white arrows. (c) Shows an overlay of DIC and fluorescent image. (d) Macrophages were infected with BCG expressing RFP and FITC-Tre labeled, Z stack max projection through cell and (e) shows RFP labeling of panel 'd.' (f) Overlay of 'd-e.' (g) (1-5) A stack through cells depicted in 'd-f' shows the differential accumulation of FITC-Tre (1 µm between slices). (h) FITC-Tre polar labeling of RFP H37Rv *Mtb* infecting macrophages. (i) RFP overlay of 'h' shows the accumulation of the probe on the cell wall, outside of the RFP label. Scale bars 5 µm.

In combination with the antibody colocalization, FITC-Tre labeling was also confirmed using macrophages infected with a red fluorescent protein (RFP) expressing *Mtb* and Bacillus Calmette-Guérin (BCG, M bovis)[10] strains treated with FITC-Tre in an analogous manner (**Figure 5.8**). Selective labeling of *Mtb* was observed within 4 hours, at FITC-Tre concentrations down to 50 µM with a large increase in fluorescence over time consistent with ongoing Ag85-mediated uptake. Optimal labeling was achieved with FITC-Tre at 200 µM and with probe incubation times >24 hours. Tagging of BCG and also of *M. avium* (not shown) confirmed that the FITC-Tre labeled a variety of mycobacteria within macrophages. Colocalization between the FITC-Tre and RFP expressing strains conclusively demonstrated that the FITC-Tre accumulation in rod-shaped objects was labeling of mycobacteria, rather than of other putative-bacteria shaped cellular components.

5.2.3 Establishing the non-toxic, selective nature of the FITC-Tre

The RFP and antibody experiments firmly established that FITC-Tre tags *Mtb* within macrophages. Next the scope and limitations of this labeling system were explored. FITC-Tre had been found by MIC to have no inhibitory effect on mycobacterial growth in cell culture (**Chapter 3**). To transfer this result to macrophage culture and to further substantiate the argument that FITC-Tre labeling was non-toxic and did not perturb mycobacterial growth within macrophages, growth inhibition curves were obtained from macrophages infected with *Mtb.* Macrophages were infected with RFP *Mtb* in 96 well format and treated with ethanol or FITC-Tre in ethanol and then monitored for RFP fluorescence increase as a metric of growth. No apparent inhibition of *Mtb* growth was observed either under low infection or high infection of bacteria (**Figure 5.5**). In addition to having no MIC, this data supports that FITC-Tre does not inhibit bacterial growth *in vitro* or *in vivo.* Throughout all labeling experiments, no significant changes in macrophage growth or viability were observed, both visually and by trypan blue staining (haemocytometer). FITC-Tre showed no toxicity towards both macrophages and the mycobacterium residing within.

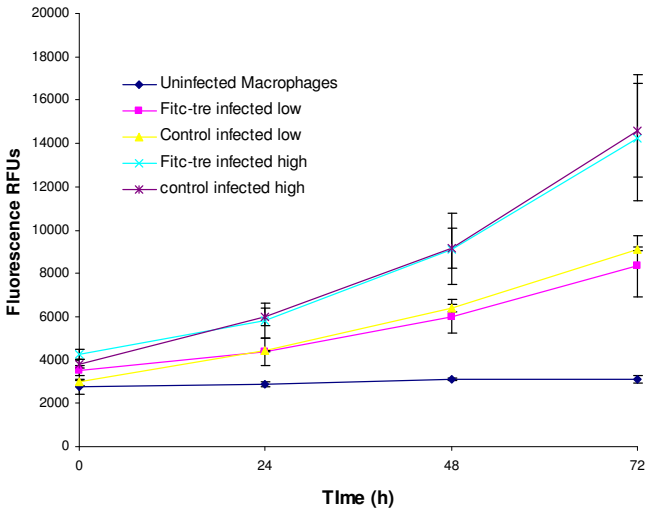


Figure 5.5. Growth inhibition curves of RFP TB residing within J774 macrophages.

5.2.4 Determining the selectivity of the labeling

Having ruled out concerns regarding toxicity, we sought to establish that the FITC-Tre labeling in macrophages was mechanism-based and linked to the incorporation of trehalose by Ag85. 2-FITC-glucose[ii] was utilized as a comparative probe, featuring similar solubility properties and the identical linkage to that found in FITC-Tre. Failure to observe significant labeling with FITC-glucose, under identical conditions to FITC-Tre labeling, suggested that the labeling depended on the presence of the trehalose moiety (**Figure 5.6**).

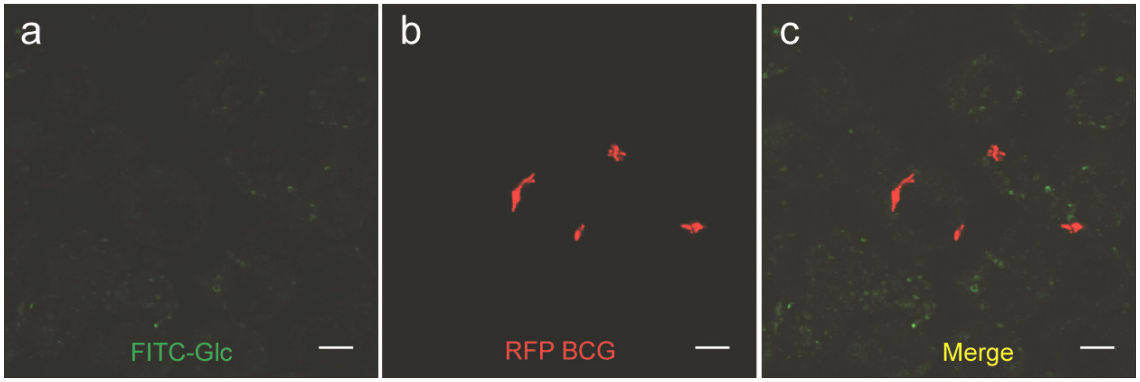


Figure 5.6. FITC-Glc labeling in infected macrophages (a) RFP BCG (excitation 594 nm and emission spectra recorded at 602-651 nm) (b) FITC-Glc label (excited at 488 nm and emission spectra were recorded at 497-557 nm) (c) overlay of RFP and FITC with FITC gain enhanced to show that there is no colocalization. Scale bars 5 µm.

It was hypothesized that unlabeled trehalose could be employed to compete with FITC-Tre labeling in a manner consistent with the *in vitro* trehalose competition experiments (**Chapter 4**). Gratifyingly, labeling of the intra-macrophage bacteria was also effectively abolished by competing administration of high concentrations (10 mM) of trehalose (**Figure 5.7**), which further supports that the FITC-Tre labeling is a trehalose-dependent process.

[ii] Synthesized by Dr. Francois D'Hooge. For synthesis see Appendix 1.

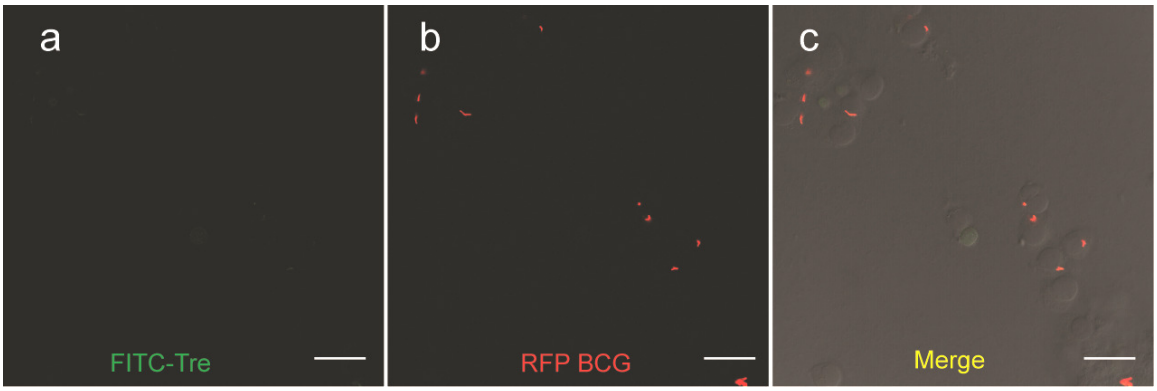


Figure 5.7. Trehalose saturation of FITC-Tre treated BCG shows no FITC-Tre labeling *in vivo*, demonstrating that the FITC-Tre labeling is trehalose dependent. (a) FITC-Tre treated BCG expressing RFP shows no green fluorescence (b) RFP channel of image in (a) and (c) shows an overlay of FITC-Tre, RFP and DIC images. Scale bars 25 µm.

As an alternative mode of fluorescence label, quantum dots conjugated to trehalose (QD-Tre, **55**) were also synthesized (**Chapter 2**). This substrate also co-localized (**Figure 5.8**) with *Mtb* in infected macrophages, corroborating a Tre-dependent and *Mtb* selective labeling that is not dependent on the source of fluorophore used. Preliminary experiments revealed that Carboxyrhodamine-trehalose (Rho-Tre, **56**) (**Chapter 2**) also labeled *Mtb*. However, these results were complicated by the background 'stickiness' of rhodamine. Due to previously reported ability of the fluorophore to adhere to mycobacterial cell walls (acid fast stains[11-13], further experiments with the Rho-Tre were discontinued. In summary, trehalose dependence of labeling suggests an Ag85 based mechanism of incorporation.

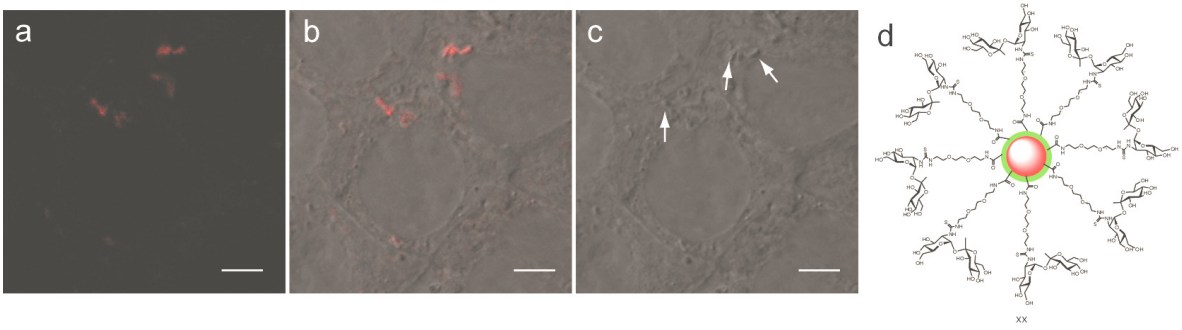


Figure 5.8. Trehalose-Quantum dot (Tre-QD) labeled *Mtb* show that trehalose labeling is independent of the type of fluorophore appending the trehalose. (a) QD labeled *Mtb* (excitation 562 nm. emission 585nm – 634 nm) (b) Overlay of QD image with DIC (c) DIC image of infected macrophages treated with Tre-QDs. Arrows indicate putative *Mtb* identified in DIC image (d) Tre-QDs with ~110 sugars/dot. Scale bars 5 µm.

5.2.5 Differential labeling by FITC-Tre of *Mtb* macrophages

In contrast to the fairly uniform labeling of bacilli observed *in vitro* (**Chapter 4**), bacterial labeling *in vivo* was not uniform, either between cells or within a cell. In both *in vitro* and *in* vivo experiments *Mtb* accumulated FITC-Tre at the poles and cell wall, consistent with the Ag85 dependent labeling hypothesis (**Figure 5.4**). However, during growth in macrophages some *Mtb* accumulated vast amounts of probe, particularly at the poles and cell wall, whereas some bacterial cells appeared to incorporate very little label. With the RFP reporter strains, we observed that even mycobacterial cells within a single macrophage displayed very different labeling intensities. We speculated that this differential FITC-Tre labeling might be relevant to the growth status of a particular bacterium within the cell reflective of the maturation status of the endosomal compartment within which they were located. Alternatively, the amount of labeling could correlate with the amount of probe available to the bacteria. Professor David Russell's lab has observed similar heterogeneity among populations of *Mtb* expressing a pH sensitive fluorescent protein construct and they also cannot fully explain the observed phenomena[iii].

Bacteria residing in phagolysosomes would be expected to have lower metabolic activity (and therefore lower Ag85 expression levels) compared to mycobacteria occupying the more hospitable phagosomes. Additionally, the Ag85 enzymes have shown decreased catalytic efficiency below pH 6[14], so decreased transmycolation would be expected in more acidic vesicles. TDM is thought to play an important role in the mycobacterial persistence within host cells. It has been shown to inhibit fusion between vesicles[15] and is potentially responsible for the inhibition of phagosomal acidification by *Mtb*[16]. Because TDM was thought to inhibit *Mtb* trafficking and the Ag85 enzymes would be more active, and produced at higher levels in the early phagosome, it was hypothesized that bacteria in the early stages of phagocytosis would accumulate greater levels of FITC-Tre and appear more fluorescent when compared with those bacteria in the phagolysosome.

[iii] Personal communication with Professor David Russell, Cornell University, unpublished.

5.2.6 Statistical analysis of BCG vs. *Mtb* fluorescence

It was initially hypothesized that measurable differences in FITC-Tre accumulation between the RFP BCG strain and the RFP TB strain would be apparent, as the less virulent BCG strain should be trafficked more rapidly by the macrophages. Macrophages were infected with RFP BCG or H37Rv and labeled with FITC-Tre and then fixed, mounted and analyzed by confocal microscopy. Z-Stacks were taken through bacteria to capture all fluorescence and allow calculation of the volume. Fluorescence intensity was quantified as relative fluorescence units (RFUs)/vol. Regions of interest (ROIs) were defined by the red channel and the statistic function in Imaris Bitplane 7.3 was used to report the green fluorescence in the red regions. Consistent with what had been hypothesized, the TB strain demonstrated greater FITC-Tre labeling (**Figure 5.9**).

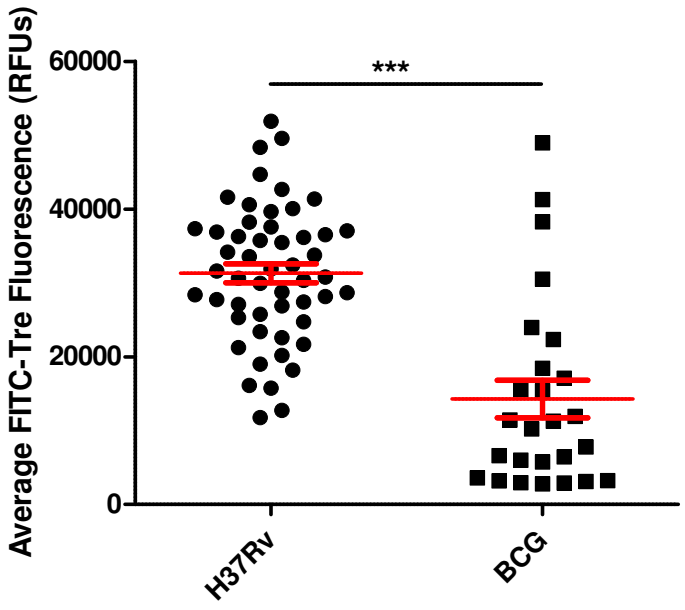


Figure 5.9. Statistically significant difference between TB and BCG labeled with FITC-Tre was observed when bacteria were treated as Regions of interest (ROIs), determined by RFP fluorescence, and the sum of FITC-Tre fluorescence was quantified for each ROI. Each datapoint represents one bacteria. P <0.0001.

Although initially very promising, these results were confounded by problems with fluorescence instability exhibited by the RFP BCG strain. After several passages of the culture, the strain appeared to be converting from an RFP to a GFP strain. In multiple separate experiments, we noticed two nearly separate populations of bacteria, one green and

one red. Many bacteria were red and not green or vice versa. The H37Rv RFP construct was found to be much more stable, producing a negligible number of bacteria that are green but not red. Missing a large subpopulation of bacteria within our data analysis almost certainly confounded the results.

In addition to the red to green conversion, the RFP BCG strain had a low level of green emission when excited with the 488nm laser, noticeable because the red fluorescence intensity actually decreased when the green laser was turned off. Also, the red and green channels had a striking similarity (crosstalk) in some cases. The level of background green fluorescence was lower than the accumulation of FITC-Tre after 24 hours labeling. However, earlier time points and attempts to quantify FITC-Tre labeling were complicated by this RFP-GFP signal.

This instability and background problem produced by the RFP BCG strain is inherent to the type of RFP encoded by the BCG strain, specifically DSRed1. DSRed1 is an early generation RFP protein that is a tetramer, composed of four GFP-like subunits, which was discovered in reef coral and found to be a distant relative of GFP[17]. A problem with the tetramer is that it can degrade to the monomeric GFP form, explaining the green background that was observed in the BCG samples[18]. Fortunately, the RFP *Mtb* strain, which encodes a monomeric version of RFP, mCherry, does not exhibit this problem. After the instability of the BCG strain was discovered, experiments with the RFP BCG were avoided. A control infection without FITC-Tre was always employed to ensure that there was no green background. With these precautions in place, the rest of the experiments proceeded without complications from RFP-GFP background.

5.2.7 Trafficking studies: Labeling of endosomes and lysosomes

The comparison between BCG and TB was suggestive, but further experimentation was required to link the FITC-Tre label to location of individual bacilli in the phagocytotic pathway. Colocalization studies with the early endosome antigen 1(EEA1) and lysosomal associated membrane protein 2 (LAMP2) were implemented, first in J774 macrophages. LAMP2 is one of two proteins LAMP1 and 2 that constitute an estimated 50% of the lysosomal membrane. EEA1 is an antibody against the early endosomes, specifically Early Endosome Antigen 1 protein that binds phospholipid vesicles containing phosphatidylinositol 3-phosphate. Labeling with fluorescent dextrin (lysosome maker) and/or transferrin

(endosome marker) was also attempted, but poor uptake and low signal limited the utility of both of these markers. Initially it was hypothesized that greater colocalization would be observed between the EEA1 marker and the FITC-Tre labeled bacteria, suggesting localization in the early phagosome. Accumulation of the LAMP2 marker was postulated to be prevalent at the bacteria that did not label with FITC-Tre, reflective of a phagolysosomal location.

As a control, we first confirmed that these antibodies do not colocalize by a dual stain with a Rabbit EEA1 antibody and a Rat LAMP2 antibody, followed by FITC-anti-rabbit and a Cy-5.5 anti-rat secondary antibodies. In some strains of J774 macrophages, LAMP2 stains the endosomes as well as the lysosomes and so it was gratifying to observe very little colocalization between the two antibodies (**Figure 5.10**). Colocalization was low, about 2.5%, confirming that the antibodies did not exhibit cross reactivity.

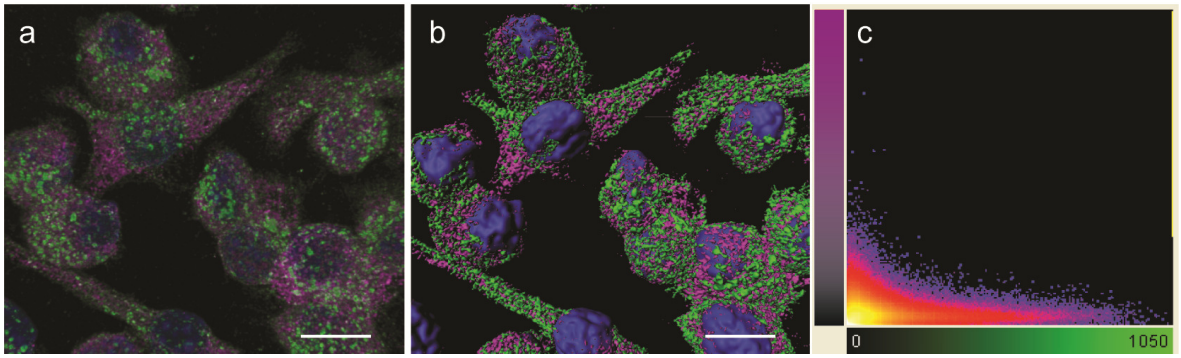


Figure 5.10. Low cross reactivity of EEA1 and LAMP2 antibodies. (a) EEA (green), LAMP2 (purple), DAPI (blue), (b) 3D reconstruction of 'a,' (c) histogram of colocalization of purple and green. Scale bars 10 µm.

In infected macrophages the antibody labeling proceeded well, but halos of antibody surrounding the bacteria were not observed, contrasting what had been expected for vesicle-shaped phagosomes. Instead fragments of the endosome and lysosomal membranes were detected. These results are consistent with those obtained previously by Deretic[19]**.** This fragmented labeling complicated correlating the endosome or lysosome labeling with bacterial color because only parts of the bacteria colocalized, i.e. colocalization of pixels, rather than of whole bacteria. Colocalization can be automated in Imaris software and the program will calculate the percentage material (takes into account both intensity and volume) or simply percentage volume that colocalizes between two different signals observed on the

microscope. The green signal and antibody signal (far red, colored magenta) and the RFP signal and magenta were evaluated for colocalization. This analysis afforded a percentage colocalization between the FITC-Tre labeled subpopulation of bacteria and the antibody or the RFP expressing entire population, neglecting a critical subpopulation, the bacteria that express RFP but do not label with FITC-Tre (red-but-not-green).

5.2.8 *Mtb* subpopulation that does not label with FITC-Tre (red-but-not-green)

Some creative image analysis techniques were attempted to try to generate a ROI for the red-but-not-green bacteria. A subtraction of the green channel from the red channel was attempted, which should have afforded the red-but-not-green ROI. However, the relative different intensities between the two channels resulted in an incomplete subtraction, leaving unwanted portions of bacteria in the image. This problem could be circumvented by generating ROIs for each color and then subtracting the green ROI from the red ROI. Because FITC-Tre labeled the cell wall of the *Mtb*, whereas the RFP was cytosolic, the green ROI often surrounded the red ROI, meaning that when the ROI subtraction was attempted, we were left with RFP centers of some bacteria that were labeled with FITC-Tre. As neither of these strategies proved feasible, all of the studies reported here reflect the colocalization between these markers and the entire bacterial population and with the FITC-Tre labeled subpopulation, unfortunately neglecting the red-but-not green subpopulation.

5.2.9 FITC-Tre labeling may be linked to mycobacterial trafficking

Preliminary colocalization studies with EEA1 and LAMP2 suggested little to no difference between the FITC-Tre bacteria and the entire population (data not shown). Several reasons for this were considered. The bacteria may continue to traffic during the labeling, which might affect the fluorescence. Phagosomal escape has both been documented[20]. One could envision a bacterium in the phagolysosome escaping to the cytoplasm and then undergoing Ag85 dependent FITC-Tre labeling. Likewise bacteria could label in the endosome and then traffic to the lysosome, thus affording green bacteria in the lysosome.

Interferon-γ (IFN), a well known activator of macrophages, was employed to try to force additional phagosome lysosome trafficking and, as a result, greater differences between the percentage colocalization with EEA1 and LAMP2. The setup of the experiment was the following: activate macrophages upon infection, wait 4 hours, treat with FITC-Tre, expose overnight and fix, permeabilize, block, mount and analyze. For each antibody 3 slides and 3-

4 fields of view per slide were evaluated in order to reduce the possibility of error from one particular slide. Unactivated macrophages were evaluated in parallel. LAMP antibody was used to confirm that the activation was successful. Macrophages elongated and flattened out with treated with IFN, which is readily apparent from a LAMP stain (**Figure 5.11**). A seeming decrease in the amount of high-level FITC-Tre fluorescence upon IFN treatment was observed. However many of the bacteria were low-level labeled, so it was hard to determine the significance of this decrease (**Figure 5.11**). For this data any ROI with a volume greater than 20 was excluded, as these represented large clumps or bacteria, which traffic differently from individual bacteria[21]. In subsequent experiments macrophage activation was conducted the day prior to infection instead of concurrent with infection, which may show larger changes in labeling.

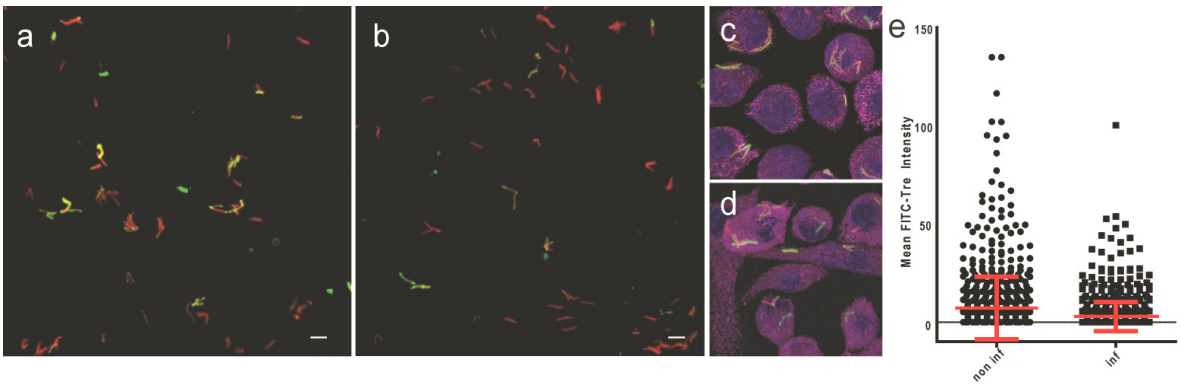


Figure 5.11. Imaging of J774 macrophages infected with RFP expressing H37Rv suggests a decrease in FITC-Tre accumulation when macrophages are activated with IFN-γ. (a) Shows cells not activated with INF-γ. Overlay of green and red channels shows slightly higher FITC-Tre probe accumulation than in (b) where an overlay of green and red channels of cells were activated with INF-γ. (c) depicts non-IFN-γ LAMP2 antibody stain. (d) IFN-γ LAMP2 antibody stain shows macrophages have flattened and elongated, confirming that activation was successful. (e) Quantification of mean FITC-Tre intensity for non-IFN and INF-γ macrophages shows a decrease in intensity under IFN-γ. Each point represents one bacterium or if touching, multiple bacteria. Scale bars 5 µm.

Under IFN-γ activation the colocalization of FITC-Tre bacteria with the lysosome marker LAMP2 was interrogated. It was hypothesized that since the FITC-Tre is an enzymatic readout, the Ag85 enzymes should be less active in the acidic lysosomal compartment and thus there should be less FITC-Tre labeling when colocalized with LAMP, particularly under INF-γ activation. However, when the amount of LAMP2 colocalization for both IFN and non-IFN treated macrophages was quantified, no statistically significant difference was observed between FITC-Tre and the population of bacteria as a whole (represented by the red

bacteria, as all bacteria are red). The Pearson scores[22] for colocalization were low, indicating that the majority of bacteria were actually not colocalized with the LAMP2. Likewise, colocalization was typically limited to a few pixels on the length of the bacteria (**Figure 5.12**). These results suggested that there is no difference in colocalization between the FITC-Tre labeling *Mtb* subpopulation and the entire bacterial population in J774 macrophages.

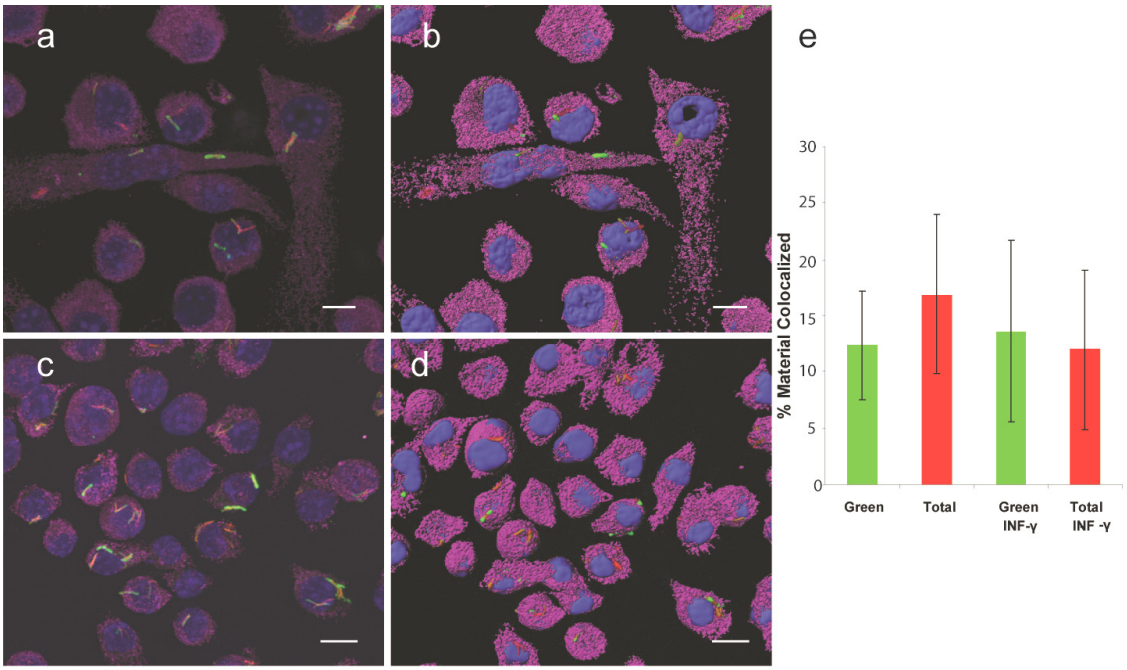


Figure 5.12. J774 macrophages infected with *Mtb* and labeled with FITC-Tre show no significant difference in colocalization with LAMP2 in the presence or absence of IFN-γ activation. (a) IFN-γ and LAMP2 treated macrophages, (b) 3D reconstruction of 'a,' (c) Non-IFN-γ and LAMP2 treated macrophages, (d) 3D reconstruction of 'c.' Scale bar 10 µm.

Colocalization between early endosome antigen 1 (EEA1) and RFP *Mtb* was quantified in a similar manner to the LAMP2 colocalization studies. Colocalization was very low in unactivated macrophages. However, in activated macrophages, the antibody labeling increased and an enhanced colocalization was observed. More colocalization of low-FITC-Tre *Mtb* with EEA1 under INF-γ treatment was the opposite of what had had been anticipated and would suggest a correlation between FITC-Tre label and increased trafficking. These results were collected over a small sample set, which could bear repeating. Part or all of this increase in colocalization may be due to increased production of EEA1 under macrophage activation[23], rather than actual differences in colocalization with bacterial subpopulations.

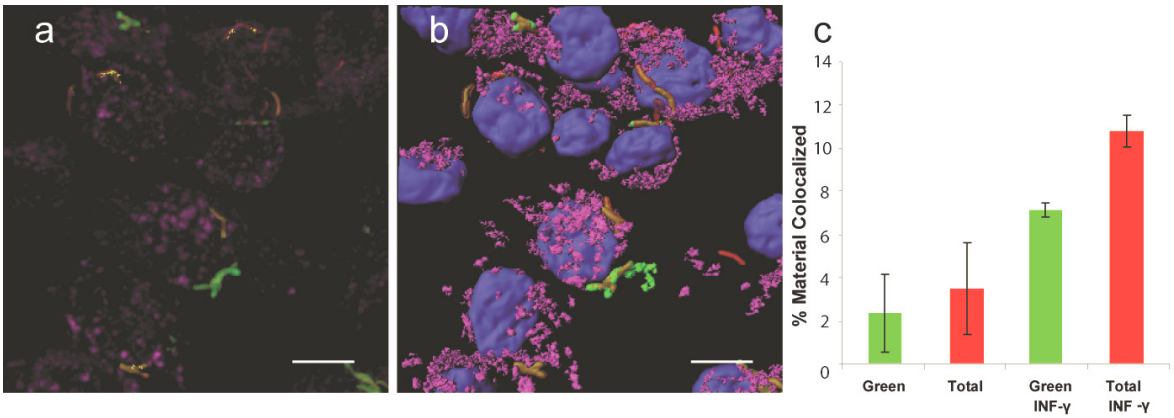


Figure 5.13. Colocalization of EEA1 and FITC-Tre or RFP shows an increase under IFN-γ treatment. (a) Macrophages are labeled with EEA1 (purple). Colocalization is shown in yellow. (b) 3D reconstruction of 'a' (c) EEA1 colocalization is quantified. Scale bar 5 µm.

Because of the difficulties ascertaining whether FITC-Tre was labeling selectively during a particular stage of infection, phagocytotic literature was surveyed to determine whether additional antibody markers existed that might prove useful for teasing apart this problem. We discovered that EEA1 and LAMP2 are markers for very early endosomes and very late lysosomes respectively (discussed in **Section 5.2.5**). These antibodies failed to capture the majority of *Mtb,* which reside somewhere between the early endosomes and the lysosomes. The literature survey revealed a wide variety of alternative markers for different stages in the endosome/phagosome pathway. A sampling of these antibodies was obtained and implemented. To summarize briefly, Rab5[24,25], Rab14[26], LAMP1[27,28], EEA1[19] and Pro-Cathepsin D[29] were selected, as these had been used previously as good endosomal, phagosomal and lysosomal markers.

Rab5 belongs to a class of over 60 low molecular weight Ras GTPases that cycle between their active GTP and inactive GDP bound forms. They function as molecular switches to promote vesicle fusion[30,31]. Duclos et al found that Rab5 regulates the 'kiss and run' fusion between phagosomes and endosomes[32]. It has been well documented that a dynamic switch exists between Rab5 and Rab7 with the loss of Rab5 and gain of Rab7 leading to phagosomal maturation[25]. Over-expression of Rab5 led to uncontrolled formation of giant phagosomes. All reports suggest high colocalization between Rab5 and the mycobacterial phagosome.

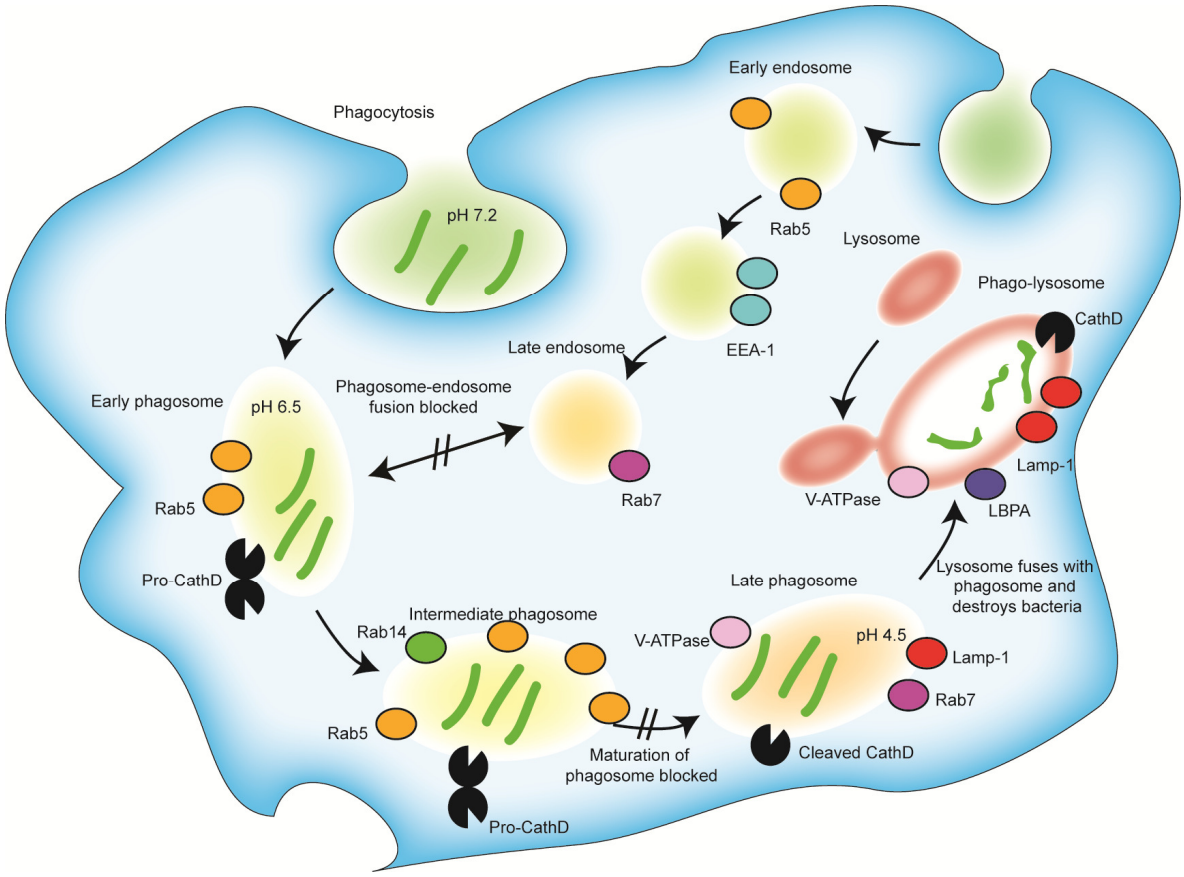


Figure 5.14. Diagram of mycobacterial phagosome formation and persistence. Based loosely on figure 2 from Kaufmann[33].

Branching out into other Rab proteins, two additional candidates were found to be important for phagosomal formation and persistence. First, Rab7 has been implicated in the arrest of phagosomal maturation, specifically in the stage of vesicle fusion regulated by Rab5 and Rab7. Via et al found that selective retention of Rab5 and exclusion of Rab7 is linked to mycobacterial persistence within the phagosome and avoidance of maturation of fusion with lysosomes[25]. Rab7 was not selected as a marker, because an increase rather than decrease in colocalization with the phagosome was desired for this study.

In a subsequent publication, Deretic's lab found that Rab14, like Rab5, is critical for maintenance of the mycobacterial phagosome maturation arrest[26]. In contrast, work from Brumell's lab[34], evaluating the whole network of Rab GTPases in reference to Salmonella only found modest colocalization between the salmonella phagosome and the Rab14 antibody. These results conflict with those from Deretic's group and so Rab14 differential colocalization was investigated.

To look at the lysosomes and bacteria that were in more acidified phagosomes and phagolysosomes, several antibodies were considered, specifically Lamp1, VATPase and cleaved Cathepsin D (CathD). LAMP1 is the lysosomal associated membrane protein, and as was discussed earlier, colocalizes with the lysosomes and phagolysosomes. V-ATPase, also known as Vacuolar-type H^+-ATPase, couples the energy from the hydrolysis of ATP to transport protons across a membrane, creating a pH gradient. It consists of 6 different domains and an antibody against the G1 subunit was obtained. Barker et al[35] uses V-ATPase, LAMP1 and cleaved CathD antibodies to demonstrate that *M. marinum* circumvents the endocytotic pathway and can reside in an intracellular compartment that is not acidic and does not fuse with lysosomes. They also compared live versus heat killed bacteria and latex beads to show an increase in colocalization for beads and dead bacteria, particularly with CathD and V-ATPase. The increases in colocalization were particularly pronounced with V-ATPase and CathD and more modest with LAMP1. All three antibodies were obtained to compare their colocalization.

Although procurement of the antibody against cleaved CathD was intended, in reality, the antibody against Pro-cathepsin D, the uncleaved form of this protease was actually obtained. This mistake was fortuitous, as the Pro-CathD has been demonstrated to be found in phagosomes. Its cleavage is one of many important steps towards phagosomal maturation. Cleaved CathD colocalizes with lysosomes, whereas the pro form colocalizes with phagosomes[29]. Presence of Pro-CathD is an important marker of phagosomal colocalization.

5.2.10 Bone marrow macrophage labeling and colocalization studies

To see whether the hypothesis that FITC-Tre labeling would correlate with phagosomal location was correct, murine bone marrow macrophages (BMMS) were obtained and half of the cells were activated with INF-γ. Subsequently both activated and unactivated BMMS were infected with the RFP H37Rv *Mtb* strain and then labeled with FITC-Tre, fixed and labeled with antibodies. BMMs were obtained by a protocol modified from what has been previously reported[36,37] from Black 6 female mice (see **Appendix 2** for full details). Visual inspection revealed enlargement of the monocytes, elongation and formation of pseudopod shape, which confirmed the generation of BMMs, under stimulation with murine M-CSF (macrophage colony stimulating factor) (**Figure 5.15**).

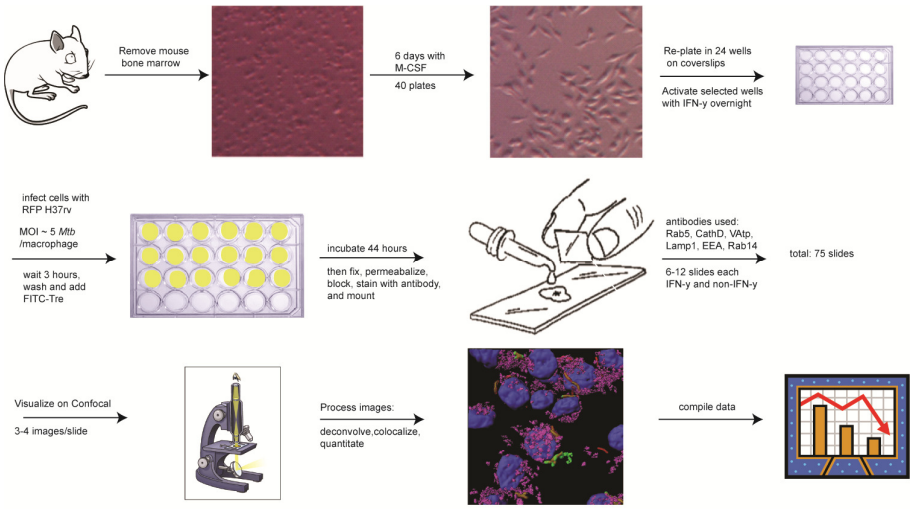


Figure 5.15. Depicts the experimental design for *Mtb* infection and labeling experiment in bone marrow macrophages (BMMS).

After infection and labeling, cells were fixed, permeabilized, stained with the target antibodies, mounted onto slides and visualized via confocal microscope (see **Appendix 2** for full details). Antibodies were either diluted 1:50 or 1:100, depending on preliminary labeling optimization. All experiments were conducted in triplicate and at least three images were collected per slide to verify results. Image averaging of four and bidirectional scanning were employed, the former to decrease the background, the latter to vastly decrease the image collection time. The same microscope and same settings were used collection of for all images so that a comparison of data across different images would be possible.

Once all of the images were collected, images were deconvolved to increase the apparent image resolution. Deconvolution[38] uses an algorithm to reduce the out of focus light in three dimensional microscope images, refocusing the light scattered during image collection. To deconvolve, sampling must be above Nyquist's criteria[39,40], meaning that sampling must be at 2.3 times the highest frequency present in the data. For microscopy in the visible light range, this works out to be a zoom of four and a z stack slice every 0.29 µm. Sampling was within these values, with a zoom of four or six and 0.29 µm slices in the stacks. Huygens deconvolution software was used with consistent deconvolution parameters to deconvolve all images within a particular antibody sample, so as to be able to compare across images.

After deconvolution, all images were imported into Imaris Bitplane 7.3 software for colocalization analysis and calculation of the mean intensity for each color within each bacteria or clump of bacteria. Each bacteria was treated as a region of interest (ROI). If two bacteria were touching or multiple bacteria were found in a clump, they were treated as one ROI because the software cannot easily distinguish objects that are touching. All of this information for both deconvolved and not deconvolved images was collected as comma separated files to be compiled into one dataset for each antibody. A brief comparison between the deconvolved and non-deconvolved colocalization results showed no significant differences in the data output, confirming that deconvolution did not alter the results reported here.

To expedite the data analysis, a series of MATLAB programs were developed in order to compile the data into matrices from comma-separated values (CSV) files. These programs directly extracted and compiled the appropriate data from the CSV files, negating the need to open or copy-paste each data point, and dramatically streamlining the data analysis (see **Appendix 2** for full details).

5.2.10.1 Colocalization results

Colocalization results revealed subtle differences between the FITC-Tre bacterial population and the *Mtb* population as a whole, specifically lower colocalization between FITC-Tre and Rab5 and LAMP1. All antibodies labeled well with the exception of the V-ATPase antibody. However, statistically significant differences in colocalization were most apparent with Lamp1 and Rab5 antibodies. When all datasets were compared, a lower level of FITC-Tre labeling was observed in *Mtb* infecting macrophages under IFN-γ stimulation, compared with those residing in unactivated cells. Taken together, these results suggest that the FITC-Tre label may serve as a marker of an early stage of phagosomal trafficking.

Low colocalization was observed both under activated and unactivated conditions for EEA1 (**Figure 5.16**). EEA1 is actually an effector of the GTPase Rab5, a known phagosomal marker[41], meaning that the presence of large amounts of Rab5 in the phagosome[41] will lead to a paucity of EEA1. Consistent with previous reports of accumulation of Rab5 in the mycobacterial phagosome, high colocalization of Rab5 was observed for both FITC-Tre bacilli and the entire *Mtb* population in both activated and unactivated macrophages (**Figure 5.17**).

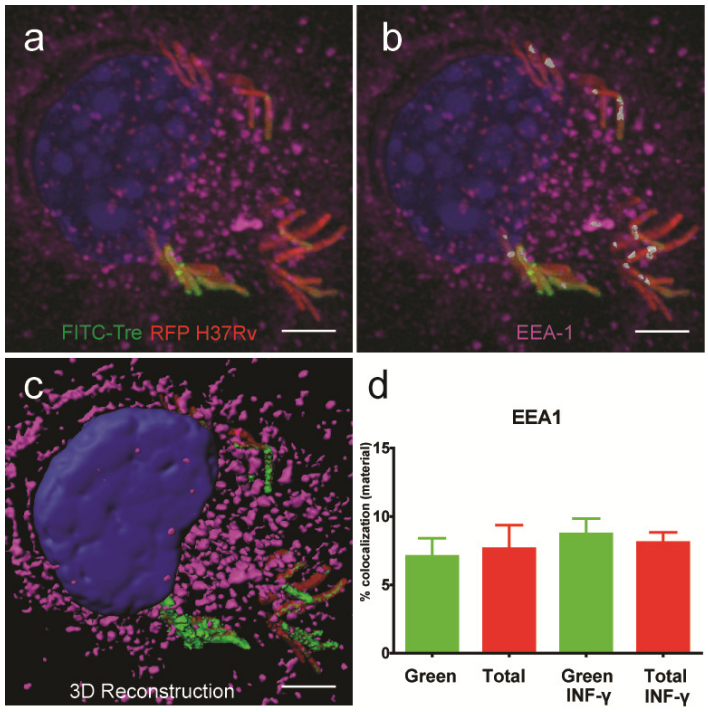


Figure 5.16. Labeling studies with anti-EEA1 in infected bone marrow macrophages (a-d) illustrates triple labeling colocalization studies with RFP, FITC-Tre and anti-EEA1 antibody in bone marrow macrophages (BMMs) activated with IFN-γ. (a) shows the merged RFP and FITC images, (b) the labeling with anti-EEA1 (magenta) with colocalization indicated in white (c) a three dimensional reconstruction with all three labels superimposed. (d) shows quantitation of colocalization of EEA1 with bacteria that were labeled with FITC-Tre (green bars) or were red (total population). Scale bars 5 µm.

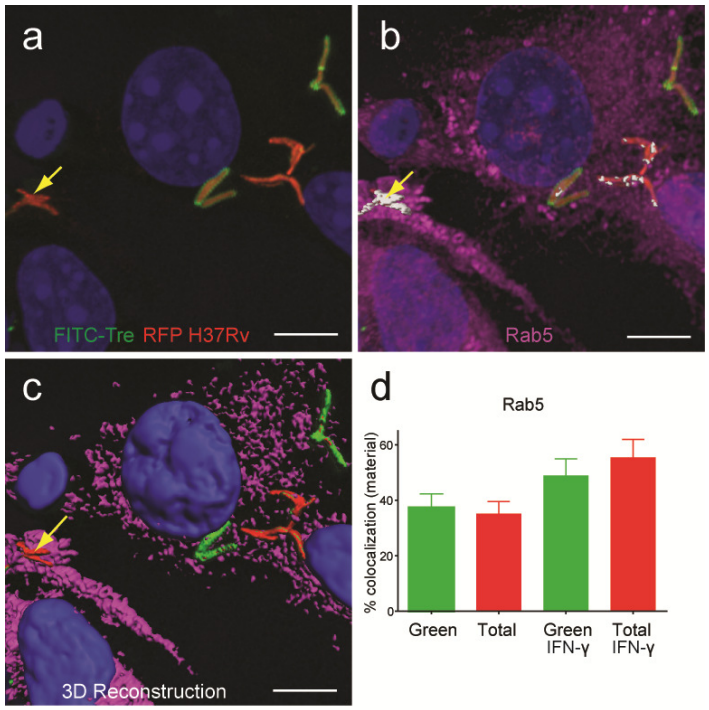


Figure 5.17. Labeling studies with anti-Rab5 in infected bone marrow macrophages (a-d) illustrates triple labeling colocalization studies with RFP, FITC-Tre and anti-Rab5 antibody in bone marrow macrophages (BMMs) activated with IFN-γ. Yellow arrows indicate red bacteria that are not green, i.e. do not take up FITC-Tre and that colocalize to a high degree with Rab5. (a) shows the merged RFP and FITC images, (b) the labeling with anti-Rab5 (magenta) with colocalization indicated in white (c) a three dimensional reconstruction of with all three labels superimposed. (d) shows quantitation of colocalization of Rab5 with bacteria that were labeled with FITC-Tre (green bars) or were red (total population). P-Values: Green vs. Total (ns), Green vs. Green+IFN-γ (0.07), Green vs. Total+IFN-γ (0.02), Total vs. Green + IFN-γ (0.04), Total vs. Total+ IFN-γ (0.008) and Green+IFN-γ vs. Total+ IFN-γ (ns). Error bars are standard error of the mean (SEM) for a minimum of 3 experiments. P value > 0.1 are notated not significant (ns). Scale bars 5 µm.

Upon IFN-γ activation, Rab5 percentage colocalization was found to be lower for highly FITC-Tre-labeled *Mtb* compared with the bulk bacterial population (**Figure 5.17**). This result contrasted with the hypothesis that FITC-Tre labeled *Mtb* should reside in the phagosome and colocalize with Rab5 to a greater extent than the entire bacterial population. Literature precedent suggested, unexpectedly, that Rab5 has also been linked to phagosome maturation. Smith et al[34] screened a panel of GTPases and identified those important for phagosomal maturation in Salmonella, namely Rab5 and Rab7 and Rab11 and Rab14 and several other GTPases to a lesser extent. Rab5a but not Rab5b or c had a correlation with accelerated maturation of Listeria phagosomes[42]. INF-γ was found to activate Rab5a, but not Rab5b or c. The Rab5 antibody employed in this study reacts with all isoforms of Rab5, suggesting that some of the increase in colocalization of the entire bacterial population under

IFN-γ activation may be due to increased colocalization with Rab5a, rather than other Rab5 isoforms.

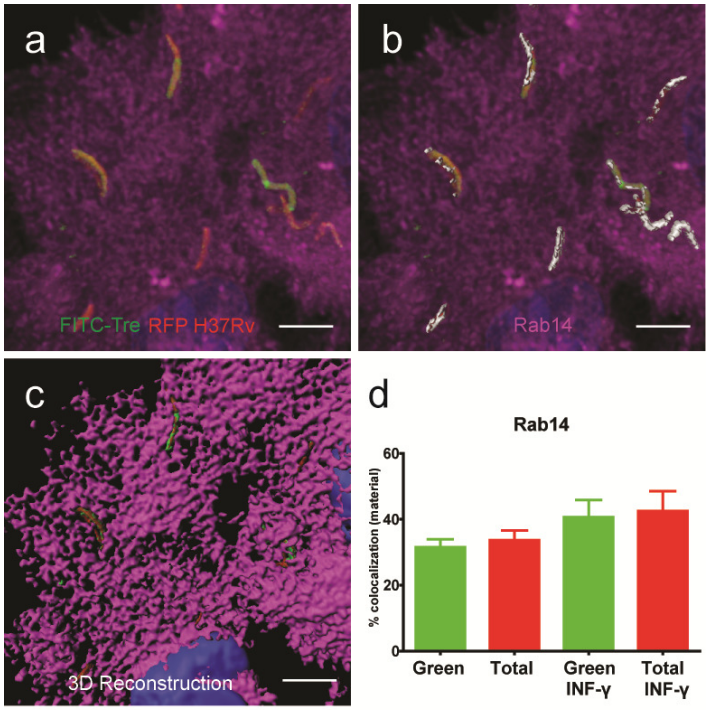


Figure 5.18. Labeling studies with anti-Rab14 in infected bone marrow macrophages (a-d) illustrates triple labeling colocalization studies with RFP, FITC-Tre and anti-Rab14 antibody in bone marrow macrophages (BMMs) activated with IFN-γ. (a) shows the merged RFP and FITC images, (b) the labeling with anti-Rab14 (magenta) with colocalization indicated in white (c) a three dimensional reconstruction of with all three labels superimposed. (d) shows quantitation of colocalization of Rab14 with bacteria that were labeled with FITC-Tre (green bars) or were red (total population). Scale bars 5 µm.

Colocalization between Rab14 and FITC-Tre-labeled *Mtb* compared to the total population of bacteria was similar for both unactivated and IFN-γ treated macrophages (**Figure 5.18**). In contrast, colocalization of Pro-Cathepsin D, another phagosomal marker, increased with FITC-Tre-positive *Mtb* relative to the whole population (**Figure 5.19**). LAMP1 was also found to colocalize to a lesser degree with highly FITC-Tre positive bacteria (**Figure 5.20**) as compared with the population average, particularly upon IFN-γ activation. Together these data are consistent with the hypothesis that *Mtb* that poorly incorporate FITC-Tre localize to Rab5-containing phagosomes that have more fully matured towards degradative lysosomes. The corollary is that those with high Ag85-activity that incorporate FITC-Tre well, and will therefore readily generate TDM, are associated with less-developed phagosomes, perhaps through TDM-associated inhibitory mechanisms[15,43].

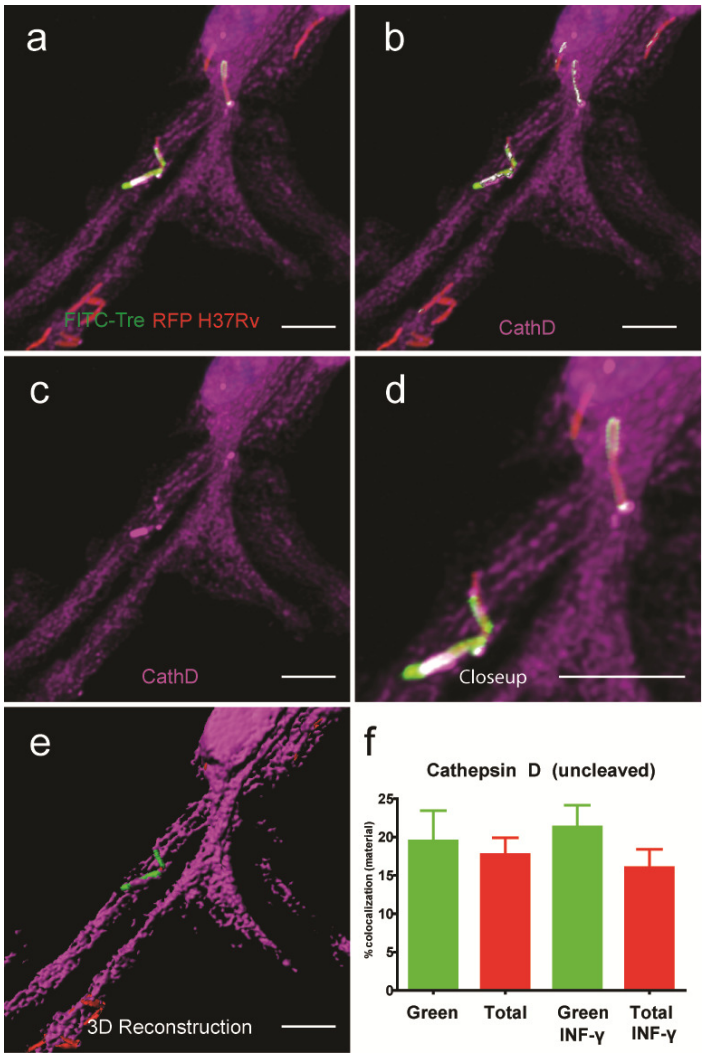


Figure 5.19. Labeling studies with anti-Pro-Cathepsin D in infected bone marrow macrophages (a-d) illustrates triple labeling colocalization studies with RFP, FITC-Tre and anti-EEA1 antibody in bone marrow macrophages (BMMs) activated with IFN-γ. (a) shows the merged RFP and FITC images, (b) the labeling with anti-CathD (magenta) with colocalization indicated in white (c) Only CathD (d) zoomed in to show large accumulation of antibody (e) a three dimensional reconstruction of with all three labels superimposed. (f) shows quantitation of colocalization of CathD with bacteria that were labeled with FITC-Tre (green bars) or were red (total population). Scale bars 5 µm.

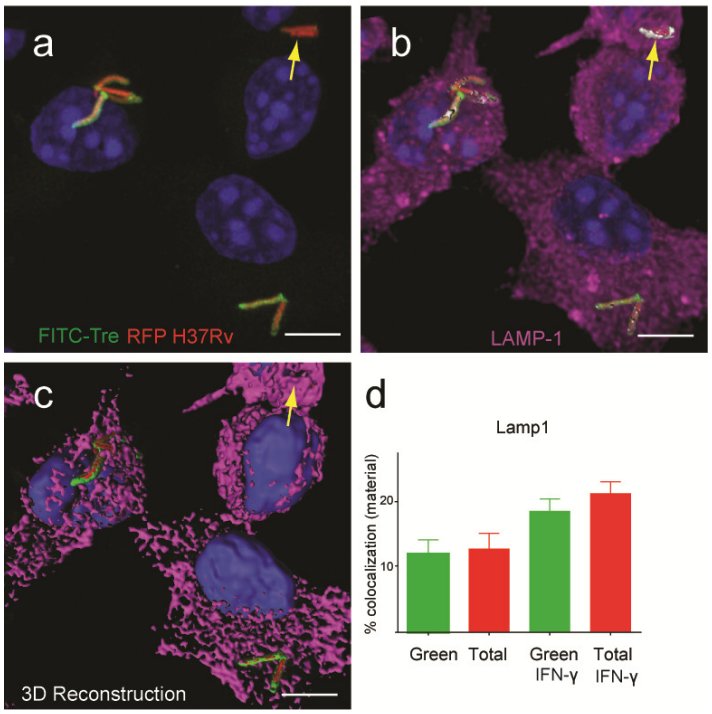


Figure 5.20. Illustrates triple labeling colocalization studies with RFP, FITC-Tre and anti-LAMP antibody in H37Rv-RFP infected bone marrow macrophages (BMMs) activated with IFN-γ. Arrows indicate red bacteria that are not labeled with FITC-Tre and that colocalize with LAMP1 (a) shows the merged RFP and FITC images, (b) the labeling with anti-LAMP1 (magenta) with colocalization indicated in white, while (c) is a three dimensional reconstruction. (d) shows quantitation of the colocalization between antibody signal and bacteria labeled with FITC-Tre (green bars) or red bacteria (total population). P-Values for anti-LAMP1: Green vs. Green+IFN-γ (0.006), Green vs. Total+IFN-γ (0.006), Total vs. Green+IFN-γ (0.01), Total vs. Total+ IFN-γ (0.001) and Green+IFN-γ vs. Total+ IFN-γ (0.1). Error bars reflect standard error of the mean (SEM) for a minimum of three experiments. P value > 0.1 are not indicated. Scale bars 5 µm.

5.2.11 IFN-γ vs. Non-INF-γ treated cells

A comparison between non IFN-γ and IFN-γ activated macrophages revealed a decrease in accumulation of FITC-Tre under activating conditions, consistent with the proposed model that increased trafficking leads to decreased accumulation of FITC-Tre. FITC-Tre bacteria, grown in both J774 and bone marrow macrophages ± IFN-γ showed decreased labeling under IFN-γ, particularly in the range of extremely bright mycobacteria (**Figure 5.21**). The difference was within the realm of error, due to the wide distribution of the labeling, but was suggestive of a decrease in label accumulation when macrophages were activated.

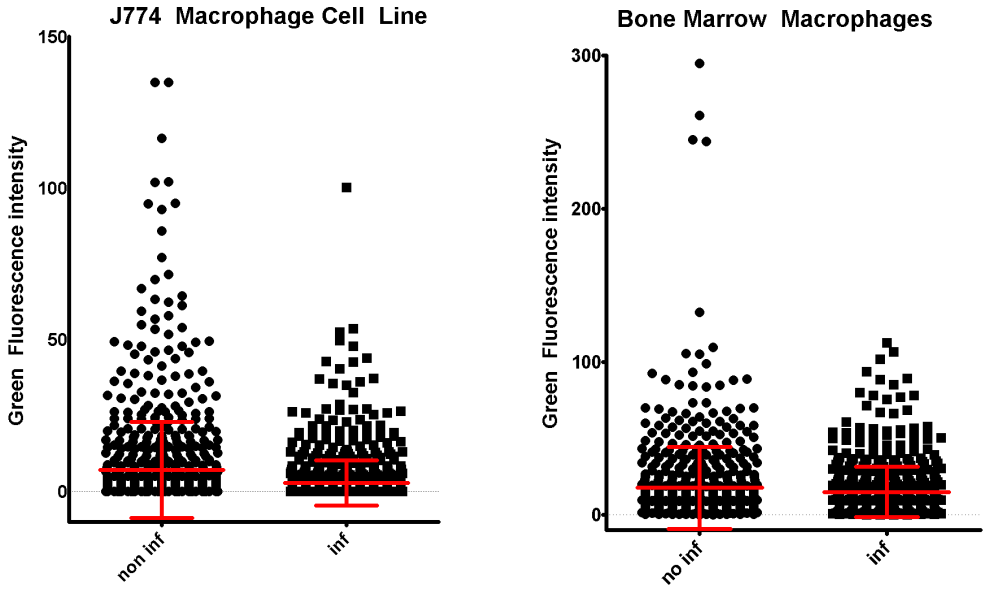


Figure 5.21. Mean green fluorescent intensity for (a) J774 infected macrophages and (b) infected BMMs. Bacteria were treated as ROIs and the mean fluorescence for each ROI was quantified and plotted. Error bars reflect the standard error of the mean.

The body of the colocalization analysis (**Figure 5.16 - Figure 5.20**) focused on colocalization of voxels within images, meaning reporting how often two channels existed within the same three dimensional space. To further this analysis, the mean fluorescence for antibody and FITC-Tre within a bacterial region of interest (ROI), as defined by areas that are labeled with RFP, was evaluated. IMARIS software can create ROIs of regions with a specific color and then export the statistic for these regions in our case mean fluorescence for FITC-Tre and antibody (633 nm). The median intensity was also calculated and comparable data was observed (not shown). These statistics were plotted as mean antibody fluorescence versus mean FITC-Tre (**Figure 5.22**) and the trends are comparable to those seen with our prior colocalization data.

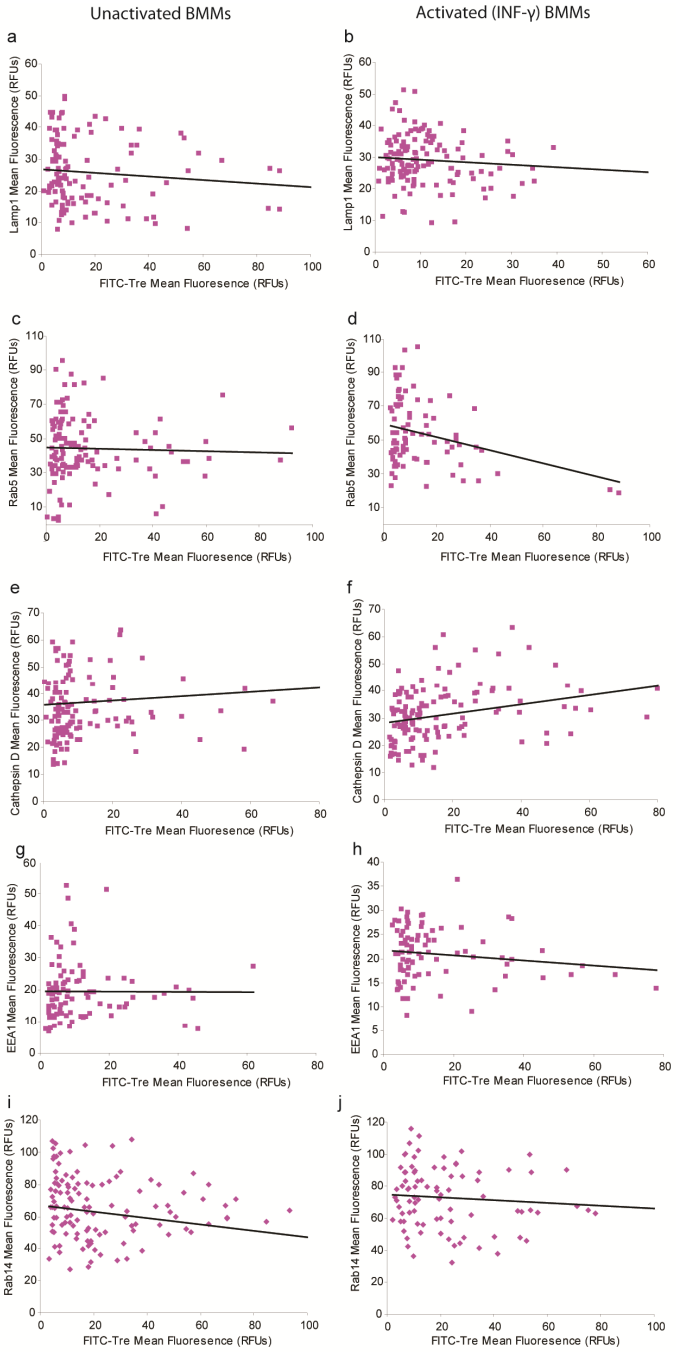


Figure 5.22. Mean antibody fluorescence and FITC-Tre fluorescence for regions of interest (ROIs) identified by RFP expressing *Mtb*. Each point reflects an individual ROI for red fluorescent protein (RFP) expressing H37Rv residing within unactivated or activated bone marrow Macrophages (BMMs). (a,b) LAMP1 (c,d) Rab5 (e,f) CathD (g,h) EEA1 (i,j) Rab14. RFUs- relative fluorescence units.

5.3 Ag85 antibodies show non-specific labeling

A conclusive link between increased FITC-Tre labeling and the Ag85 enzymes was sought through the use of Ag85 antibodies. A general polyclonal antibody against all three isoforms of Ag85 (Colorado State) and Anti-Ag85A and Anti-Ag85B monoclonal antibodies (Abcam) were obtained. Infection and labeling was conducted in J774 macrophages (**Figure 5.23**). Uninfected cells were also labeled as a control in order to evaluate how much background might be caused by non-specific binding to the cells.

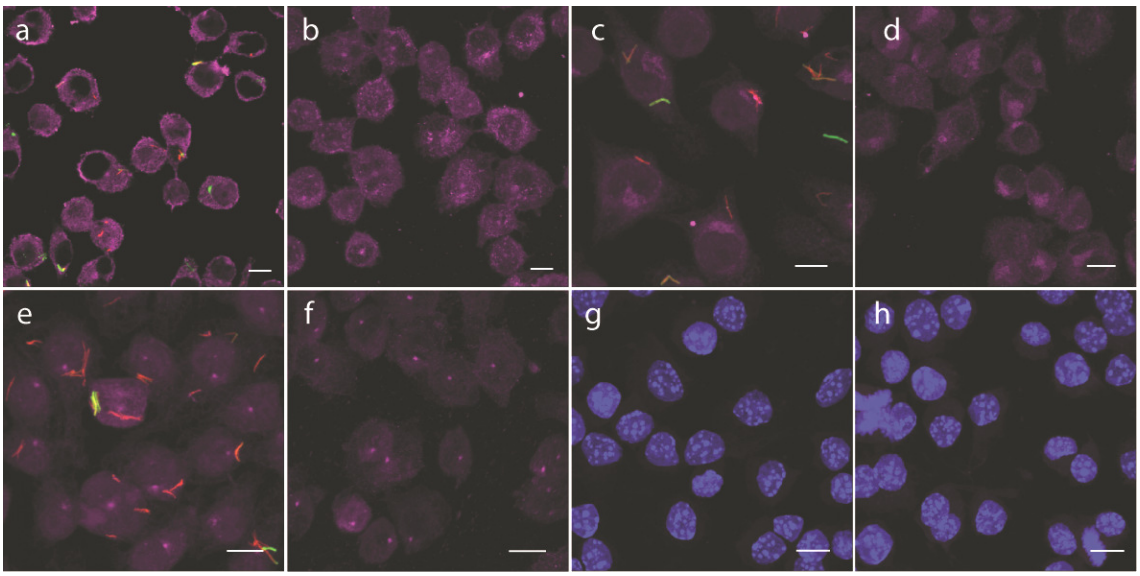


Figure 5.23. Ag85 antibodies fail to show utility in immunocytochemistry. (a) J774 macrophages infected with RFP H37Rv and labeled with FITC-Tre and Ag85 antibody (Colorado State). (b) Uninfected macrophages labeled with Ag85 antibody (Colorado State). (c) J774 macrophages infected with RFP H37Rv and labeled with FITC-Tre and Ag85A antibody (Abcam). (d) Uninfected macrophages labeled with Ag85A antibody. (e) J774 macrophages infected with RFP H37Rv and labeled with FITC-Tre and Ag85B antibody (Abcam). (f) Uninfected macrophages labeled with Ag85B antibody. (g-h) Secondary antibodies labeling J774 macrophages in the absence of primary antibodies shows low background. Magenta- Anti-Ag85 antibodies, Blue-Dapi, Red-RFP, Green-FITC-Tre. Scale bars 5 µm.

High background in both infected and uninfected macrophages and negligible colocalization of the antibodies with the *Mtb* was observed, suggesting that these antibodies were not compatible with immunocytochemistry. Slight accumulation of antibody in small vacuoles adjacent the nucleus was observed in several images from both macrophages infected and not infected with bacteria. To confirm that the high background was not due to the secondary antibodies, very low background was observed when the secondary antibodies were used to label cells directly without first staining with primary antibodies (**Figure 5.23 g,h**). Although

these results did not link FITC-Tre labeling to the Ag85 accumulation directly, they do reveal the utility of the FITC-Tre probe as an auxiliary for Ag85 activity, under experiments that are not compatible with the currently available antibodies.

5.3.1 Attempts at *in vivo* labeling

The FITC-Tre labeling methodology was attempted in a mouse model of *Mtb. Mtb* infected (3 weeks post aerosol infection) and uninfected mice were dosed with FITC-Tre through tail-vein injection to determine whether the compound would label *Mtb in vivo.* Such labeling would be useful in the development of novel TB-diagnostics and methods for monitoring the disease through a course of therapy. Mice were sacrificed and lungs, liver and spleen harvested and treated with standard flash freeze and mounting protocol (**Appendix 2**).

Labeling in the mouse model was inconclusive and complicated by a low infection level. No labeling was observed within liver or pancreatic tissues. Low level infection is characteristic of the pancreas and so some labeling was expected within pancreatic tissues. However, the bacterial load may have simply been too low to detect any bacteria. Little background fluorescence was observed in the lungs of the control mice (**Figure 5.24)** and in the low dose mice (data not shown). In contrast, the mouse treated with the double high dose demonstrated apparent fluorescence in the lung tissue (**Figure 5.24**). Conclusive proof that the fluorescently labeled objects were in fact *Mtb* was not available because the anti-TB antibody system labeled tissue very poorly and attempts to employ an auramine-rhodamine stain revealed cross-talk between the auramine signal and the FITC-Tre signal. The healthy appearance of the lung tissue slices and the lack of dense granuloma-like region was suggestive of a very low infection level. It was hypothesized that the lack of FITC-Tre labeling was either due to a paucity of bacteria, failure of the probe to access the lesions or potential degradation of the compound.

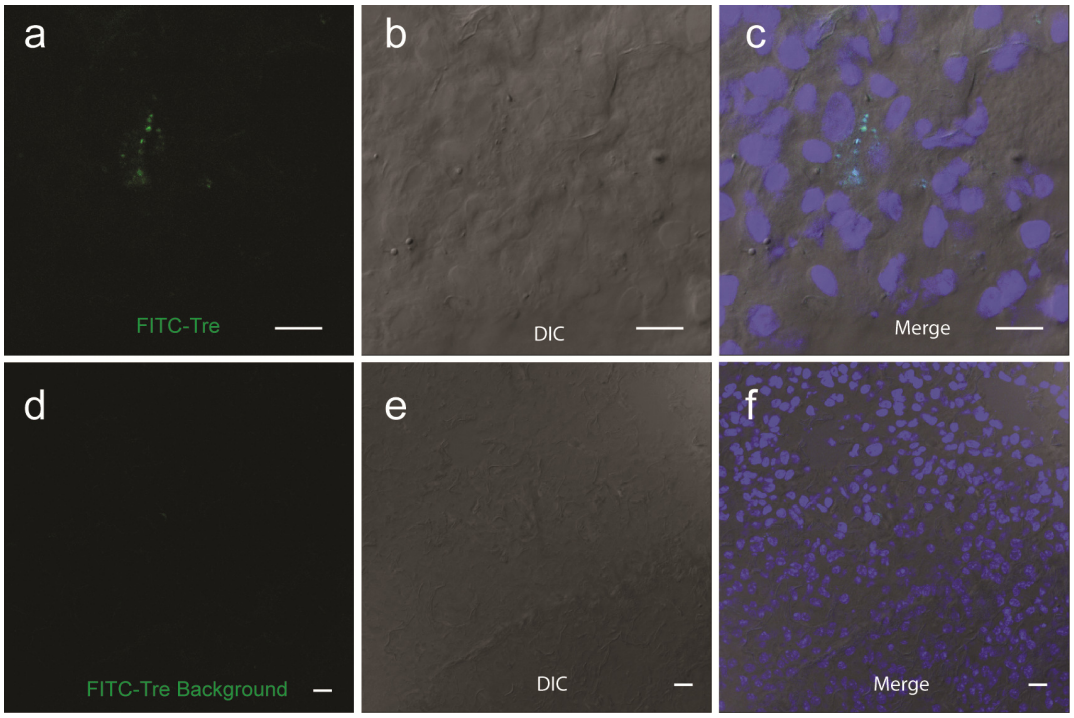


Figure 5.24. Potential *in vivo* labeling of an infected mouse (high dose FITC-Tre) 'a-c' or an uninfected mouse 'd-f' (a) FITC-Tre channel of lung infected with H37Rv (b) DIC image of 'a' (c) Merge of FITC-Tre and DIC image with DAPI staining the nuclei (d) FITC-Tre channel of healthy lung (e) DIC image of 'd' (f) Merge of FITC-Tre and DIC image with DAPI staining the nuclei. Scale bars 10 µm.

5.3.2 Attempts at tissue labeling

Current diagnosis of *Mtb* relies on primitive techniques, including chest X-rays and acid-fast stains of sputum. Acid fast stains are problematic because a significant portion of bacteria do not appear in the stain, so called non-acid-fast mycobacteria (NAF)[44]. Non-acid-fast cell wall deficient variants of *M. tuberculosis* have been found in the caseous lesions of tuberculosis patients. The low sensitivity of sputum microscopy and the low bacterial counts characteristic of *Mtb* infection have resulted in an infection detection rate of less than fifty percent[45]. Because of these problems with acid-fast stains, a better method for the detection *Mtb* in tissue was sought. It was postulated that the application of FITC-Tre to fresh, infected tissue, or application of FITC-Tre and Ag85 to fixed tissue, might allow for fluorescent labeling of the TDM within the tissue.

Tissue labeling was attempted both by simply adding FITC-Tre and by adding both Ag85C and FITC-Tre with the hypothesis that the endogenous Ag85 in fresh tissue could catalyze the transesterification of FITC-Tre, whereas exogenous Ag85 would need to be added to fixed tissue. Small, fluorescent labels were potentially observed in lung tissue, both with and

without Ag85. Also, large regions of fluorescent debris were observed, consistent with incomplete washout of the probe and adherence of fluorescein to tissue, particularly regions of partially dried tissue.

Putative granuloma regions were identified during microscopy of tissue labeling experiments to increase the chances to observe labeled bacteria. In the mouse tissue regions of extensive DAPI staining and DNA were identified as these are characteristic of mouse granulomas. By appearance of the tissue as well as by the auramine rhodamine stain, the mice were not heavily infected. For the rabbits, regions of low level of DAPI staining and an almost fuzzy appearance were selected because these are the hallmarks of rabbit granulomas. The rabbit tissue appeared to be more heavily infected than the mouse tissue. Some localized labeling of objects roughly the size of mycobacteria was observed. However, the anti-TB antibodies, which would have been the ideal colabel to pinpoint whether the observed labeling was legitimate, failed to label the tissue. Three different antibodies (Abcam Ab905-500, Axel Yv54601, and rabbit serum from an infected rabbit) were employed. All were primary antibodies, used with the xenon fragments as a secondary label. Further work in the labeling of tissue with FITC-Tre should employ more effective primary and secondary antibodies or RFP expressing strains as a secondary label.

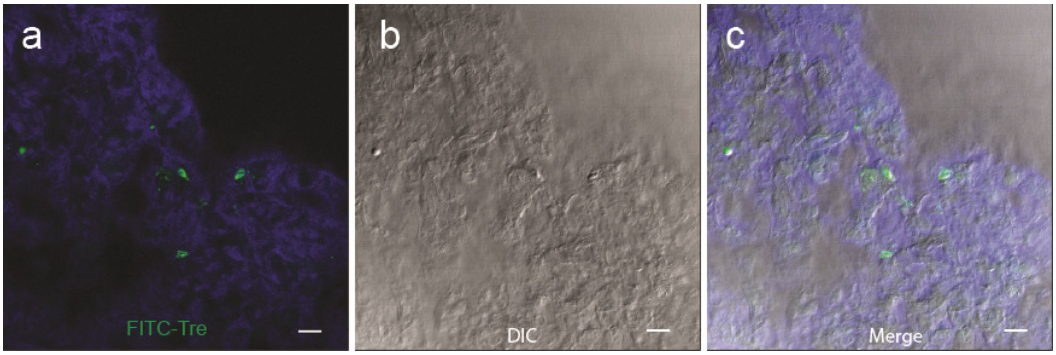


Figure 5.25. Rabbit lung treated with FITC-Tre and Ag85C overnight at room temperature. (a) FITC-Tre (green) and DAPI (blue) stain the infected rabbit lung slice (b) DIC image shows the edge of a putative rabbit granuloma and (c) shows an overlay of 'a-b.' Scale bars 10 µm.

5.3.3 ^{18}F-Trehalose and PET imaging of infected rabbits

FITC-Tre had been originally designed as a method to validate the development of trehalose based positron emission tomography (PET) probes for PET/CT imaging, and even with the inconclusive results from the accumulation of FITC-Tre in tissue, we turned our attention towards the synthesis and implementation of ^{18}F-trehalose as a PET probe. ^{18}FB-Trehalose (for synthesis see **Chapter 2**) was neutralized, filter sterilized and injected into a rabbit, 122 days post infection. The rabbit was imaged twice, 45 minutes and 2 hours post injection[iv]. After the first scan, no radiolabeling of lesions was detected. However, 2 hours post injection, lesions were detected as radiolabeled (**Figure 5.26**). The liver was highly radioactive with a standard uptake value (SUV) of about 13. Lesions only had a SUV of about 3-4. Although less than the liver, this SUV was still much higher than the background. These preliminary results suggest that trehalose probes might have the potential as PET probes. The continuation of this work has been conducted by others in the Barry lab and has featured scans of uninfected animals as control and the development of several other ^{18}F-trehalose analogues (2-F, 4-F, 6-F).

[iv] Work done with Dan Schimel, Danielle Wiener and Dr. Laura Via, NIAID, NIH

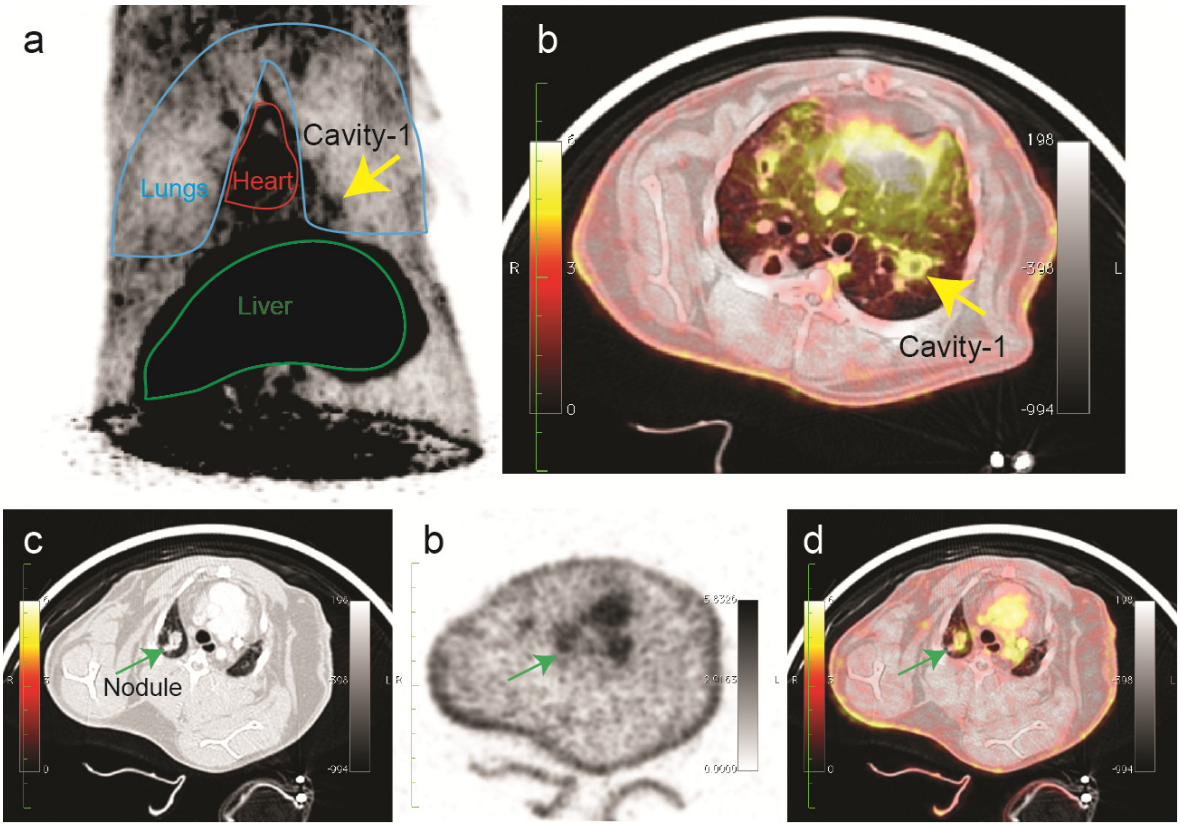


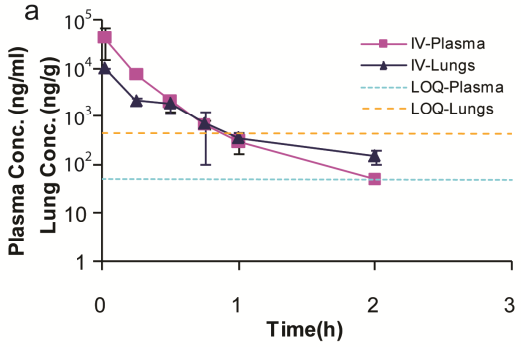
Figure 5.26. PET/CT of infected rabbit 122 hours post infection, 2 hours post injection. (a) 3D reconstruction of right torso of rabbit showing ^{18}F activity collected by PET; (b) A large cavity in lower left lung is labeled with ^{18}F-trehalose after 70 min uptake time images (c) Pet scan of nodule in upper right lobe shows activity or Standard Uptake Value (SUV) of 3.4 (axial view) (d) CT scan of image in 'c' (e) is an overlay of 'c-d.' Lung background has an SVU of 0.6. PET scale: SVU 0 (dark red)-6 (pale yellow). CT scale: Hounsfield units -1000 (black-air) - 0 (light grey-water)-+1000 (white-bone). Lesions are labeled and indicated by arrows. Image generated in Osiris.

Diffuse lesions appeared to label more easily, possibly due to probe tissue penetration. Likewise the multiple small nodules and diffuse infiltrates labeled extensively. The labeling was time dependent and occurred over a longer timeframe than previous results with ^{18}FDG, which supports the hypothesis of an Ag85 dependent labeling mechanism. Potentially, a longer residence time for the probe, to allow for greater clearance could lead to more labeling and a higher signal to noise ratio (S/N). Alternatively, switching to a longer half-life isotope, such as bromine or copper might allow for a longer residence time prior to imaging and more labeling.

5.3.4 Pharmacokinetics and pharmacodynamics (PK/PD)

To further determine whether ^{18}F-fluorobenzyl-trehalose (FB-Trehalose) offered favorable metabolic properties for mycobacterial PET imaging, pharmacokinetic and dynamic (PK/PD) analysis was conducted. In collaboration with Novartis Institute for Tropical Diseases (NITD) (Ann Goh, Veronique Dartois and Suresh Lakshminarayana) PK/PD was conducted on the FITC-Tre and FB-Tre. Compounds were dosed by saline IV injection and PK/PD was analyzed by mass spectrometry. This analysis revealed that the FB-Tre clears slowly from the lungs, whereas the FITC-Tre has a relatively quick clearance (**Figure 5.27**, **Figure 5.28**).

a



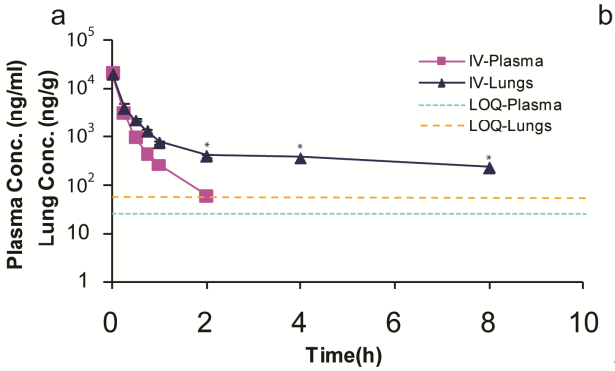
b

IVPK parameters	Plasma Values	Lung Values
V_{ss}	54.59 mL/kg	788.77 g/kg
CL	393.30 mL/h/kg	962.39 g/h/kg
$T_{1/2}$	0.16 h	0.46 h
Exposure (AUCinf)	7619.41 ng.h/mL	2993.50 ng.h/mL
Theoretical Dose	5.0 mg/kg	
True Dose	3 mg/kg	

Figure 5.27. (a) PK/PD profile of FITC-Tre reveals rapid clearance from plasma and lungs with an exposure ratio (lung to plasma): 0.4 (b) IVPK Plasma and lung parameters . V_{SS}- Apparent volume distribution, CL- Systemic clearance, $T_{1/2}$-halflife, AUCinf-Area under curve with last time point extrapolated.

These results may hint at why the PET studies with the FB-Tre indicated a lot of non-specific labeling and background radioactivity in the lungs. The FB-Tre had an exposure ratio of lung to plasma of 1.8 whereas the FITC-Tre has an exposure ratio of only 0.4. Part of the rapid clearance of the FITC-Tre could be due to gluconylation of the fluorescein, which has been well-documented previously[46]. Unfortunately, the mass for the gluconylation product was not tracked, so it is not yet certain whether this is the explanation for the rapid clearance of the FITC-Tre. These results suggest that neither of these compounds has the ideal PK/PD profile for a drug-like compound and, presumably, a PET Probe. However, one of the problems in focusing on PK/PD is that it is unclear what the ideal PK/PD profile is for the PET tracer. These PK/PD results do support that a more hydrophilic trehalose derivative (2-F trehalose in place of FB-Trehalose) may clear more rapidly from the lungs, resulting in reduced background radioactivity.

a



b

IVPK parameters	Plasma Values	Lung Values
V_{ss}	237.23 mL/kg	1100.47 g/kg
CL	1229.59 mL/h/kg	669.13 g/h/kg
$T_{1/2}$	0.40 h	0.31 h
Exposure (AUCinf)	3819.49 ng.h/mL	6920.14 ng.h/mL
Theoretical Dose	5.0 mg/kg	
True Dose	4.7 mg/kg	

Figure 5.28. (a) PK/PD profile of FB-Tre reveals clearance from plasma and adherence in lungs with an exposure ratio (lung to plasma): 1.8. Starred (*) data points were excluded in PK parameter regression. (b) IVPK plasma and lung parameters for FB-Tre. V_{SS}- Apparent volume distribution, CL- Systemic clearance, $T_{1/2}$-halflife, AUCinf-Area under curve with last concentration extrapolated.

5.4 Conclusions

FITC-Tre is the first small molecule, mycobacteria selective probe that will label *Mtb* within macrophages. Unlike other possible probes, the FITC-Tre mode of action is based on the activity of an enzyme unique to this genus of organisms, providing a mechanism based readout of Ag85 activity *in vivo*. Our understanding of the macrophage infection process is incomplete[47], so the ability to follow viable *Mtb* during infection shown here, could prove extremely useful as a means to further understand the bacterial transit to the phagosome as well as other intracellular compartments. The colocalization studies between FITC-Tre-labeled *Mtb* and various markers of phagosome maturation suggest that FITC-Tre may preferentially label those bacteria that resist acidification and phagosome-lysosome fusion. An alternative hypothesis that cannot be eliminated is that incorporation of FITC-Tre into the membrane of a subset of bacilli alters the course of development of their phagosome, as has been seen with *Brucella pertusis* prelabeled with FITC before infection[48]. In our experiments FITC-Tre was added after the infection has been established, and is in low concentration relative to TDM. Therefore, we expect the influence of the FITC-Tre-DM to be slight. Preliminary results in the development of FITC-Tre and FB-Tre for *in vivo* and in tissue labeling are suggestive, but will require additional work to fully develop. Other trehalose-tagged molecules, potentially labeled at different positions or different tags on the trehalose scaffold, may also prove useful as additional *Mtb* probes. PET scans with the first ^{18}F-trehalose derivative (FB-Trehalose) suggest that trehalose PET probes may provide a method for TB-selective PET imaging. In summary fluorescent and radioactive trehalose show utility for *in vivo* labeling of *M. tuberculosis.*

5.5 References

1. Beatty, W.L. et al. Trafficking and release of mycobacterial lipids from infected macrophages. *Traffic* **1**, 235-47 (2000).
2. Thanky, N.R., Young, D.B. & Robertson, B.D. Unusual features of the cell cycle in mycobacteria: polar-restricted growth and the snapping-model of cell division. *Tuberculosis (Edinb.)* **87**, 231-6 (2007).
3. Provvedi, R., Boldrin, F., Falciani, F., Palu, G. & Manganelli, R. Global transcriptional response to vancomycin in *Mycobacterium tuberculosis*. *Microbiology* **155**, 1093-1102 (2009).
4. Glatman-Freedman, A., Martin, J.M., Riska, P.F., Bloom, B.R. & Casadevall, A. Monoclonal antibodies to surface antigens of *Mycobacterium tuberculosis* and their use in a modified enzyme-linked immunosorbent spot assay for detection of mycobacteria. *J. Clin. Microbiol.* **34**, 2795-2802 (1996).
5. Nordin, A., Rossetti, C. & Rahim, N. Disseminated tuberculosis infection: a 'super' 18F-FDG PET/CT appearance. *Eur. J. Nucl. Med. Mol. Imaging* **36**, 882-882 (2009).
6. Govindarajan, M., Nagaraj, Ravikumar, Kalyanpur, A. & Kallur. Tuberculosis - The spoil sport. *Indian J Radiol Imaging* **18**, 265-6 (2008).
7. Tufariello, J.M., Chan, J. & Flynn, J.L. Latent tuberculosis: mechanisms of host and bacillus that contribute to persistent infection. *The Lancet Infectious Diseases* **3**, 578-590 (2003).
8. Siddiqui, A.H., Perl, T.M., Conlon, M., Donegan, N. & Roghmann, M.C. Preventing nosocomial transmission of pulmonary tuberculosis: when may isolation be discontinued for patients with suspected tuberculosis? *Infect. Control Hosp. Epidemiol.* **23**, 141-4 (2002).
9. Grange, J.M. The humoral immune response in tuberculosis: its nature, biological role and diagnostic usefulness. *Adv. Tuberc. Res.* **21**, 1-78 (1984).
10. Abadie, V. et al. Neutrophils rapidly migrate via lymphatics after *Mycobacterium bovis* BCG intradermal vaccination and shuttle live bacilli to the draining lymph nodes. *Blood* **106**, 1843-1850 (2005).
11. Darrow, M.A. Staining the tubercle organism in sputum smears. *Stain Technol.* **24**, 93 (1949).
12. Abbe, E. Beitrage zur Theorie des Mikroskops und der Mikroskopischen Wahrnehmung. *Mikrosk. Anat.* **9**, 55 (1873).
13. Neelsen, F. Ein casuistischer Beitrag zur Lehre von der Tuberkulose. *Centrabl. Med. Wissenschaften* **28**, 497-501 (1883).
14. Sathyamoorthy, N. & Takayama, K. Purification and characterization of a novel mycolic acid exchange enzyme from *Mycobacterium smegmatis*. *J. Biol. Chem.* **262**, 13417-23 (1987).
15. Spargo, B.J., Crowe, L.M., Ioneda, T., Beaman, B.L. & Crowe, J.H. Cord factor (alpha,alpha-trehalose 6,6'-dimycolate) inhibits fusion between phospholipid vesicles. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 737-740 (1991).
16. Indrigo, J., Hunter, R.L., Jr. & Actor, J.K. Cord factor trehalose 6,6'-dimycolate (TDM) mediates trafficking events during mycobacterial infection of murine macrophages. *Microbiology* **149**, 2049-59 (2003).
17. Matz, M.V. et al. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat Biotech* **17**, 969-973 (1999).
18. Sacchetti, A., Subramaniam, V., Jovin, T.M. & Alberti, S. Oligomerization of DsRed is required for the generation of a functional red fluorescent chromophore. *FEBS Lett.* **525**, 13-19 (2002).
19. Fratti, R.A., Backer, J.M., Gruenberg, J., Corvera, S. & Deretic, V. Role of phosphatidylinositol 3-kinase and Rab5 effectors in phagosomal biogenesis and mycobacterial phagosome maturation arrest. *J. Cell Biol.* **154**, 631-644 (2001).
20. van der Wel, N. et al. *M. tuberculosis* and *M. leprae* translocate from the phagolysosome to the cytosol in myeloid cells. *Cell* **129**, 1287-98 (2007).
21. Cougoule, C., Constant, P., Etienne, G., Daffé, M. & Maridonneau-Parini, I. Lack of Fusion of Azurophil Granules with Phagosomes during Phagocytosis of *Mycobacterium smegmatis* by Human Neutrophils Is Not Actively Controlled by the Bacterium. *Infect. Immun.* **70**, 1591-1598 (2002).
22. Soper, H.E., Young, A.W., Cave, B.M., Lee, A. & Pearson, K. On the distribution of the correlation coefficient in small samples. Appendix II to the papers of a "Student" and R. A. Fisher. A cooperative study. *Biometrika* **11**, 328-413 (1917).
23. Jutras, I. et al. Modulation of the Phagosome Proteome by Interferon-γ. *Mol. Cell. Proteomics* **7**, 697-715 (2008).
24. Gorvel, J.P., Chavrier, P., Zerial, M. & Gruenberg, J. rab5 controls early endosome fusion in vitro. *Cell* **64**, 915-25 (1991).

25. Via, L.E. et al. Arrest of Mycobacterial Phagosome Maturation Is Caused by a Block in Vesicle Fusion between Stages Controlled by rab5 and rab7. *J. Biol. Chem.* **272**, 13326-13331 (1997).
26. Kyei, G.B. et al. Rab14 is critical for maintenance of *Mycobacterium tuberculosis* phagosome maturation arrest. *EMBO J.* **25**, 5250-5259 (2006).
27. Sturgill-Koszycki, S. et al. Lack of acidification in Mycobacterium phagosomes produced by exclusion of the vesicular proton-ATPase. *Science* **263**, 678-81 (1994).
28. Huynh, K.K. et al. LAMP proteins are required for fusion of lysosomes with phagosomes. *EMBO J.* **26**, 313-324 (2007).
29. Ullrich, H.J., Beatty, W.L. & Russell, D.G. Direct delivery of procathepsin D to phagosomes: implications for phagosome biogenesis and parasitism by Mycobacterium. *Eur. J. Cell Biol.* **78**, 739-48 (1999).
30. Martinez, O. & Goud, B. Rab proteins. *Biochim. Biophys. Acta* **1404**, 101-12 (1998).
31. Zerial, M. & McBride, H. Rab proteins as membrane organizers. *Nat. Rev. Mol. Cell Biol.* **2**, 107-17 (2001).
32. Duclos, S. et al. Rab5 regulates the kiss and run fusion between phagosomes and endosomes and the acquisition of phagosome leishmanicidal properties in RAW 264.7 macrophages. *J. Cell Sci.* **113**, 3531-3541 (2000).
33. Kaufmann, S.H.E. How can immunology contribute to the control of tuberculosis? *Nat. Rev. Immunol.* **1**, 20-30 (2001).
34. Smith, A.C. et al. A network of Rab GTPases controls phagosome maturation and is modulated by *Salmonella enterica* serovar Typhimurium. *J. Cell Biol.* **176**, 263-268 (2007).
35. Barker, L.P., George, K.M., Falkow, S. & Small, P.L. Differential trafficking of live and dead *Mycobacterium marinum* organisms in macrophages. *Infect. Immun.* **65**, 1497-1504 (1997).
36. Celada, A., Gray, P.W., Rinderknecht, E. & Schreiber, R.D. Evidence for a gamma-interferon receptor that regulates macrophage tumoricidal activity. *J. Exp. Med.* **160**, 55-74 (1984).
37. Zhang, X., Goncalves, R. & Mosser, D.M. The isolation and characterization of murine macrophages. *Curr Protoc Immunol* **Chapter 14**, Unit 14 1 (2008).
38. Landmann, L. Deconvolution improves colocalization analysis of multiple fluorochromes in 3D confocal data sets more than filtering techniques. *J. Microsc.* **208**, 134-47 (2002).
39. Nyquist, H. Certain Topics in Telegraph Transmission Theory. *American Institute of Electrical Engineers, Transactions of the* **47**, 617-644 (1928).
40. Shannon, C.E. Communication in the Presence of Noise. *Proceedings of the IRE* **37**, 10-21 (1949).
41. Kelley, V.A. & Schorey, J.S. Mycobacterium's Arrest of Phagosome Maturation in Macrophages Requires Rab5 Activity and Accessibility to Iron. *Mol. Biol. Cell* **14**, 3366-3377 (2003).
42. Alvarez-Dominguez, C. & Stahl, P.D. Increased expression of Rab5a correlates directly with accelerated maturation of *Listeria monocytogenes* phagosomes. *J. Biol. Chem.* **274**, 11459-62 (1999).
43. Indrigo, J., Hunter, R.L., Jr & Actor, J.K. Cord factor trehalose 6,6'-dimycolate (TDM) mediates trafficking events during mycobacterial infection of murine macrophages. *Microbiology* **149**, 2049-2059 (2003).
44. Bishai, W.R. Rekindling old controversy on elusive lair of latent tuberculosis. *The Lancet* **356**, 2113-2114 (2000).
45. Perkins, M.D., Roscigno, G. & Zumla, A. Progress towards improved tuberculosis diagnostics for developing countries. *The Lancet* **367**, 942-943 (2006).
46. Webb, J.M., Fonda, M. & Brouwer, E.A. Metabolism and excretion patterns of fluorescein and certain halogenated fluorescein dyes in rats. *J. Pharmacol. Exp. Ther.* **137**, 141-147 (1962).
47. Schorey, J.S., Carroll, M.C. & Brown, E.J. A Macrophage Invasion Mechanism of Pathogenic Mycobacteria. *Science* **277**, 1091-1093 (1997).
48. Weingart, C.L. et al. Fluorescent Labels Influence Phagocytosis of *Bordetella pertussis* by Human Neutrophils. *Infect. Immun.* **67**, 4264-4267 (1999).

6 Chapter 6: Differential roles of Ag85s and related enzymes

Chapter 6 dissects the functional roles of the Ag85 enzymes both through enzyme assays and via whole cell *Mtb* lipid profiling and microscopy. Overexpression of a panel of Ag85B and C mutants revealed a secondary site that regulates catalytic activity and dimeric or monomeric nature of the proteins. Mutation of the secondary site Ag85C residues into the Ag85B protein construct afforded a gain of catalytic activity on par with Ag85C function. Conversely, mutation of the 'B' residues into the Ag85C caused nearly a twenty fold loss of activity. Molecular dynamics (MD) simulations confirm that this site helps to regulate catalytic activity. Overexpression of the Ag85 proteins as tetracycline inducible, GFP fusion proteins shows that these enzymes can be used to modulate the incorporation of mycolates into TDM, TMM and the cell wall in an isoform dependent manner. The sum of this data suggests that the Ag85 enzymes serve unique purposes within the mycobactia

Linked to the development of trehalose probes based on the Ag85 activity (**Chapters 2-5**), a better understanding of the scope of Ag85 enzymes catalytic and immunological purposes will facilitate the production of improved trehalose-based mycobacterial labelling. Because the Ag85 enzymes are important immune epitopes and proposed drug targets and are implicated in the synthesis of the mycobacterial cell wall, a more complete analysis of these enzymes may further our ability to diagnose and treat *Mtb.*

6.1 Introduction

Described in the previous chapters as a means to artificially coopt the native bacterial cell wall biosynthesis labelling mycobacteria by fluorescent and radioactive probes, the natural functions of the Antigen 85 (Ag85) enzymes are still poorly understood. The Ag85 enzymes were discovered in 1987[1], as a novel mycolyl transferase enzyme from *M. smegmatis*. These enzymes have been identified as the serine esterases responsible for the transesterification of trehalose monomycolate (TMM) to trehalose dimycolate (TDM) and the attachment of mycolates from TMM to the arabinogalactan. Three isoforms of Ag85 (A,B and C) have been identified in *Mtb*, although additional isoforms of Ag85 are found in related organisms, *M. smegmatis* and *C. glutamicum*. An additional, inactive isoform, Ag85D (serine mutation) is also present in the genome, but is thought to, perhaps, be a genomic artifact or immune decoy. These proteins have been crystallized and mutants of each isoform of the enzymes have been raised[2,3]. However, due to the seeming 'redundancy' or ability to complement each

other, these individual mutants failed to fully dissect the functional roles of Ag85 in the mycobacterial cell wall biosynthesis.

The state of the current literature on the Ag85 proteins proposes a basic framework of how these proteins function, as serine esterases, transferring mycolates to additional units of trehalose and the cell wall. They are also thought to have fibronectin binding capacity, which may be important for infection and persistence within the host, although this has not been fully documented[4]. The Ag85 proteins are the largest protein component of mycobacterial culture supernantant[5]. They are trafficked in host cells and are recognized by the immune system[6]. To date, the dogma surrounding the Ag85 enzymes suggests functional redundancy, specifically that Ag85A, B and C serve more or less the same function in cell wall biosynthesis and that expression of different isoforms under different environmental conditions allows the mycobacteria to evade a host immune response towards these proteins. This model is incomplete, as it does not explain the findings that T cell epitopes are actually hyperconserved by mycobacteria[7], implying that the bacteria benefit from provoking an immune response. Related organisms, such as the soil bacteria, *M. smegmatis* and *C. glutamicum*, encode for six Ag85 proteins[8]. The minimalist organism, *M. leprae,* produces all three isoforms of Ag85[9]. Reductive evolution of *M. leprae* is thought to have eliminated most non-essential genes from the genome[10]. All three of the Ag85 enzymes retain their place in the *M. leprae* genome, implying a non-redundant functionality of each of these enzymes within bacterial survival and proliferation.

This body of evidence poses the question: If the Ag85 enzymes are not functionally redundant, then what unique purposes do they serve? A hypothesis that emerges from the Ag85 enzyme literature is: *Mtb* would not produce such vast quantities of three closely related proteins unless these proteins served some distinctive and important purposes for the survival and persistence of the organism within its human host.

6.2 Results

6.2.1 Identification and analysis of key residues in Ag85 secondary site loop

Although the Ag85 enzymes have a high degree of identity (70-80% sequence identity) with high identity in the active sites, the catalytic efficiencies vary dramatically between enzyme isoforms and to date no publications have explained, why this relative difference in activity is present. A comparison of the sequences of Ag85A, B and C identified key residues that

varied between isoforms and were conserved between various strains of mycobacteria (**Figure 6.1**). Several of these residues were located not in the active site, as had been expected, but instead in a secondary site. In previous reports this secondary site cocrystallized with a second trehalose unit and straightened when Ag85C was crystallized bound to a diethylphosphonate covalent inhibitor and has been annotated as a membrane binding site (**Figure 6.1**)[11,12]. Sequence analysis of this secondary site revealed greater conservation of residues between Ag85A and B as compared with Ag85A and C or B and C. One of the unresolved Ag85 dilemmas was the dramatic difference in catalytic activities between the three isoforms. It was found (**Chapter 4**) that Ag85C was 20-50 fold more active than either Ag85A or B with B the least active of the three. Because the Ag85C sequence diverged from Ag85A/B at this secondary site, it was hypothesized that some of these residues may be important for catalytic activity.

According to Ronning et al[11] a significant straightening movement of helix 8 (a component of the secondary site) was observed when the enzyme was crystallized with a diethylphosphate covalent inhibitor. This movement also causes the catalytic aspartate to flip out of the active site into a potential inactive conformation (**Figure 6.1**). There is only one reported structure (1DQY) at which this movement is observed; as such it was assigned as a putative transition state. No further analysis was conducted in the original publication to understand the implications of this movement. This movement may be caused by residues in the secondary binding site, which is located adjacent to this helix. In fact, the publication on the Ag85B crystal structure suggested that residues in this secondary site, particularly Phe230 were responsible for transferring the TMM into the active site via flexibility at the secondary site but the authors did not follow up on this postulate. Many of the Ag85 structures are uploaded to the Protein Data Bank (PDB) as dimmers (1DQZ, 1F0N,1F0P, 1SFR), where this same helix abuts the other half of the dimer, with the dimeric face occurring at this putative secondary site. One can envision a mechanism, whereby the movement of helix 8 (1DQY), forces the 2nd half of the dimer to rotate, shuffling substrates through the active site. Alternatively, the dimerization allows for binding to the divalent arabinan headgroups. However, beyond these crystollagraphic clues, no further literature linked this secondary site to regulation of Ag85 activity.

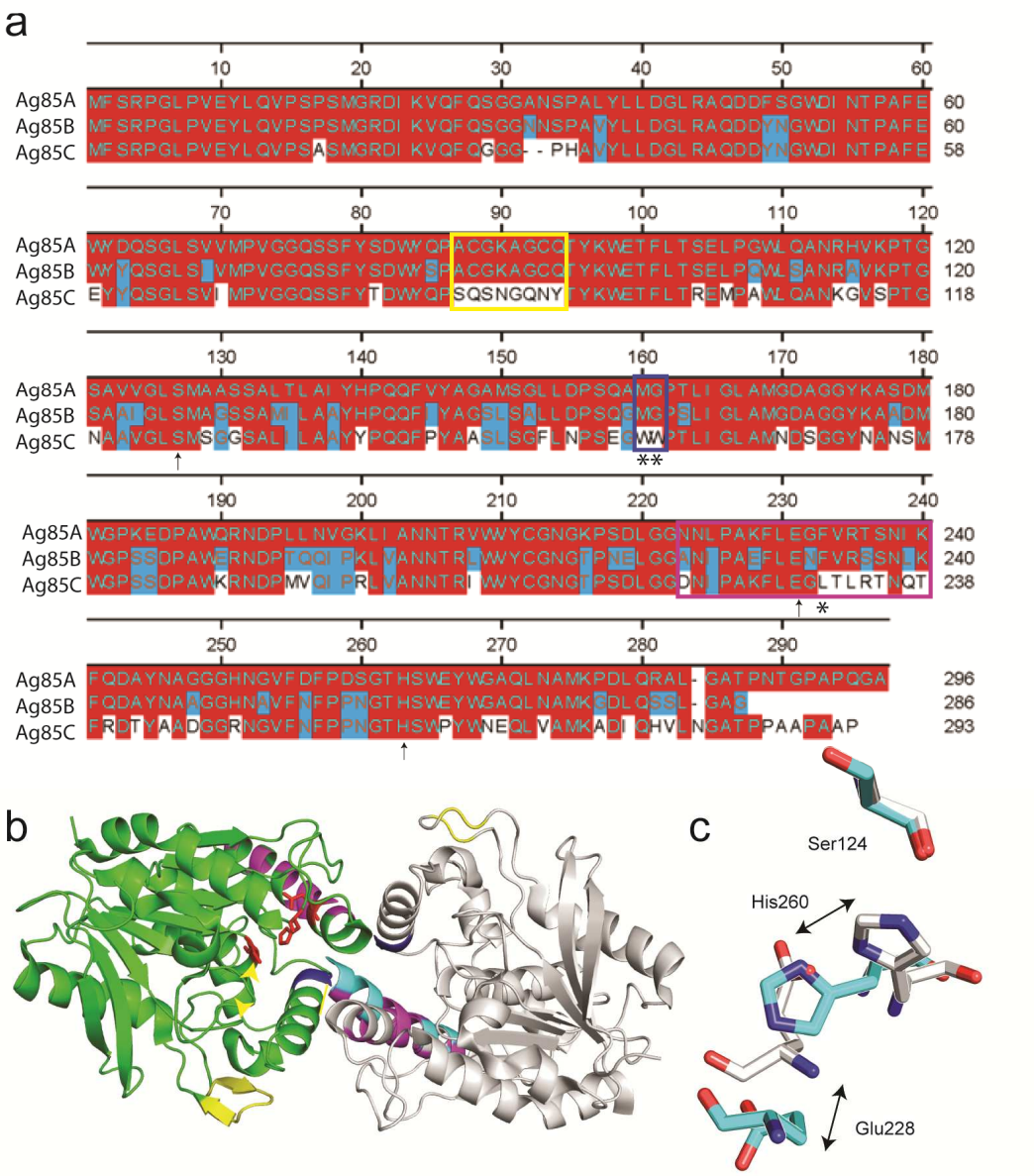


Figure 6.1. (a) Key residues (boxed) that are not conserved between Ag85C and Ag85A, are all located at the putative dimer or membrane binding interface. Residues highlighted in red match the Ag85A sequence, whereas residues in blue show residues that match Ag85B and Ag85C, but differ from Ag85A. Residues in white are Ag85C residues that differ from both Ag85A and B. Box colors correspond to the portions of Ag85 highlighted in 'b.' Trp157, Trp158 and Leu230 are indicated by stars. Arrows indicate catalytic triad. Residue numbers are for the truncated proteins expressed without the signal peptides. (b) Ag85C (1DQZ)- white , Ag85B (1F0P)-green Ag85C bound to diethylphosphate-cyan (1DQY) Non-conserved residues are highlighted in purple, blue and yellow with colors corresponding to the colored boxes in 'a'. The catalytic triad is highlighted in red (c) depicts the movement of the catalytic triad that occurs when Ag85C is acylated with diethylphosphate inhibitor.

A comparison of the sequences of *M. avium*, *M. smegmatis*, *M. avium* Paratuberculosis, *M. marinum*, *M. ulcerans* and *M. tuberculosis* H37Rv, for Antigen 85A, B and C indicated several residues, both on the helix 8, as well as on the helix in the dimer portion that is adjacent to helix 8 (see **Appendix 2** for sequence comparison). These residues (**Figure 6.1**) appeared directly in the secondary binding site and were well conserved. Work in Corynebacteria provided some of the motivation for this sequence analysis. Originally it was believed that Corynebacteria expressed only one mycolyl transferase, PS1. Subsequently, 5 additional mycolyl transferases were identified in Corynebacteria. These were found to fall into two classes, which can only partially complement for one another, hinting at structural differences that might relate to complementation[8].

Many of the residues that differed between Ag85A, B and C were actually functionally similar, eg serine swapped for threonine or lysine for arginine. Three residues that were conserved in all Ag85A and B proteins and that differed in both residue identity, hydrophobicity and size, from the conserved Ag85C residues were selected for prelimary studies: In Ag85C residues tryptophans (Trp) 157 and 158 are replaced by a methionine (Met) 159 and a glycine (Gly) 160 in Ag85A and B. Replacement of two adjacent Trp residues with Met and Gly is a relatively large, non-conserved substitution, which suggested that these residues might be functionally important. On the adjacent helix, flanking the two Trp, the Ag85C leucine 230 is replaced with phenylalanine 232 in both Ag85A and B. This amino acid change is not nearly as substantial as the tryptophans, but the positioning of this residue across from both Trp Trp residues suggested that interactions between these three amino acids could prove important. Asparagine 230 in Ag85B was also suggestive, because in Ag85A and C this residue is a glycine. These four non-conserved residues, not found in the active site, were hypothesized to explain the difference in Ag85 catalytic activity.

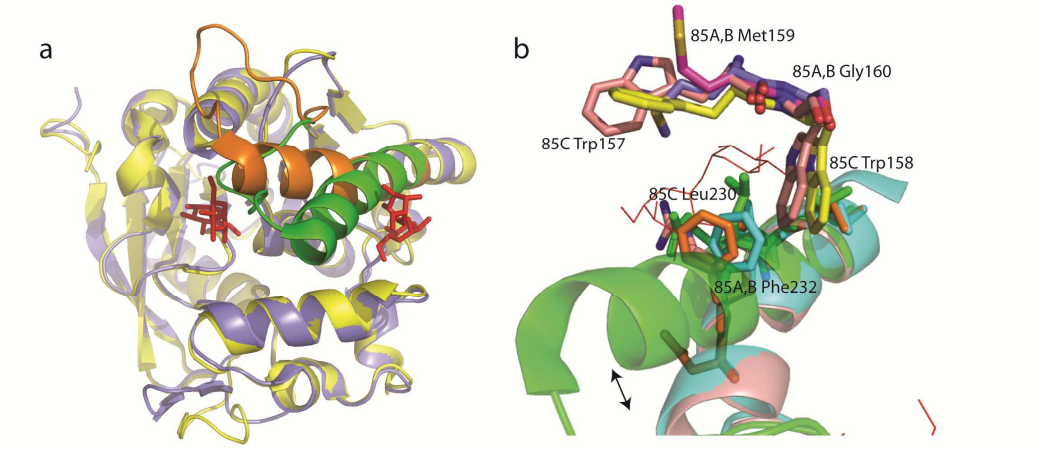


Figure 6.2. (a) Overlay of crystal structures of Ag85B and C, bound to diethyl phosphate (green, yellow) and bound to trehalose (orange and blue) shows the movement of the loop identified to be important in the secondary binding pocket interacting the two units of trehalose (red). (b) Considerable flexibility in secondary binding site. Tan-Ag85C without inhibitor. Green-with inhibitor. Blue-Ag85A, Orange-Ag85B. Arrow indicates direction of motion of Trp residues encountered during molecular dynamics simulations. Trp207 passes outside Leu282 in simulation. Notably Ag85A and B lack these two Trp residues.

6.2.2 Molecular Dynamics[i]

The importance of Trp157, Trp158 and Leu230 (Ag85C residues detailed above) were investigated first by molecular dynamics simulations and then by site directed mutagenesis and enzyme assays. Initial 20 nanosecond MD simulations were conducted on the Ag85C and B structures PDB 1DQZ and 1F0P. In these simulations considerable flexibility was observed in helix 8, consistent with the movement in the crystal structure. In simulations of both Ag85B and Ag85C bending and straightening of helix 8 was observed to be linked to Leu230 and Trp158 flipping or acting as a 'switch,' where these two hydrophobic residues were alternating positions in the secondary site pocket and their switching either straightened or flexed helix 8 (yellow) (**Figure 6.3**). Also consistent with the crystal structure, the active site glutamate flipped away from the histidine as the helix straightened, presumably temporarily inactivating the enzyme (**Figure 6.1**). Comparable to Ag85C simulations, in the Ag85B dynamics the corresponding residues, Phe231 and a neighboring, conserved residue Leu163 were also switching and helix 8 was bending and straightening in tandem with the switch. These simulations substantiated the hypothesis that Trp158 and Leu230 and their corresponding Ag85A/B residues regulated the active site and catalytic activity. The

[i] All Md simulations were conducted with Dr. Michael Dolan (NIBIB).

simulations also revealed a potential explanation for the difference in activity between the Ag85A, B and C enzymes. The Trp Trp Leu switch in Ag85C occupied more space and appeared to switch more slowly, resulting in a greater percentage of enzyme in the putative 'active' conformation. In contrast the Ag85AB Phe, Leu switch was slightly less hindered and the altered position of the Leu residue facilitated more rapid switching and a greater proporation of enzymes existing in a seemingly less productive conformation.

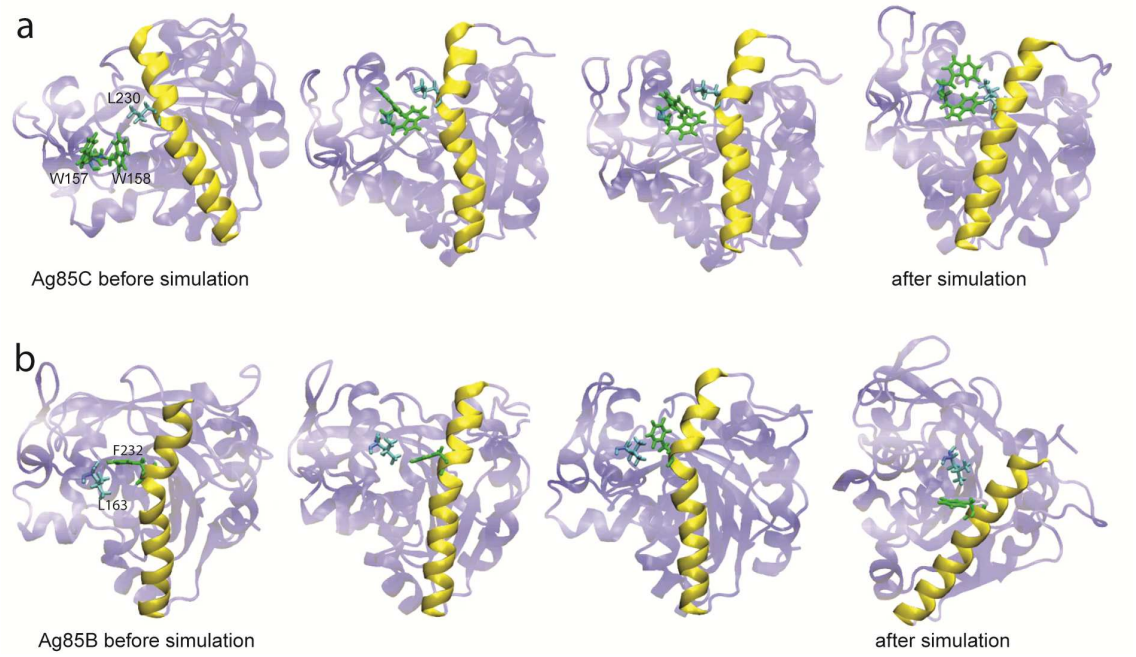


Figure 6.3. Molecular dynamics simulations confirm that helix 8 (yellow) bending and straightening is linked to three key 'switch' residues. (a) Ag85C (1DQZ) with helix 8 highlighted in yellow, Trp157 and Trp158 in green and Leu230 in blue (b) Ag85B (1f0n), Phe232 in green and Leu163 in blue.

Throughout the simulations, the other residue of interest, Trp 157, projected outwards into the buffer, an unusual position for the hydrophobic tryptophan residue. The Ag85 enzymes are actually annotated as a membrane binding protein by the University of Michigan Orientation of Proteins in Membranes (OPM) database[13]. The OMP database protein model proposes that helix 8 (magenta) (Figure 6.4) should sit at the interface between membrane and protein. The presence of solvent, exposed, hydrophobic tryptophan residues on this helix suggests that the Ag85 proteins should either be associated with another protein or with a membrane or other hydphobic surface to protect these residues from the solvent.

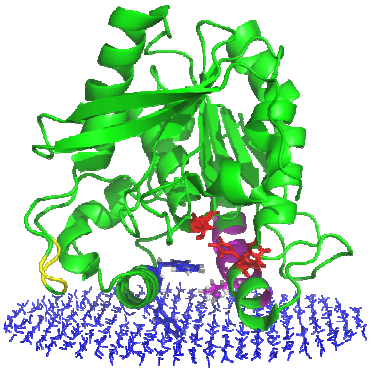


Figure 6.4. Model of Ag85C binding to a membrane (blue-water molecules) as annoted by the University of Michigan Orientation of Proteins in Membanes (OPM) database[13]. Yellow, blue and pink indicate significant non-conserved residues. Catalytic triad are indicated in red.

Molecular dynamics simulations run with Ag85C interacting with a simulated phospholipid bilayer (**Figure 6.5**) suggested that Ag85C, particularly Trp 157, interacts favourably with lipid membranes. In these simulations the Ag85C enzyme was oriented with the helix 8 face towards the bilayer to minimize simulation time, and then the interactions with the charged, phospholipid headgroups were observed over a 50 nanosecond simulation. Although not truly representative of the mycobacterial membrane, a phosphotidyl choline phospholipid bilayer was employed for ease of simulation. During this simulation Trp157 was observed to interact with the membrane, pointing away from the enzyme and down into the bylayer, an action that appears to anchor the enzyme to the membrane. Other charged residues, including Arg233 were also observed interacting with the membrane. This simulation suggested that this secondary site may serve a vital role in anchoring the Ag85C to the cell wall, allowing for transmycolation of the arabinogalactan to occur. Potentially, the replacement of Trp157 with Met159 in Ag85AB may not anchor the enzymes with the same efficiency, thus reducing the transmycolation efficiency.

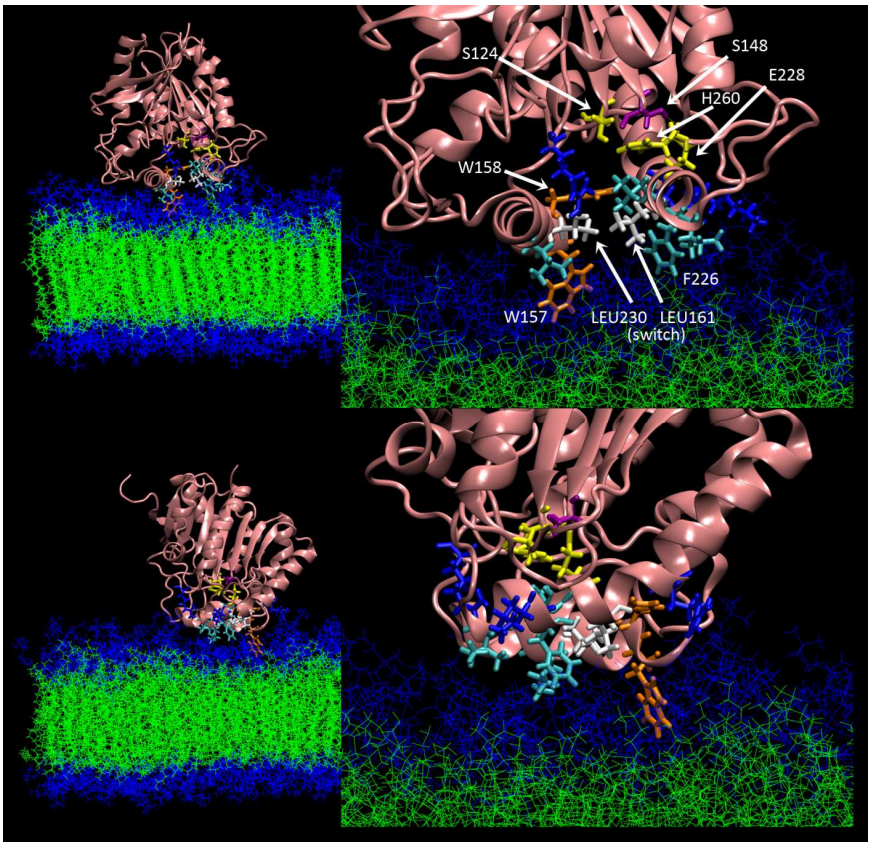


Figure 6.5. Molecular dynamics simulations of Ag85C interacting with a phospholipid bilayer reveals membrane binding interactions with Trp157.

6.2.3 Mutagenesis of Ag85C

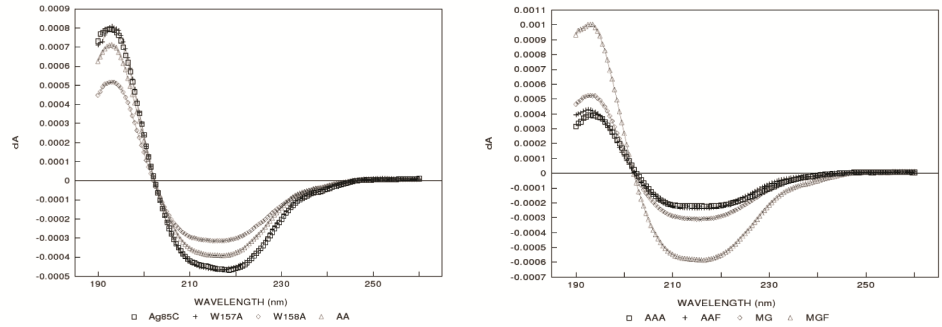
Site-directed mutagenesis confirmed that the secondary site residues are critical for Ag85 function. Using the Stratagene Quikchange® Lightening kit, a series of mutants, where the secondary site residues in Ag85C were systematically replaced by either alanine or the corresponding residues in Ag85A/B were generated (**Table 6.1**). Troubles with the mutagenesis protocol were initially encountered, but through increased primer length, decreased transformation volume and use of highly competent cells, all target mutants were generated. Double mutants of adjacent residues were generated with primers containing both desired mutations. Triple mutants were obtained through two rounds of mutagenesis.

Table 6.1. Mutants of Ag85C in pET-23b vector generated by site directed mutagenesis.

Entry	Mutant of Ag85C	Plasmid
1	Ag85C	pMRLB16
2	W157A	pKMB1
3	W158A	pKMB2
4	L230F	pKMB3
5	W157A, W158A	pKMB4
6	W158G, W157M	pKMB5
7	W157A, W158A, L230A	pKMB6
8	W157A, W158A, L230F	pKMB7
9	W157M, W158G, L230F	pKMB8

All proteins were overexpressed and found to be soluble (see **Appendix 2** for full details). The correct folding of these proteins was confirmed by Circular Dichroism (CD) spectroscopy (**Figure 6.6**). A concentration independent CD method was employed[14] and it was found that the CD spectra overlayed well and the resulting calculated values for percentage helix, sheet and turn matched the wildtype Ag85C protein, within error. These results, combined with the solubility of the expressed protein, confirmed that the Ag85C mutant proteins were correctly folded and that observed catalytic differences in the mutants would not be due to protein misfolding and aggregation, but rather to small changes in the secondary structure, caused by the change in specific amino acids.

a

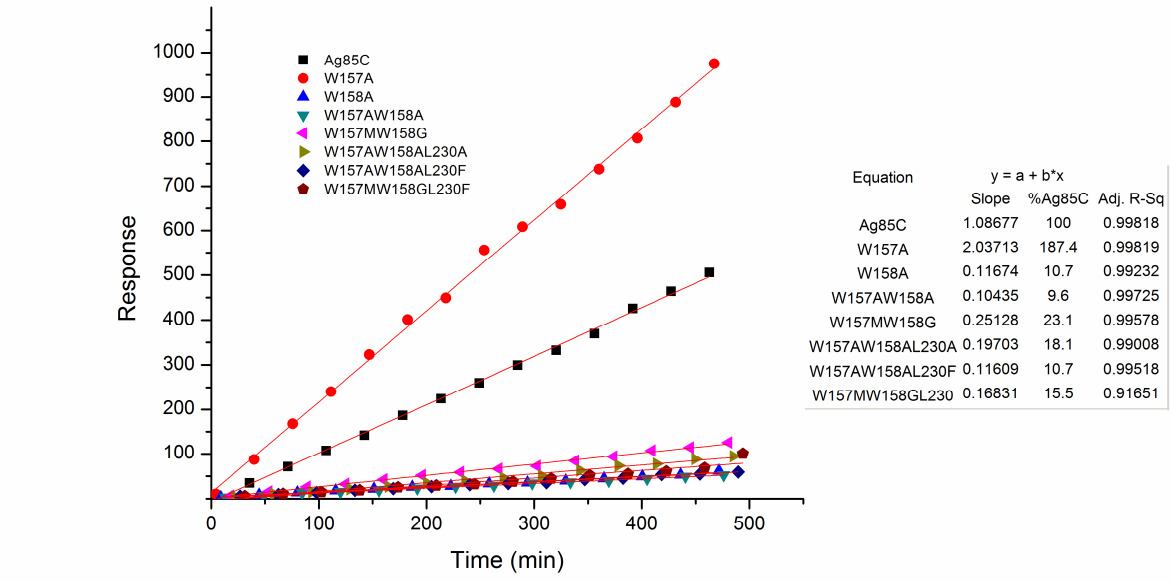


b

Protein	%Helix	%Sheet	%Turn	%Coil
Ag85C	41	18	13	28
W157A	42	16	13	28
W158A	37	19	13	29
W157A, W158A	45	18	13	26
W157A, W158A, L230A	42	14	13	29
W157A, W158A, L230F	44	16	13	28
W157M, W158G	40	16	13	29
W157M, W158G, L230F	40	18	13	29

Figure 6.6. (a) CD spectrums overlay well, supporting similar folding of mutant proteins, with comparable peak minima at 220 and 210 nm. Differences in peak intensisties are due to protein concentration. (b) Values calculated from spectra in 'a' confirm that protein secondary structures are within error.

The Ag85 activity assay (**Chapter 4** and Barry et al[15]) revealed that mutations in the Ag85C secondary site modulate enzyme activity with a large loss of activity observed for most amino acid changes. Catalytic efficiencies were measured by response rates of the mono-hexanoyl trehalose produced during transesterification of trehalose and dihexanoyl trehalose. Mutation of any residue, excluding W157A, dramatically reduced the catalytic efficiency of Ag85C to a level that is on par with Ag85B (**Figure 6.7**). In an interesting turn, W157A actually appeared to double the relative rate of activity. These results further suggested, but did not conclusively prove that this secondary site was important for activity.



Equation	y = a + b*x		
	Slope	%Ag85C	Adj. R-Sq
Ag85C	1.08677	100	0.99818
W157A	2.03713	187.4	0.99819
W158A	0.11674	10.7	0.99232
W157AW158A	0.10435	9.6	0.99725
W157MW158G	0.25128	23.1	0.99578
W157AW158AL230A	0.19703	18.1	0.99008
W157AW158AL230F	0.11609	10.7	0.99518
W157MW158GL230	0.16831	15.5	0.91651

Figure 6.7. Relative response of all Ag85C mutants reveals loss of activity when secondary site is mutated.

Molecular dynamics was used to compare the Ag85C W157A, W158A and W157A and W158A mutants with wildtype. From these simulations the W157A remained permanently in an 'on' arrangement with helix 8 constantly bent during the simulation (**Figure 6.8**). These results were consistent with the gain of activity observed in the enzyme assays (**Figure 6.7**). In the W157A and W158A double mutant, the lack of two large hydrophobic groups in the hydrophobic channel identified above seems to allow the two helices (9 and 8) to collapse towards each other, but even with this collapse the catalytic triad still cycles from active to inactive conformations. This collapse may inactivate the enzyme or cause a reduction of activity, consistent with the enzyme assay data. However, the mutations could cause a loss of function for other reasons in addition to the proposed switching mechanism. CD spectroscopy suggested that gross conformational changes were not occurring, but more localized changes would not be detected. MD simulations are suggestive, but not conclusive without further assays.

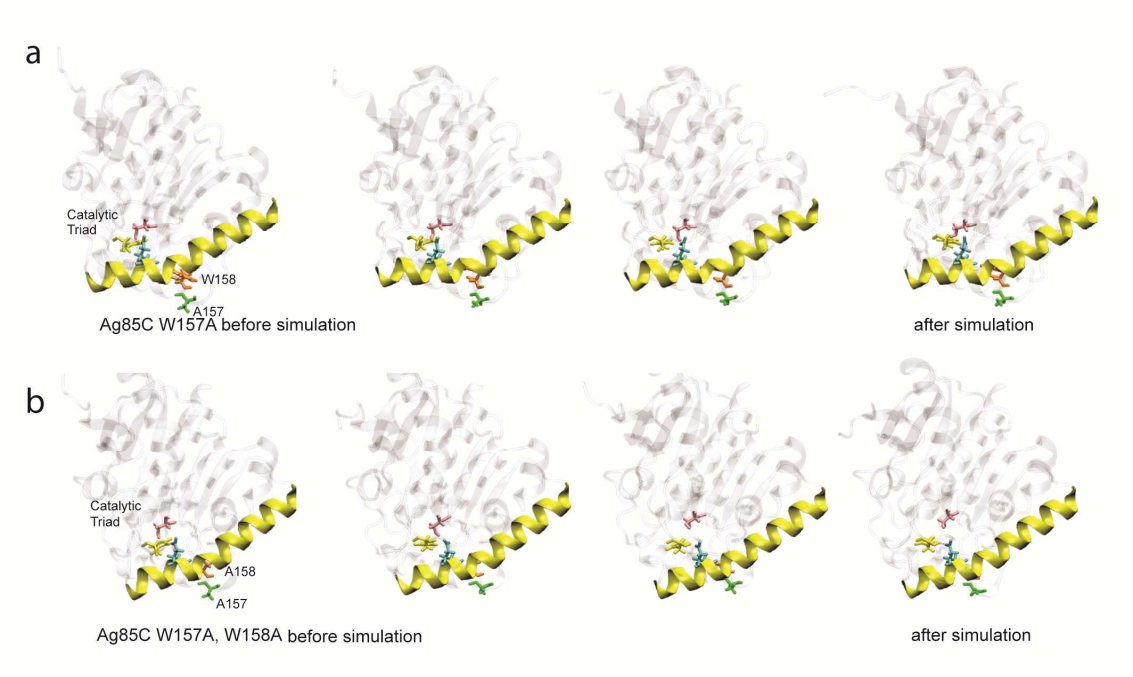


Figure 6.8. Molecular dynamics simulations of Ag85C mutants reveals that W175A mutation causes helix 8 to remain bent. (a) Ag85C protein containing W157A mutation (b) Ag85C containing both W157A and W158A. In contrast with 'a' helix 8 straightens in the double mutant.

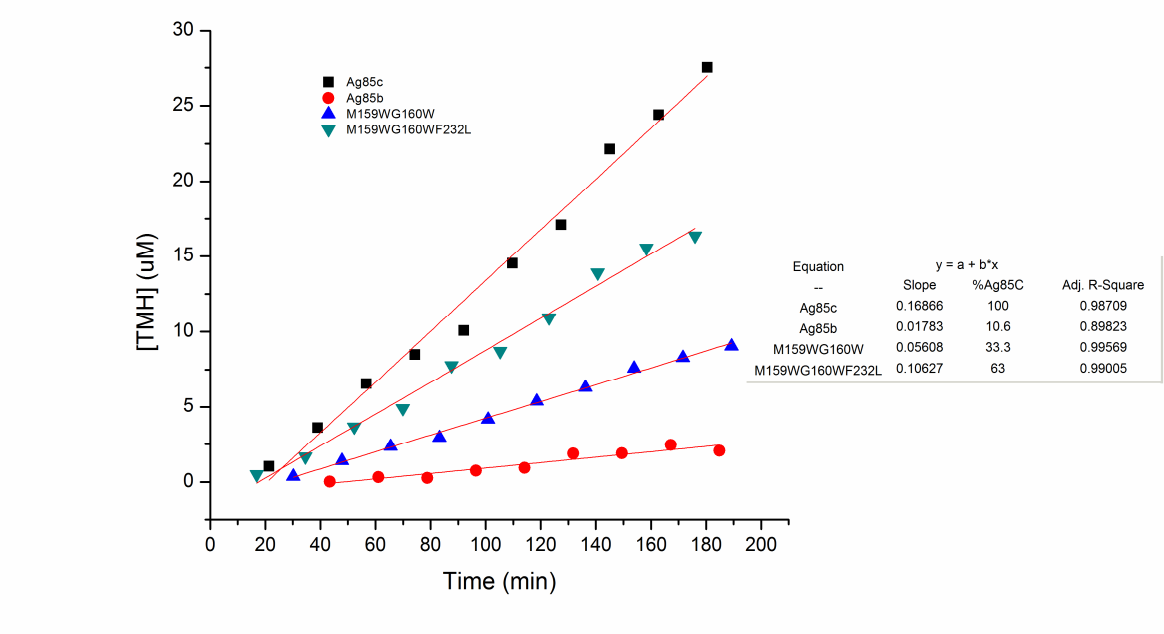
To further link the three switch residues (Trp157, Trp158 and Leu230) to Ag85 function, additional acitivity assays were required. It was hypothesized that if the Ag85C residues were mutated into the Ag85B protein, a gain of function should be observed, consistent with the loss of function observed for the reverse experiment. Compared with the loss of function observed with Ag85C, a gain of function would be more compelling evidence that these residues were critical for catalytic activity. The Trp Trp, Leu residues from Ag85C were systematically mutated into the Ag85B construct (**Table 6.2**).

Table 6.2. Mutants of Ag85B generated in pET-23b construct to confer Ag85C phenotype.

Entry	Mutant of Ag85B	Plasmid
1	Ag85B	pMLRB47
2	M157W G158W	pKMB9
2	M157W G158W F230L	pKMB10

These mutants were then expressed, purified and assayed for catalytic activity. Consistent with the hypothesis, systematic replacement of the B 'switch' residues with the corresponding residues from Ag85C afforded a B mutant with catalytic activity on par with

Ag85C (**Figure 6.9**). Taken with the Ag85C mutagenesis discussed in the previous section, these results confirmed that the secondary site regulated catalytic activity of the Ag85 enzymes and is a significant contributor to the relative differences in catalytic efficiency observed in *in vitro* Ag85 enzyme assays.



Equation	y = a + b*x		
--	Slope	%Ag85C	Adj. R-Square
Ag85c	0.16866	100	0.98709
Ag85b	0.01783	10.6	0.89823
M159WG160W	0.05608	33.3	0.99569
M159WG160WF232L	0.10627	63	0.99005

Figure 6.9. Relative rates of catalysis of Ag85B mutants reveal gain of activity when the Ag85C residues are mutated into Ag85B.

6.2.3.1 Dimerization state of mutants

With an *in vitro* explanation that linked secondary site residues to the relative differences in Ag85A,B and C activity, we turned our attention towards analyzing how the secondary site applies to *Mtb* cell wall biosynthesis. During several Ag85 purifications, a potential dimeric band was observed in SDS PAGE gels, particularly when insufficient sodium dodecyl sulphate (SDS) was added to the protein loading aliquot. Although the title 'Antigen 85 complex' implies protein association, this name actually reflects the triplet banding of the similarly sized Ag85A,B and C proteins, under SDS gel electrophoresis. Little research has been conducted as to whether these proteins are monomeric, dimeric or multimeric in solution and within mycobacterial culture. One report suggests that by 'natural PAGE,' lacking SDS, the Ag85A B and C proteins are monomeric[16]. To our knowledge no other publications have reported on the presence or absence of protein-protein interactions between these enzymes.

Native PAGE gels revealed that Ag85B putatively exists primarily as a dimer and Ag85C and Ag85A appear largely monomeric (**Figure 6.10)**. As these results were inconsistent with prior reports in the literature, this discovery was probed more fully. Mixtures of AG85A-B, A-C and B-C were analyzed by native PAGE and no apparent change in dimerization was observed, suggesting that homo and not heterodimers are the predominant form of the enzymes. The dimeric state of Ag85C was not dramatically changed when the Ag85B residues were mutated into the protein (**Figure 6.10**). None of these mutants, including the mutant where all three target residues W157M, W158G and L230F had been mutated to the Ag85A/B phenotype, showed any increased dimerization. Many of the mutants exhibited an apparent decrease in dimerization.

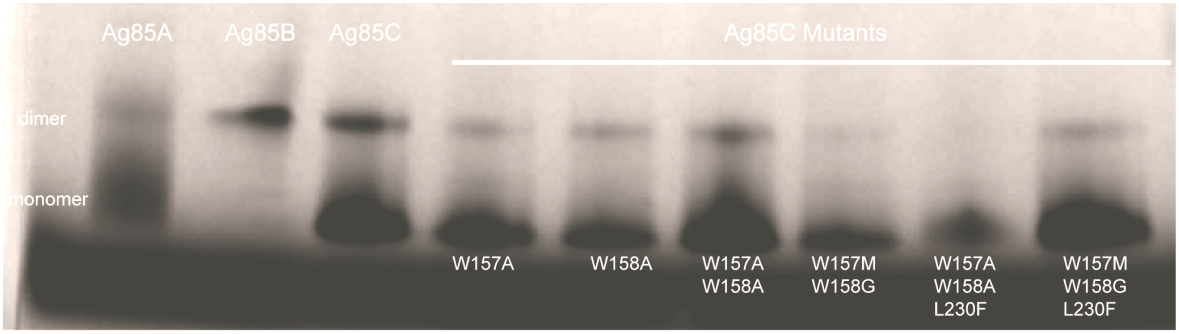


Figure 6.10. Native PAGE reveals dimeric nature of Ag85B. Mutants generated in Ag85C construct fail confer homodimer status on Ag85C.

In contrast with the Ag85C mutants, the mutations into Ag85B caused the dimer to dissociate. When the Ag85B mutants were run on native PAGE, the mutants were observed to be largely monomeric, whereas the parent Ag85B existed almost exclusively as a dimer (**Figure 6.11**), both at high and lower protein loading. These results suggest that the three residues of the 'switch' are necessary but not sufficient for Ag85 dimerization.

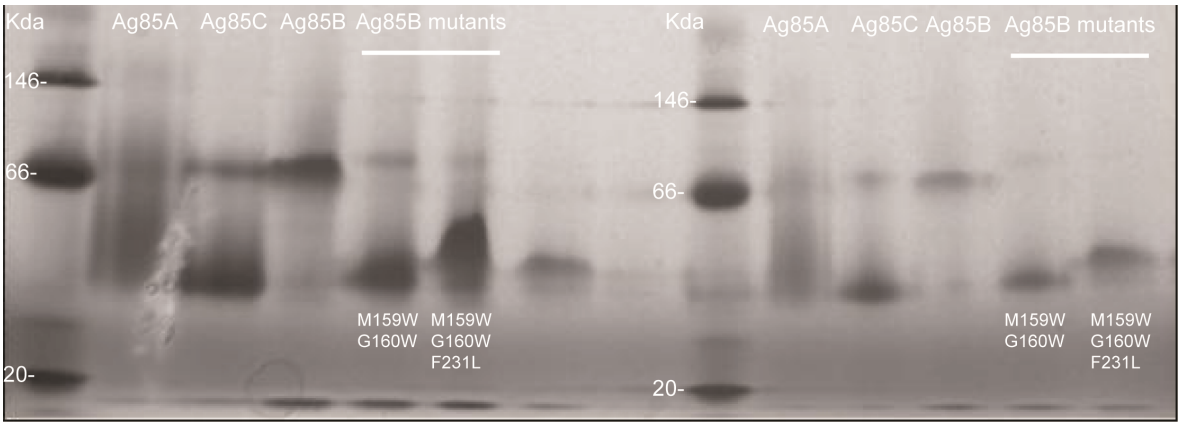


Figure 6.11. Native PAGE gel reveals that WW and WWL mutants in AG85B break up Ag85B dimerization. Lanes 2-6 (5 × loading) lanes 10-15 (1 × loading volume). Protein concentration does not affect dimeric state. Molecular weight of Ag85A, B and C is ~32 Kda.

6.2.4 Overexpression of Rv1288 and Rv0519c

In addition to the key differences in secondary site residues, the phylogenetic analysis (detailed in **Section 6.2.1**) also identified two additional putative mycolyl transferases encoded in the genome. Recent reports have suggested that both the *M. smegmatis* and *C. gluctamicum*[17] genomes encode for additional, putative trehalose acyltransferases. It was inferred that *Mtb* may also express additional trehalose acyl transferases. Upon blasting the sequences of Ag85A, B and C against the *Mtb* H37Rv genome, two additional genes (Rv1288 and Rv0519c) were identified as being distantly related to Ag85 (20% homology). Rv1288 was previously identified as a potential trehalose acyltransferase in 2005[18]. When the protein sequences were aligned with the sequences of the Ag85s, we found that both Rv1288 and Rv0519c have conserved putative catalytic triads. To fully explore the scope of the Ag85 enzymes, Rv1288 and Rv0519c were first modelled and it was determined that structural homology could exist between both enzymes and Ag85. Next, Rv1288 and Rv0519c were expressed, purified and assayed and it was found that Rv0519c exhibited novel trehalose acyl transferase activity, identifying this enzyme as a potential novel mycolyl transferase.

LOMETS threading[19] was employed to model the protein structures, yielding 10 PDB files for each of these annotated proteins all with high confidence value. These structures were then overlayed to ascertain whether the models were reasonable. Ramachandran plots were obtained for the models to evaluate their plausibility. Models for both Rv0519c (**Figure 6.12**) and Rv1288 **(Figure 6.13**) had good Ramachandran plots with most residues falling

within acceptable regions. Notably, and particularly present with Rv1288, both proteins have long tails that could not be fit by LOMETs. These regions are illustrated by the long tails (**Figure 6.12** and **Figure 6.13**). The amino acids of these tails were found to be very hydrophobic, so these regions could potentially be membrane domains that are not well modeled by the LOMETS service.

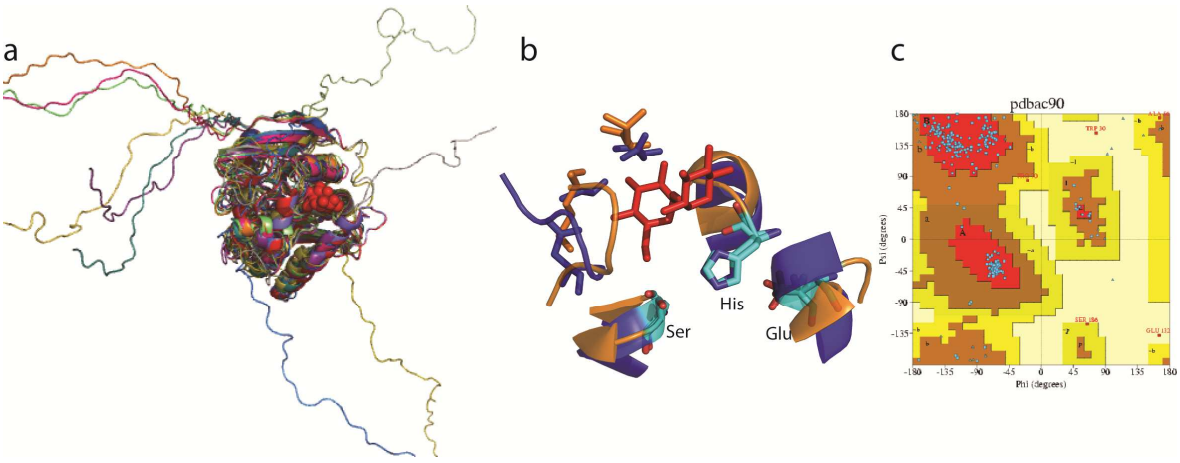


Figure 6.12. LOMETS modelling of Rv0519c suggest high structural homology to Ag85. (a) overlay of Ag85B and ten threadings for Rv0519c (b) overlay of active sites of Rv0519c model(orange) and Ag85B (blue) (c) Ramachandran plot.

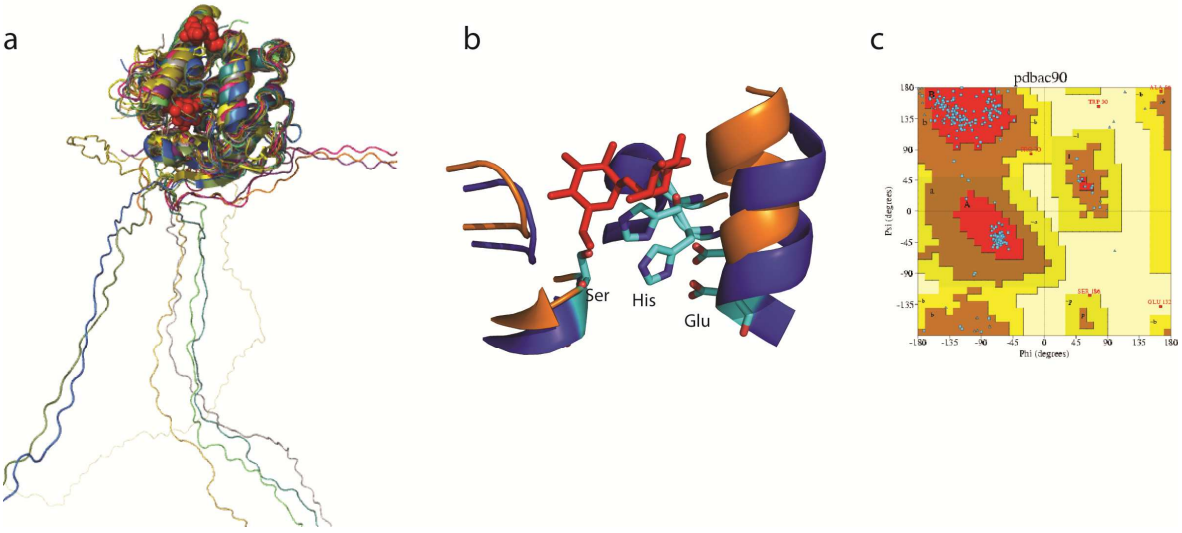


Figure 6.13. (a) overlay of Ag85B and ten threadings for Rv1288 (b) overlay of active sites of Rv1288 model (orange) and Ag85B (blue) (c) Ramachandran plot.

Upon blasting the Rv1288 sequence further, it became apparent that the tail domain contained three putative LysM domains[20]. The LysM domains are known to be involved in peptidoglycan binding, which supports the argument that this is a membrane binding domain.

With the assistance of Dr. Michael Dolan, we were able to model this tail portion as three Lysm domains and superimpose them over Ag85B bound to trehalose (**Figure 6.14**).

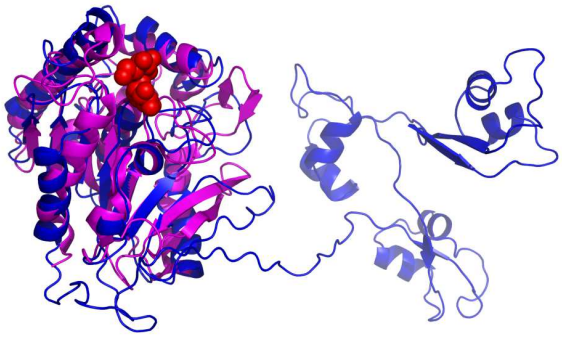


Figure 6.14. Rv1288 (blue) overlayed with Ag85B (pink) with trehalose (red spheres). Three LysM domains are modeled (blue at the right). Pdb 1F0P.

This preliminary *in silico* characterization of Rv0519c and Rv1288 supported that these enzymes could be trehalose acyl transferases with putative catalytic triads and secondary structures modelled as similar to the Ag85 proteins. Protein expression was attempted and the soluble proteins obtained were assayed for mycolyl transferase activity. Truncated genes, with hydrophobic 'tails' removed, were obtained from Genscript and genes were cloned into pET-23b at BamH1 and HindIII. Initial attempts at overepression of Rv0519c and Rv1288 afforded no expression for products of the correct size for both genes.

It was realized that the original cloning strategy had afforded an out of frame gene product in pET-23b, explaining the lack of protein expression (see **Appendix 2** for full details). When the Rv0519c and Rv1288 genes were cloned into pET-23a vector, robust protein overexpression was observed under induction with IPTG. The BamH1 and HindIII cloning strategy also allowed for cloning into the PMAL-C2X vector, generating C-terminal Maltose Binding Protein (MBP) fusion protein constructs. Rv0519c was found to be soluble, both as an MBP fusion and with a His-tag (pET-23a) (**Figure 6.15** and **Appendix 2**). Rv1288 was found entirely in inclusion bodies when expressed with a His-tag. Most of the Rv1288-MBP was found in the insoluble fraction and it was possible that the small portion found in the soluble lane had been forced into solution by the MBP and was not folded properly. Further Rv1288-MBP protein purification attempts in BL21 DE3 *E. coli* overexpressing chaperones and under slow expression conditions at cold temperatures (11 °C) afforded no increase in soluble protein. Attempts to purify Rv1288-hexahis from overproducing *M. smegmatis* strain

(*M. smegmatis* seemed to produce soluble protein (See **Section 6.2.6**)), yielded no detectable purified protein. The cloning of alternate truncations should be considered in future efforts to express and purify Rv1288.

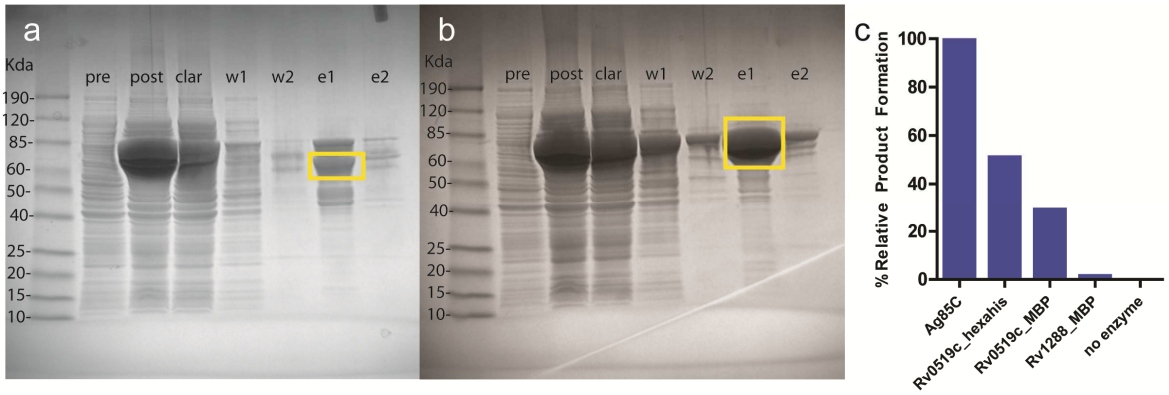


Figure 6.15. (a) Expression and purification of maltose binding protein (MBP) fusion with Rv1288 (b) expression and purification of MBP fusion with Rv0519c. Pre-pre induction with IPTG, post-post induction, clar-clarified lysate, w1-wash1, w2-wash2, e1-elute1, e2-elute2 (c) enzyme assay shows activity for Rv0519c.

In enzyme assays we found substantial trehalose hexanoyl transferase activity for Rv0519c and negligible activity for the soluble portion of Rv1288-MBP (**Figure 6.15c**). These results confirmed that Rv0519c can transfer acyl chains onto trehalose. However, it remains to be seen whether the long, hydrophobic mycolates will be processed by Rv0519c or if this enzyme is actually responsible for the transesterification of other acyl trehalose species. With the addition of future, corroborating mycolyl transfer assays, this will be the discovery of a novel trehalose mycolyl transferase.

6.2.5 Transcriptional profiling of Ag85

Given the knowledge that a secondary site in the Ag85B and C enzymes modulates *in vitro* activity and dimerization and that other Antigen 85 homologues may exist in the *Mtb* genome, the question was posed as to whether these differences would translate to mycobacterial culture. Whole cell assays were employed to further understand the 'redundancy' of the Ag85 enzymes. Transcriptional profiling under drug treatment and overexpression of the Ag85 constructs in *M. smegmatis* and *Mtb* were employed to further evaluate the differential roles of the Ag85 enzymes.

Transcriptional profiling of the Ag85 enzymes from previously reported studes[21] and current, unpublished microarray data in the Barry lab, revealed stark transcriptional differences between the *ag85B* and *ag85A/C* profiles. *ag85B* (Rv1886c) is strongly downregulated under drugs that affect mycolate synthesis (Isoniazid (INH) and SQ109). In contrast, *ag85C* (Rv0129c) and, more modestly, Ag85A were upregulated under INH treatment. *Ag85C* was upregulated under treatment with both SQ109 and INH. These transcriptional profiling differences were consistent between experiments and were statistically significant (**Table 6.3**).

Table 6.3. Ag85B is strongly downregulated under treatment with both INH and SQ109. Numbers reflect the log of change in gene expression level[ii].

Protein	Gene	10X MIC Isoniazid, 6h	10X MIC SQ109, 6h
Ag85A	Rv3804c	1.05±0.07	-0.12±0.13
Ag85B	Rv1886c	-2.56±0.61	-4.27±0.72
Ag85C	Rv0129c	0.98±0.56	1.82±0.39

Boshoff et al[21] revealed that other drugs, which inhibit cell wall biosynthesis, Eta (Ethambutol) and PZA (Pyrazinamide) affected gene expression similarly to INH and SQ109, although the effect of PZA was less pronounced. In contrast Rifampicin treatment, which targets the RNA polymerase, afforded a downregulation of both *ag85A* and *C* and a more modest decrease in *ag85B* expression. In this report, starvation caused an increase in *ag85B* expression as did low pH and non-replicating persistent state (NRP) (**Table 6.4**).

Table 6.4. From the supplementary information of Boshoff et al[21], differential expression level between *ag85A/C* and *ag85B* were revealed under multiple stress conditions. Eta-Ethambutol, Pza-Pyrazinamide, NRP- Non-replicating persistence. Rif-Rifampicin Numbers reflect the log of change in gene expression level.

Protein	Gene	pH4_8	starvation	Eta	NRP	PZA	Rif
Ag85A	Rv3804	-0.329301	-0.83035	0.7862365	0.5607483	-0.098797	-1.5472
Ag85B	Rv1886	-0.895514	0.9641262	-1.8102	1.3701901	-0.772561	-0.61325
Ag85C	Rv0129c	1.2958096	0.3138429	1.5675449	-2.289001	0.7150194	-1.32642

[ii] RNA microarrays were conducted on H37Rv *Mtb* by Dr. Kriti Arora and Dr. Helena Boshoff.

This data may have implications in the individual roles of Ag85A, B and C. INH treatment decreases the level of mycolates present in the cell. A compensatory upregulation of Ag85 would be expected in response to INH to salvage the remaining mycolates for TDM and TMM synthesis. SQ109 is an anti-mycobacterial drug that is thought to function through inhibition of the biosynthesis of TDM, potentially through inhibition of the MMPL3 TMM transporter[iii]. In addition to INH, Ag85B downregulation was also observed under SQ109 treatment, further supporting the hypothesis that Ag85B is serving a different function in cell wall biosynthesis than that of Ag85A or C. The absence of strong *ag85B* downregulation under RIF treatment supports that this response is not simply due to antibiotic-induced stressed, but rather is specific to cell wall biosynthesis. The evidence that Ag85B was strongly downregulated under INH and SQ109 and modestly upregulated under pH and starvation stress, is suggestive of a role in catabolism of TDM, TMM and/or mycolates. Mycolates store huge quantities of energy and possibly the bacteria would contain a mechanism to utilize this energy store. One can hypothesize that Ag85B may be recycling the mycolates back from TDM and TMM. Under conditions of low mycolate production (INH or SQ109) this process might be downregulated. Alternatively, Ag85B could be involved in a particular step towards TDM and TMM synthesis, such as orienting the TMM in the correct positioning towards the cell wall or shuttling substrates to particular positions in the outer cell wall. The dimeric state of Ag85B and relatively slow catalytic efficiency *in vitro* hint at a role that does not involve directly interacting with the cell wall and necessitates limited enzymatic turnover. This microarray data is suggestive of differential roles for Ag85B, relative to A and C; it remains to be seen whether the gene expression profiles correspond with lipid profiling and overexpression experiments.

An alternate explanation for the Ag85B downregulation is that this it was a consequence of the location of Rv1886 in the Ag85B operon. Unlike the other Ag85 proteins, the gene for Ag85B is part of an operon (**Figure 6.16**). This operon contains a chorismate mutase (Rv1885c) and RpfC (Rv1884c). Rpfs, resuscicitation-promoting factors, are muralytic enzymes that increase the cultivability of dormant bacteria, ie recovery from a latent state[22]. *Mtb* genome encodes two chorismate mutases[23], enzymes that catalyze the rearrangement of chorismate to prephenate in the biosynthesis of phenylalanine. Schneider et al found that *ag85B* is regulated by aromatic amino acid levels, as part of the operon with Rv1885c (the

[iii] Unpublished data from Dr. Kapil Tahlan

chorismate mutase), and that these three genes do act as an operon, although there is a second promoter after *ag85B*[24]. This second promoter suggests that the expression of these three genes may not be entirely linked, although the authors did find strong co-expression of all three genes. Analysis of the microarray data for the downstream genes from Ag85B suggested that both Rv1885c and Rv1884c were also downregulated under INH and SQ109, but more modestly than Rv1886c (85B), suggesting that the primary output of the downregulation is a reduction of the level of Ag85B expression.

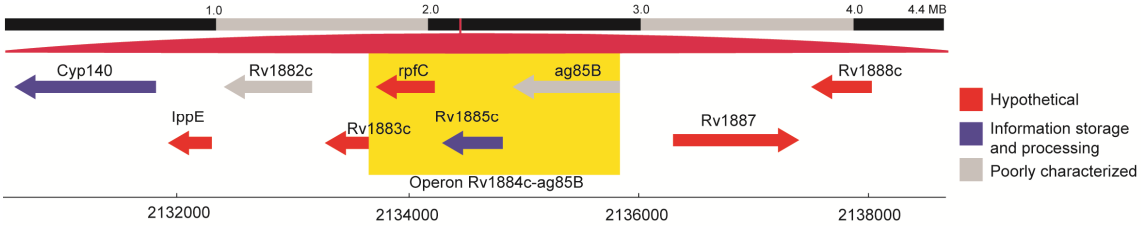


Figure 6.16. Rv1886 (ag85B) is part of an operon.

An additional piece of the Ag85B puzzle is that the promoter for *ag85B* is targeted by MtrA, the essential response regulator[25]. Unlike *ag85B,* the authors found that the promoters of *ag85A* and a*g85C* are not targeted by MtrA. *MtrA* is an essential gene in mycobacteria. Expression of the *fbpB* gene is found to be down-regulated in cells overproducing wild type MtrA, but not when they produce the non-phosphorylated form. *MtrA* also binds to the *oriC* origin of replication and the authors suggest that this may be a means to limit DNA replication under conditions of stress and/or *in vivo.* This mechanism is not fully characterized, but could prove an additional component to the Ag85B story.

6.2.6 Overexpression of His-tagged Ag85 constructs in *M. smegmatis* and *Mtb*

The genes for Ag85A, B, C, Rv1288 and Rv0519c with C-terminal His-tags were cloned and overexpressed in *M. smegmatis* and *M. tuberculosis.* It had been anticipated that overexpression could provide a handle for protein pulldowns, allowing identification of putative protein partners for the Ag85 proteins. These experiments confirmed that recombinant *Mtb* Ag85 proteins were overexpressed and secreted by *M. smegmatis*. Preliminary protein pulldown experiments suggested that no strong protein-protein interactions exist for Ag85C. The Rv0519c and Rv1288 proteins can also be overexpressed and Western blots confirm, consistent with the structural analysis, that these are cell wall associated proteins.

The vector pVV16, a derivative of pMV261, encoding a C-terminal His-tag was chosen as the target overexpression vector and was obtained from the Colorado State TBVTRM contract. This vector contains a cloning site with a variety of restriction sites and a C-terminal hexahis tag (**Figure 6.17**).

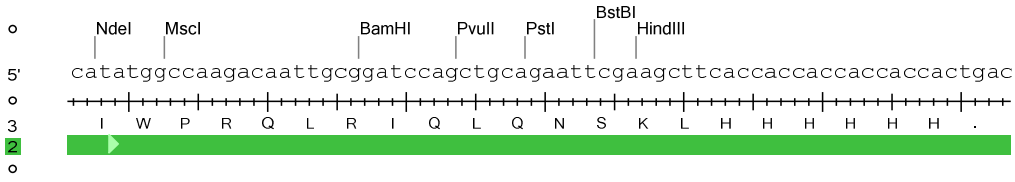


Figure 6.17. The cloning site of pVV16 shows unique restriction sites and His-tag.

After a period of cloning optimization a method for the successful cloning of genes into pVV16 was developed and implemented. Genes were cloned at the NdeI and HindIII sites, with care to ensure that they were in frame with the His-tag (see **Appendix 2** for full details). The Ag85A coding sequence contains a HindIII site, so the BamHI site was employed as an alternate ligation site. All genes were amplified by PCR from genomic DNA and then cloned into cloning vectors, either PCR 2.1 TA or Topo PCR 2.1 TA, obtained from Invitrogen. Colonies were carefully screened to avoid point mutations. The target genes were then excised from the cloning vector and ligated into the expression vector, pVV16. These vectors were then electroporated into electrocompetent mc^2155 *M. smegmatis.* Electroporation was successful and individual clones were then picked and cultured in liquid medium. Through this approach His-tagged overexpression strains for Ag85A, Ag85B, Ag85C, Rv1288 and Rv0519c have successfully been created. Truncated versions of Rv1288, lacking the peptidoglycan binding domain, and Ag85C, lacking the signal peptide (**Table 6.5**) were also cloned. These truncated versions may be useful controls in future experiments.

Table 6.5. Plasmids for cloning into hexahis vector pVV16. Restriction sites NdeI and HindIII. BamHI was employed in place of HINDIII for Ag85A. PGB-peptidoglycan binding

Gene cloned	Cloning vector PCR 2.1 or TOPO TA	pVV16	Overexpression in *M. smegmatis*
Ag85A	pKMB17	pKMB24	✓
Ag85B	pKMB18	pKMB25	✓
Ag85C	pKMB19	pKMB26	✓
Ag85C without signal peptide	pKMB20	pKMB27	✓
Rv0519c	pKMB21	pKMB28	✓
Rv1288	pKMB22	pKMB29	✓
Rv1288 without PGB domain	pKMB23	pKMB30	✓

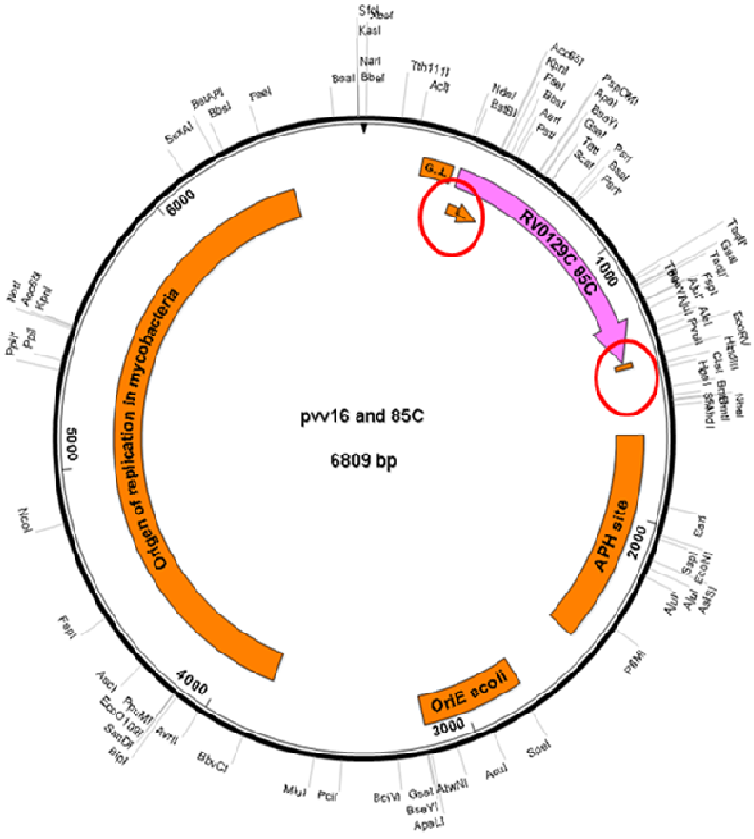


Figure 6.18. pVV16 expressing Ag85C. Plasmid map shows GroEL promoter and His-tag, flanking the cloned gene. Circles indicate signal peptide and His-tag.

The His-tagged protein expressing strains were generated for several reasons. It was thought that analysis of the lipid profiles of the overexpression strains could prove informative. Because these constructs are constitutively expressed, there was a good chance that the bacteria would compensate for additional protein and wild-type lipid profiles would be observed. The true strength of the His-tag is that it provided a handle from which to identify whether the proteins have been successfully overexpressed, what subcellular fraction they localize to and whether they are potentially interacting with other proteins.

SuperSignal West HisProbe western blotting kit was obtained from Pierce. The kit employs a novel nickel conjugated horseradish peroxidase that allowed for 1 step labeling and eliminated the need for a secondary antibody labeling. Overexpressing strains were lysed and the lysate, supernatant and soluble portion of the lysate were run on a protein gel. Western blotting protocols in the kit were followed precisely and good labeling of the His-tagged proteins was observed (**Figure 6.19**) with the majority of the Ag85C detected in the culture supernatant. Purified hexahis Ag85C was used as a positive control and lysate from wildtype *M. smegmatis* was employed as a negative control (See **Appendix 2** for further details). The His-tagged labelled proteins were identified well with only one background band, which was identified to most likely be the bovine serum albumin (BSA) from the culture medium, although it may also be another crossreactive protein (**Figure 6.19**). This background band disappeared when an anti-His-tag antibody (Genscript, A00186) was employed instead of the Pierce SuperSignal kit.

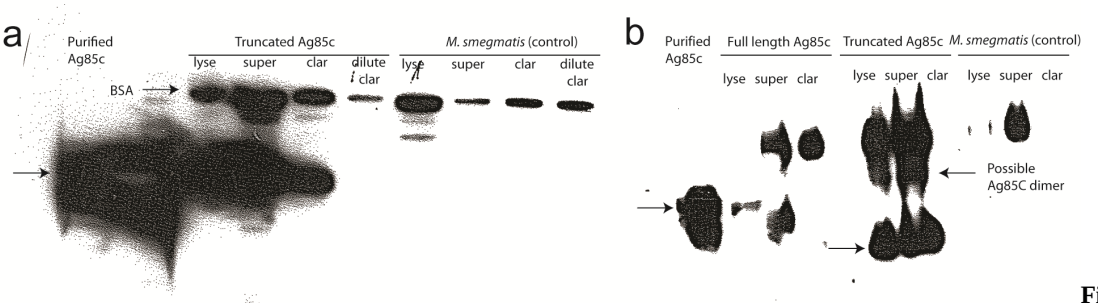


Figure 6.19. Western blots of hexahis tagged Ag85C overexpressing strain of *M. smegmatis* show overproduction of target protein (a) shows expression of a truncated Ag85C that lacks its signal peptide. BSA and Ag84C are indicated by arrows. Purified Ag85c is included in lane 1 as a control (b) Western blots of native PAGE gel lysate from *M. smegmatis* expressing full Ag85c or truncated Ag85c, lacking its signal peptide. Arrows indicate Ag85c . The gel may have run slightly crooked, explaining the slight shift. Between full and truncated Ag85C. Lyse- whole cell lysate, Super-supernatant, clar-clarified fraction. Truncated Ag85C is present in the clarified fraction, whereas the full length Ag85C localizes to the supernatant.

Western blots of lysates from *M. smegmatis* expressing either a full Rv1288 construct or a truncated Rv1288 (**Figure 6.20**) revealed that Rv1288 was most likely cell-wall associated (present in whole cell lysate, but absent from the supernatant and the clarified lysate). In contrast, the truncated Rv1288 appeared in the clarified fraction, suggesting that at least part of the truncated Rv1288 was expressed in a soluble form in *M. smegmatis.* Protein purification attempts of this truncated construct from *M. smegmatis* proved unsuccessful, affording none of the target protein.

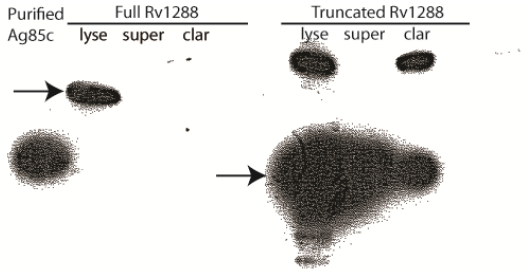


Figure 6.20. Western blots of lysate from *M. smegmatis* expressing full Rv1288 or truncated Rv1288, lacking its peptidoglycan domain. Lyse- whole cell lysate, Super-supernatant, clar-clarified fraction .

It was hoped that the hexahis constructs would facilitate identification of proteins that interact with mycolyl transferases. Native PAGE or crosslinking were first considered for Ag85 binding partner identification. A native PAGE of lysates from overexpressing strains was blotted for His-tags and no protein interaction bands were identified. The probe did label the insoluble fraction of the bacteria slightly, but this was very likely non-specific labeling. Next, crosslinking of the culture with gluteraldehyde prior to lysis and gel elextrophoresis was attempted. This experiment yielded very little western blot signal (**Figure 6.21**). It is possible that the gluteraldehyde interferes with the antibody labeling process. In this experiment higher detergent concentration (4% Triton X-100) was used, which seemed to help solubilize the Rv1288, which had been only detected at low levels previously. A benchmark His-tagged protein ladder (invitrogen LC5606) was used to identify the molecular weight of the crosslinked proteins. This ladder worked well with the western blotting system. Loading of ladder was reduced in further experiments prevent signal saturation.

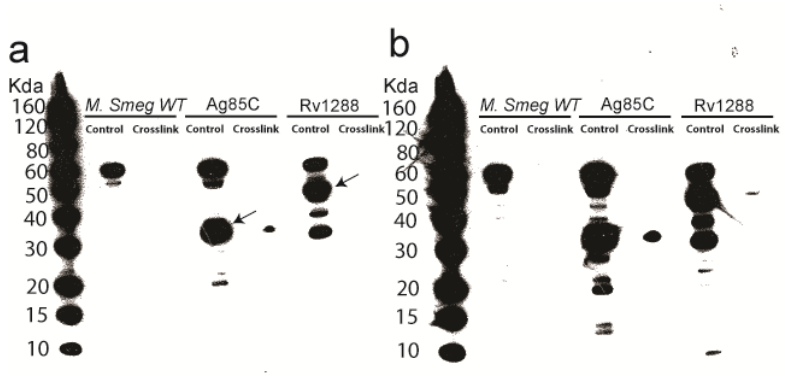


Figure 6.21. Western blots of mc^{2}155 *M. smegmatis* either overexpressing no protein (control) or Ag85c or full length Rv1288. Proteins with a MW about 60,000 Kda were presumed to be BSA (MW 66,399). Only small amounts of target proteins (arrows) were detected in crosslinked lanes. (a) blot developed for 10 seconds. (b) blot developed for 1 minute shows some apparent degradation of proteins. WT-wildtype

The absence of crosslinking products in the soluble fraction could be explained by crosslinking of the target proteins to insoluble, membrane bound proteins. A membrane solublization protocol (Colorado State TBVTRM) was employed, utilizing SDS, to solubilze protein fractions after crosslinking. *M. smegmatis* was crosslinked with gluteraldehyde, lysed and the soluble lysate and membrane component were fractionated by centrifugation. Preliminary experiments with the Ag85C overexpression strain, compared with mc^{2}155 *M. smegmatis*, revealed potential bands of interest that were only present in the crosslinked Ag85C strain (**Figure 6.22).** This target band (**Figure 6.22**) was excised from the gel, exposed to a tryptic digested and analyzed by MS proteomics. The proteomics analysis failed to reveal any proteins unique to the Ag85C crosslinked strain.

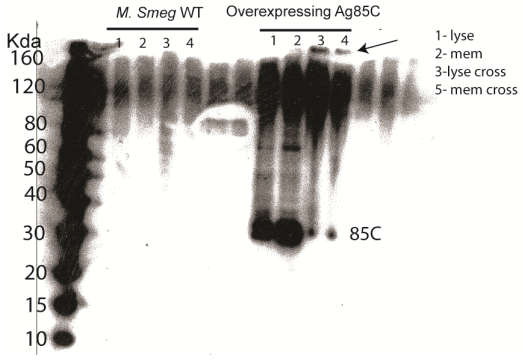


Figure 6.22. Crosslinking experiments on Ag85C strain reveal a potential band of interest (arrow). Lyse-whole cell lystate, mem-membrane fraction, cross- culture was crosslinked prior to lysis. WT- wildtype.

The protein interaction experiments are still ongoing but to date, have not yet yielded any protein binding partners, suggesting that the Ag85C protein-protein interactions are weak, transient and/or negligible. Although these experiments did not identify protein binding partners, use of the His-tag constructs has validated the overexpression and the secretion of the Ag85 proteins in *M. smegmatis*, suggesting appropriate folding and function. Minimal non-specific antibody labeling was observed under optimized labeling conditions. Some apparent protein degradation was observed. However, with better protease inhibitors, this could likely be prevented. These experiments hint at the scope of possibilities with overexpressing His-tagged constructs.

6.2.7 Overexpression of Ag85 as Tet inducible GFP fusion proteins

As the constitutive His-tagged overexpression constructs were not conclusive of any differences between the Ag85 enzymes, a more tunable, tetracycline (Tet) inducible expression system. The parent plasmids for this system were a kind gift from Dr. Sabine Ehrt (full details of these constructs can be found in **Appendix 2**). The original Tet system involved expressing two vectors in mycobacteria concurrently, one which consisted of a pMC1 plasmid to express the tetracycline repressor (tetR) and a second, pSE100, containing the Tet regulated promoter and gene of interest. To streamline this system, a single vector containing both components was designed by Dr. Ehrt. With this second generation system a *E. coli* cloning vector is used, containing the TetR construct and the operon (vectors either pTC-2M1L or pTE-2M1X) and the gene of interest is first cloned into cloning vector and then the whole cassette containing the gene of interest is excised with the rare cutters PacI/NotI and then ligated into pTC- or pTE-msc1. To further simplify the cloning it we discovered that this process can simply been done in one vector, either pTC-2M1L (integrating) or pTE-2M1L (replicating).

A sequential cloning directly into the pTE-2M1L instead of into the shuttle vector was chosen. For the large number of genes that were needed for this study this approach reduced the number of cloning steps. pTE-2M1L contains three unique cloning sites within the lacZ gene (SphI, EcoRV and NotI). It was found to be possible to clone the *gfp* or *rfp* at the EcorRV and NotI sites, generating pTE-2M1L-RFP or pTE-2M1L-GFP (**Figure 6.23**). In a subsequent cloning step the genes of interest were cloned upstream of the *rfp* or *gfp*, between SphI and EcoRV (SmaI was used for Ag85B and C, as they contained internal EcoRV sites). This strategy allowed the production of Tet inducible GFP fusion proteins of Ag85A,B,C and

Rv1288 and Rv0519c as well as Ag85B mutants at the secondary site that showed gain of activity *in vitro* (**Section 6.2.3**).

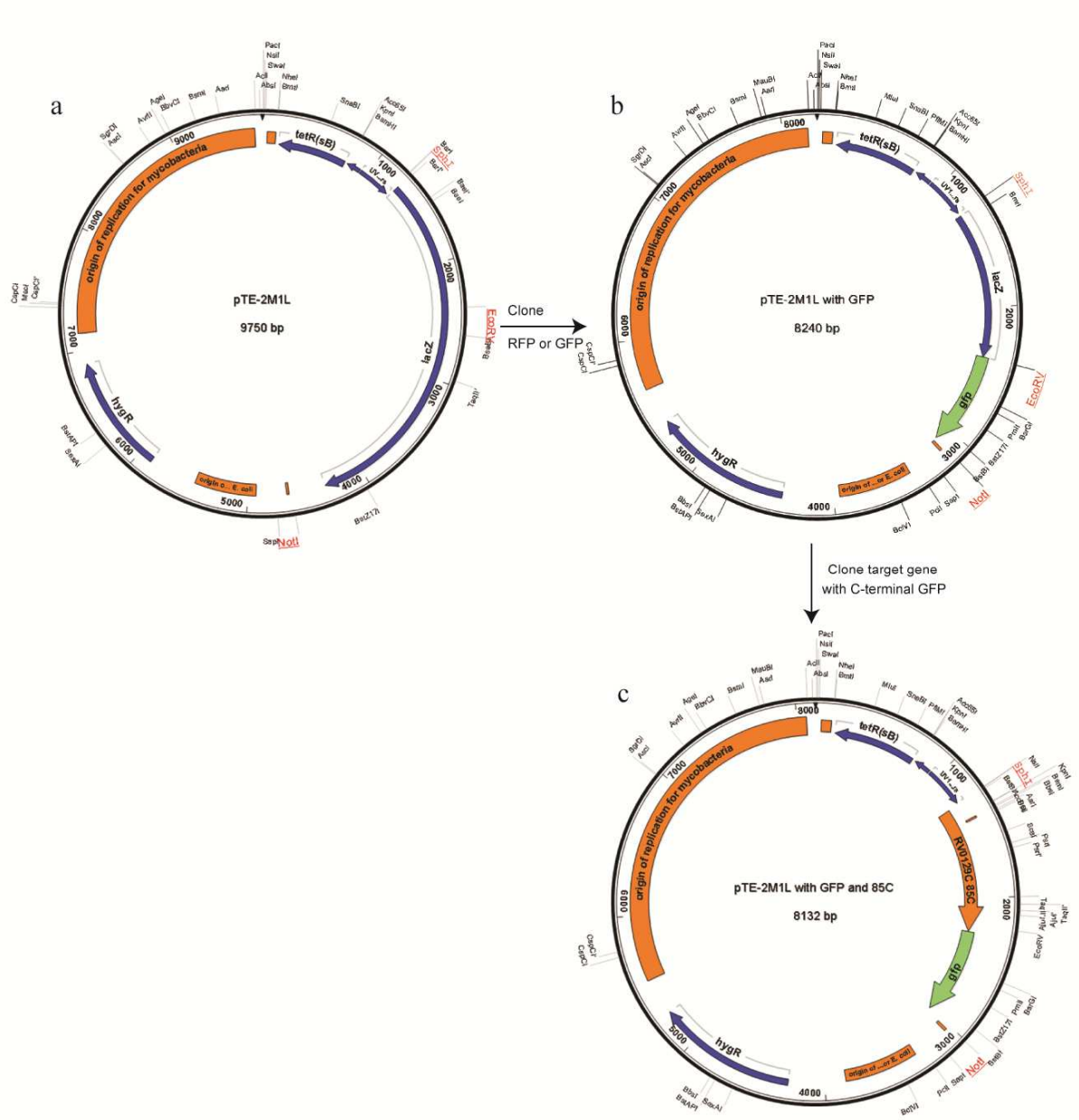


Figure 6.23. The cloning strategy for the generation of our target fusion proteins in tet inducible replicating vector (a) pTE-2M1L. (b) *gfp* or *rfp* were amplified by polymerase chain reaction (PCR) to contain EcoRV and NotI restriction sites and cloned at the respective restriction sites in pTE-2M1L. (c) The genes of interest have been amplified by PCR to include SphI and EcoRV restriction sites and ribosome binding site (RBS) and cloned at the respective sites in pTE-2M1L-GFP. For Ag85C and B, SmaI was utilized in place of EcoRV as both genes have internal EcoRV sites. Only unique restriction sites are shown. Sites employed in the cloning detailed are highlighted in red.

Future diversification of the library, including the generation of double overexpression strains, was incorporated into the cloning strategy. pTC-2M1L is the integrating equivalent of pTE-2M1L. pTC vector does not have a unique SphI site at the start of the *lacZ* gene. Because of the unique NotI and PacI sites in both vectors, in future work, the cloned cassette can be excised from the pTE vector and ligated it into the pTC vector. The pTE vector encodes hygromycin (Hyg) resistance, whereas the pTE encodes kanamycin (Kan) resistance. Strains expressing two proteins could be generated by sequential electoporations with the overexpression plasmids, followed by selection for double overexpressors with both antibiotics. In summary this cloning strategy should allow for the rapid diversification of the *gfp* and *rfp* fusion protein vector library.

Plasmids expressing *rfp* and the *gfp* were available in the Barry lab in the pMV261 overexpression vector, containing the heat shock promoter, GroEL. These *rfp* and *gfp* constructs were validated to express robustly in mycobacteria, having previously been used in the Barry lab to generated RFP and GFP *Mtb* strains. Primers were designed to amplify these sequences with EcoRV and NotI sites. PCR proceeded smoothly (See **Appendix 2** for full details). Consideration was given to what *rfp* would be appropriate for this overexpression. Sequencing of the *rfp* revealed that it was most similar to mCherry, which would be monomeric. However, the construct was not 100% homologous to mCherry, which caused concern because some of the earlier RFP generations (DsRed1) are actually tetrameric in solution, which makes them incompatible with the production of fusion proteins, as they will form insoluble aggregates (discussed in **Chapter 5)**. Because of these concerns with tetrameric RFP, the cloning of GFP fusion proteins, which were undoubtably monomeric, was the initial focus.

After restriction digest, The *rfp* and *gfp* were successfully ligated into the pTE-2MIL vector, generating pTE-2M1L-GFP and pTE-2M1L-RFP, with the remaining portion of *lacZ* fused to the *rfp* and *gfp*. The *lacZ-rfp* and *gfp* constructs were electroporated into *M. smegmatis* and overexpression was observed by microscopy and was tightly regulated by anhydrotetracycline (ATc) (**Figure 6. 24**).

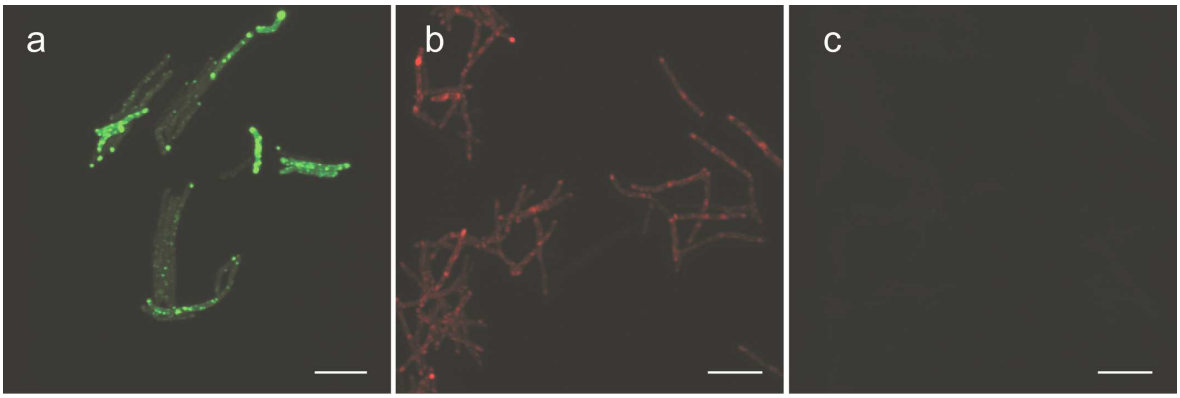


Figure 6. 24. Tet inducible GFP and RFP strains of M. Smegmatis (a) Strain expressing pTE-2M1L-GFP shows accumulation of GFP in inclusion bodies. (b) Strain expressing pTE-2M1L-RFP. (c) Very low background in the absence of anhydrotetracycline (ATc). Scale bars 5 µm.

Initial Tet induction experiments with *rfp* and *gfp* construction showed significant accumulation of protein into large vacuoles. However, subsequent experiments with reduced induction time (3 hours compared to 24 hours) showed much less 'inclusion body' like formation (**Figure 6.25**). These experiments confirm that the tet regulation is effective and that tetracycline dependent expression of GFP and RFP is possible in *M. smegmatis.*

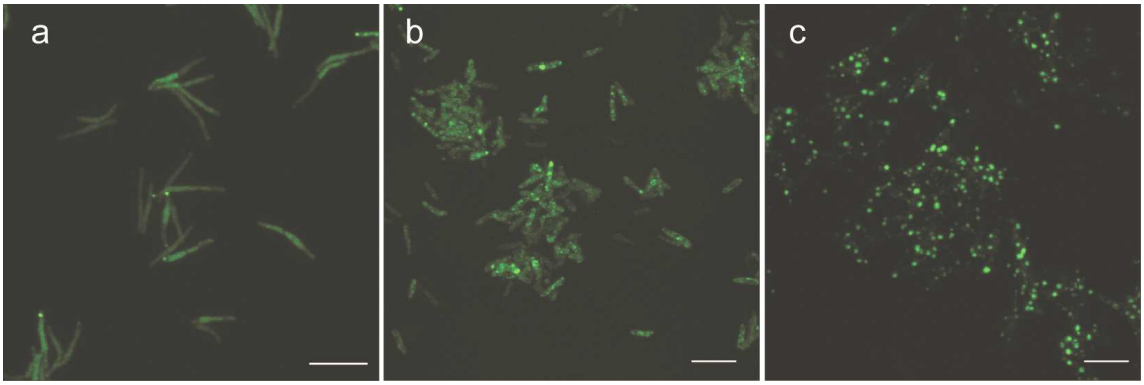


Figure 6.25. Tet induction of GFP-LacZ strain in *M. smegmatis* in tween free medium shows Gradual accumulation of protein into inclusion bodies. (a) 3 h induction (b) 18 h induction (c) 24 h induction. Scale bars 5 µm.

With *rfp* and *gfp* TET vectors in hand, we turned our attention towards the cloning of the target genes upstream of the fluorescent reporters. Problems were initially encountered with the ligation of the genes at the SphI and EcoRV sites, potentially because blunt and 3' overhang restriction sites can yield reduced ligation efficiency. With longer ligation time at

colder temperature, the desired pTE-85C-GFP fusion vector was obtained and electroporated it into *M. smegmatis.* Upon induction with ATc, no fluorescence was observed. After some confusion it was realized that through cloning at the SphI site, the ribosome binding site (RBS directly downstream of the SphI restriction site) had been accidentially removed and, without the RBS, expression was dramatically reduced (**Figure 6.26**).

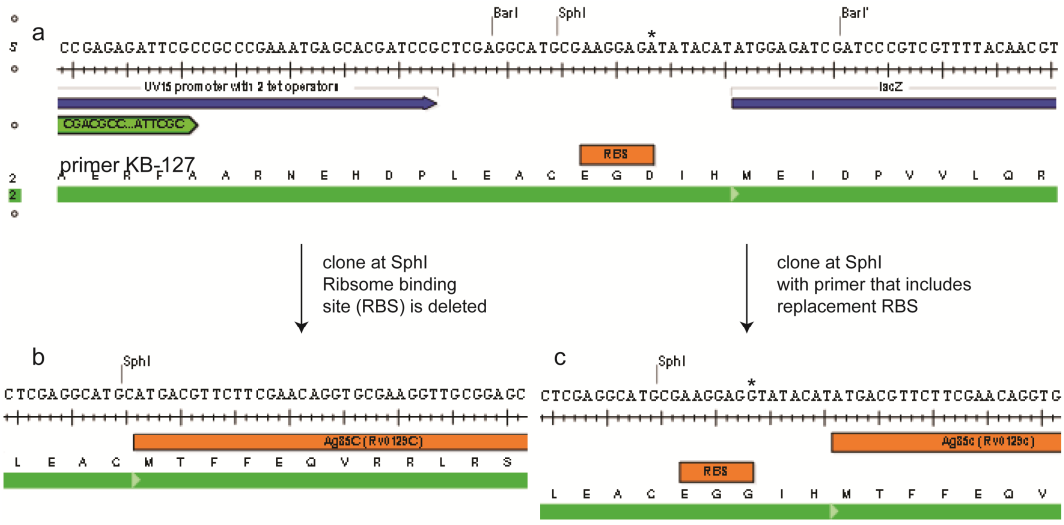


Figure 6.26. Demonstrates the new cloning approach to include the ribosome binding site in the primer, so as to retain it during cloning. Here 85C is cloned into PTE-2M1L and includes the improved RBS. (a) tet operator and ribosome binding site (RBS) in PTE and PTC-2MIL plasmids (b) Cloning at SphI site leads to deletion of the RBS site downstream of the restriction site. (c) New strategy used primers that included the RBS ensures that RBS is maintained in the final construct. Bases inducated with stars (*) are those changed in the RBS to enhance ribosome binding.

To salvage the cloning of the genes into the *gfp* fusion vectors, primers that contained, at the 5' end, the SpH site and the sequence for an improved ribosome binding site were employed. The following sequence was used 5'-GGGGCATGC G**AAGGAGG**TATACAT ATG ...-3'. The first six bases correspond to the SphI site. The only difference from the pTE-2M1L parent sequence was that one A was changed to a G (the last G in the bolded part of sequence-3). This sequence was reported[iv] to afford better expression than the original RBS in pTE-2MIL. With these new primers all genes were reamplified by PCR and cloned, first into cloning vectors and then into the pTE-2MIL-GFP and then electroporated into *M. smegmatis* and H37Rv *Mtb*. With the RBS site in the sequence all genes expressed robustly under anhydrotetracyline dosage.

[iv] Personal communication with Dr. Sabine Ehrt suggested this improved sequence.

Preliminary results in *M smegmatis* strains suggested that overexpression of the Ag85 proteins could differentially affect the incorporation of mycolates into the cell wall (See **Section 6.2.9.2** below). It was hypothesized that some of the observed incorporation might be a result of chaperoning by Ag85B and not as a direct result of catalysis of the transmycolation reaction. To more fully dissect the differences in lipid incorporation by Ag85, the catalytically inactive serine-alanine (SA) mutants of all three isoforms were generated by site directed mutagenesis and cloned into the tet inducible construct. Additionally the mutants at the secondary site of Ag85B (notated WW, WWL and WWLG) that conferred increased catalytic activity were also cloned into the tet construct to attempt to observe whether the Ag85 B mutant could afford an Ag85C-like lipid profile. For a full summary of all genes cloned into the Tet inducible system see **Table 6.6.**

Table 6.6. Plasmids for cloning into the tet inducible vector pTE-2M1L. Restriction sites EcorV and NotI were used for *rfp* and *gfp*. SphI and EcorV or SmaI (Ag85B,C) were used for target proteins. SA-catalytic serine to alanine mutant. Check (✓) marks indicate that the strain overexpressing the particular plasmid indicated was generated. X indicates that plasmid or strain was not generated.

Gene cloned	PCR 2.1 or TOPO TA	PTE-2M1L	PTE-2M1L-Gene-GFP	Expression in *M. Smeg*	Expression in *Mtb*
GFP	pKMB31	pKMB36	NA	✓	✓
RFP	pKMB32	pKMB37	NA	✓	✓
Ag85A	pKMB39	X	pKMB50	✓	✓
Ag85B	pKMB40	X	pKMB51	✓	✓
Ag85C	pKMB41	X	pKMB52	✓	✓
Ag85A SA	pKMB42	X	pKMB53	✓	✓
Ag85B SA	pKMB43	X	pKMB54	✓	✓
Ag85C SA	pKMB44	X	pKMB55	✓	✓
85B WW	pKMB45	X	pKMB56	✓	✓
85B WWL	pKMB46	X	pKMB57	**X**	✓
85B WWGL	pKMB47	X	pKMB58	**X**	✓
Rv1288	pKMB48	X	pKMB59	✓	✓
Rv0519c	pKMB49	X	pKMB60	✓	✓

6.2.8 Microscopy of Tet inducible plasmids

One possible difference between the Antigen 85 proteins was subcellular localization, with the individual isoforms localizing to different portions of the cell wall (eg the poles, snap or elsewhere). The fluorescent Tet strains provided tools to interrogate protein localization. **Figure 6.27** reveals that Tet induction worked with robust induction under anhdrotetracycline and very low background in the absence of ATc. Similar to the *lacZ-gfp* strain, the *rv1288-gfp* fusion strain displays robust labelling of the bacteria, presumably accumulating on the cell wall. Microscopy revealed that the Ag85C was secreted efficiently and under growth in

7H9 medium (contains Tween 80 and BSA) with very low accumulation within/on the bacteria. Halos of Ag85C-GFP around clumps bacteria were apparent in some images. BSA and tween may strip weakly associated protein from the outer cell wall, removing any apparent localization. To address the detergent effect *M. smegmatis* strains were grown in tween-free GAS medium and were induced for 18 hours, at which point cells were fixed and imaged by confocal microscopy. Even in medium without tween and BSA, relatively little adherence of the target proteins to the cell wall was observed (**Figure 6.28**). These experiments confirmed that Ag85-GFP fusion proteins were secreted and had low association with the mycobacterial cell wall. To our knowledge there are few reports[26,27] of secretion of a GFP tagged protein in mycobacteria and previous examples employed vastly truncated proteins fused to GFP. Personal communications with others in the field[v] suggested that appropriate secretion of fusion proteins by mycobacteria was a relatively rare occurrence.

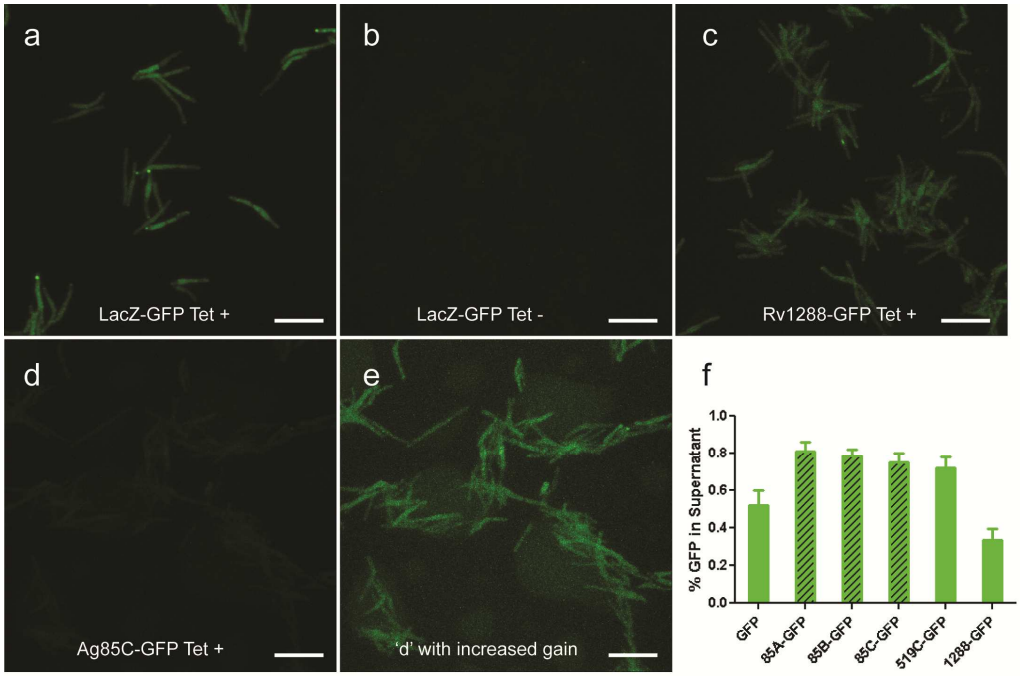


Figure 6.27. Anhydrotetracycline (ATc) induction of *M. smegmatis* strains expressing GFP fusion proteins revealed robust protein synthesis for all constructs over 3 hour induction, but limited protein localization for the Ag85C overexpression strain. (a) LacZ-GFP overexpression shows localization to the cytoplasm. (b) In the absence of ATc, low background expression is observed. (c) Rv1288-GFP shows localization to the cell wall and/or cytoplasm. (d) Microscopy of the Ag85C-GFP expression construct at the same gain as in 'a-c' shows lower fluorescence. (e) When the gain is increased, Ag85C-GFP fluorescence can be observed on *M. smegmatis* as well as in halos surrounding the mycobacteria. (f) Quantification of the presence of fluorescent protein in supernantant shows high levels of secretion of the Ag85 proteins. Scale bars 5µm.

[v] Personal communications with Professor David Russell suggested to us that bacterial attempts at secretion of other GFP fusion proteins had resulted protein accumulation at the mycobacterial poles, where the protein secretion machinery is thought to reside.

The GFP Tet inducible constructs offered a means to quantify the subcellular localization of the fusion proteins through a comparison of the relative fluorescence in whole cell culture versus in culture supernatants. Rv1288-GFP fluorescence was largely absent from culture supernatants. In contrast, Ag85-GFP fluoresnence localizes to the supernatant, confirming that the fusion proteins are secreted; LacZ-GFP fluorescence was found in both supernatant and pellet. Some supernatant fluorescence, even in the absence of protein secretion, is expected, due to bacterial death and lysis, particularly with soluble cytosolic proteins. The low supernatant fluorescence of the Rv1288-GFP suggests that this protein is cell wall associated; even with bacterial lysis, this protein is not released into the culture medium.

Microscopy of cultures grown in GAS medium (BSA and tween free) revealed low localization of all Ag85 isoforms. The Ag85-GFP proteins showed much lower accumulation within the mycobacteria relative to LacZ-GFP, presumably due to their extracellular location. Consistent with literature reports[28] about the localization index of the Antigen 85 proteins, the Ag85A strain accumulated more protein than Ag85C or B strains (**Figure 6.28**). It is unclear whether this greater Ag85A accumulation was due to a difference in relative rates of protein secretion or actually due to a greater Ag85A affinity for the cell wall. The Ag85B strain displayed negligible accumulation of fluorescence, even in tween free medium (**Figure 6.28b**). Ag85A and C strains accumulated of protein in either inclusion bodies or at the poles or, perhaps, both (**Figure 6.28a,b and d,e**). The Ag85C strain images featured reproducible GFP halos surround the bacteria, suggesting that the protein may not associate directly with the cell wall but exist as a cloud surround the bacteria (**Figure 6.28d,e**). These images suggest that Ag85AB and C are differentially distributed between the cell wall and culture supernatant.

The overexpression of Ag85AB and C fused to GFP in *M. smegmatis* suggested greater accumulation of Ag85A and Ag85C on the cell wall, relative to Ag85B. Although purely speculative, these results may relate to the dimeric state of the protein, with the Ag85B dimer unable to productively interact with the mycobacterial cell because its membrane interacting face is dimerized with another Ag85B unit. Greater or less cell wall localization may also have implications in which of the Ag85 isoforms catalyzed the transmycolation of the TDM and arabinogalactan with greater localization facilitating more Ag85 substrate access. One

caveat of these results is that *M. smegmatis*, although closely related to *Mtb*, may feature undiscovered differences in the Ag85 protein functions relative to *Mtb*.

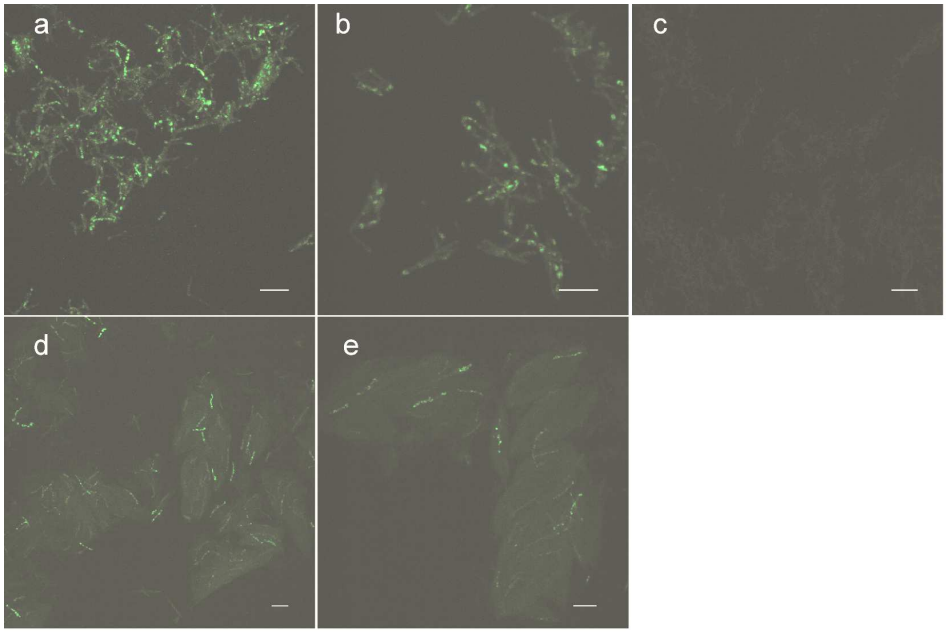


Figure 6.28. ATc induction of Ag85A,C and C strains in *M. smegmatis* in tween free medium shows higher cell wall accumulation with Ag85A (a,b), followed by Ag85C (d,e). Low affinity was observed for Ag85B (c). GFP halo appears to surround Ag85C bacteria. Scale bars 5 µm.

6.2.9 Overexpression of Tet inducible constructs in *Mtb*

With the GFP Tet induction suggestive of differential accumulation of Ag85A,B and C in *M. smegmatis*, the reproducibility of these results was tested in *M. tuberculosis.* The GFP fusion protein vectors were electroporated into H37Rv and all target strains were obtained (**Table 6.7**). These strains were inducible by anhydrotetracyline and, consistent with the *M. smegmatis* results, the majority of the fluorescence for the Ag85 fusions was found in the supernatant.

Microscopy revealed protein induction over a 72 hour period. The H37Rv strains were grown in GAS media without tween and were induced with ATc for 18 hours at which point the cultures were fixed, washed and mounted. The LacZ-GFP control showed low level fluorescence, relative to background and the Ag85A, B and C strains (not shown) showed no difference between induced and uninduced cells. The lack of fluorescent accumulation was most likely due to the short induction time limiting the protein accumulation. This experiment was repeated with a longer induction time in strains expressing a constitutive RFP marker

alongside the Tet inducible GFP fusion. With a 72 hour induction time, as had been previously suggested[29], robust protein overexpression was observed in the GFP Tet strain with some accumulation of fusion protein in inclusion bodies. Even at these long induction times, negligible Ag85 protein localization was observed (not shown).

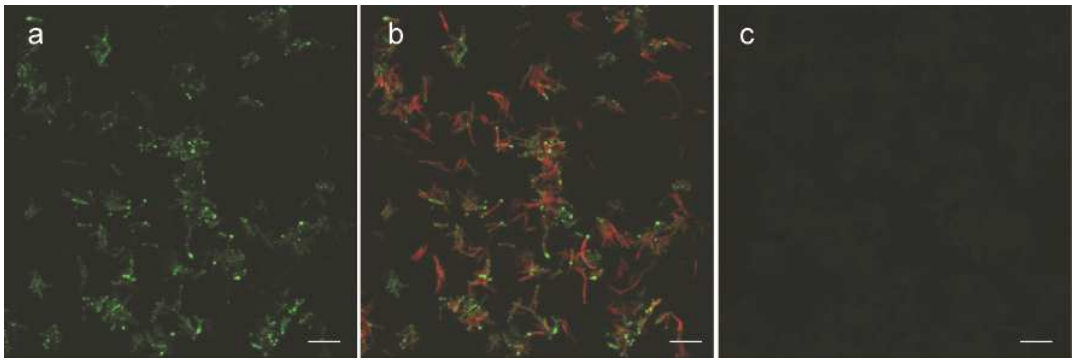


Figure 6.29. Induction of GFP-tet-H37Rv reveals modest fluorescence after 18 hours of growth. (a) H37Rv GFP strain after 72 hour induction anhydrotetracycline (ATc). (b) strain coexpresses RFP, which shows accumulation of GFP in inclusion bodies. (c) Background fluorescence in the absence of AhT. Scale bars 5 µm.

Previous work with Tet inducible strains of *Mtb* suggested that induction of the Ag85 strains infecting macrophages was possible[29]. It was hypothesized that induction of the Ag85 proteins within macrophages would afford better accumulation of the Ag85 proteins, as they would be trapped within the phagosomes of the macrophages. J774 macrophages were infected with all the GFP fusion protein strains, induced for 72 hours with anhydrotetracyline, and fixed, stained with and Anti-Tb antibody (Axel Yv54601) and analyzed by confocal microscopy. All strains were infected at the same multiplicity of infection (MOI, 5-10 *Mtb*/macrophage) and fluorescent *Mtb* were visible with all infected strains (**Figure 6.30, Figure 6.31**).

Far fewer fluorescent bacteria were observed in the Ag85-GFP (**Figure 6.31**) strains compared with the LacZ-GFP strain (**Figure 6.30**). Because LacZ-GFP is not secreted and instead is trapped within the *Mtb*, all bacteria show up labelled. In contrast, the Ag85 can be shed from the outer portion of the bacteria and trafficked by the macrophage. The paucity of Ag85-GFP positive *Mtb* can be rationalized either by detection of only the bacteria that are currently growing and producing protein at the time of harvest or of those bacteria that produced the greatest amount of Ag85-GFP. An additional explanation was that the overexpressed Ag85-GFP was somehow toxic to the mycobacteria within macrophages,

causing the plasmid to be lost over the 72 hour infection/induction period. This explanation seems unlikely to us, as the LacZ-GFP expression was very stable. Future plating and PCR experiments will be employed to rule out plasmid loss.

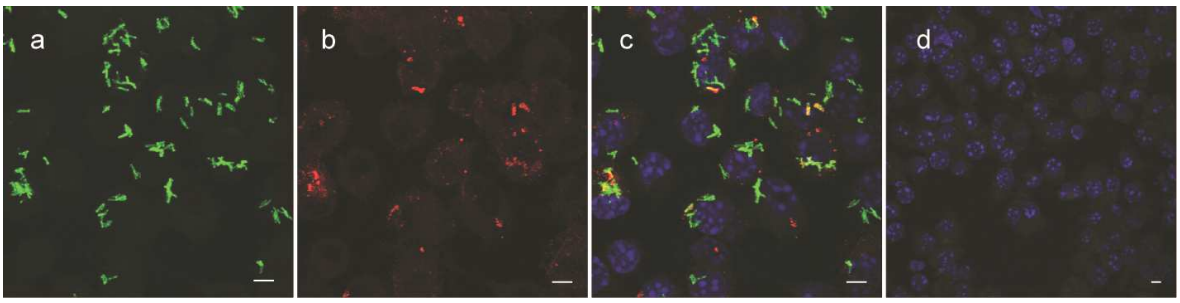


Figure 6.30. (a) LacZ-GFP expressing strain infecting J774 macrophages (b) Anti-tb antibody (Axel Yv54601) (c) Overlay of 'a-b' with DAPI staining the nuclei overlay of 'a-b' show reasonable colocalization of antibody and bacteria. Antibody staining has been previously shown to not label all bacteria with equal avidity. (d) Background signal of macrophages infected with LacZ-GFP expressing strain but not induced with ATc and not labelled with antibody. Scale bars 5 µm.

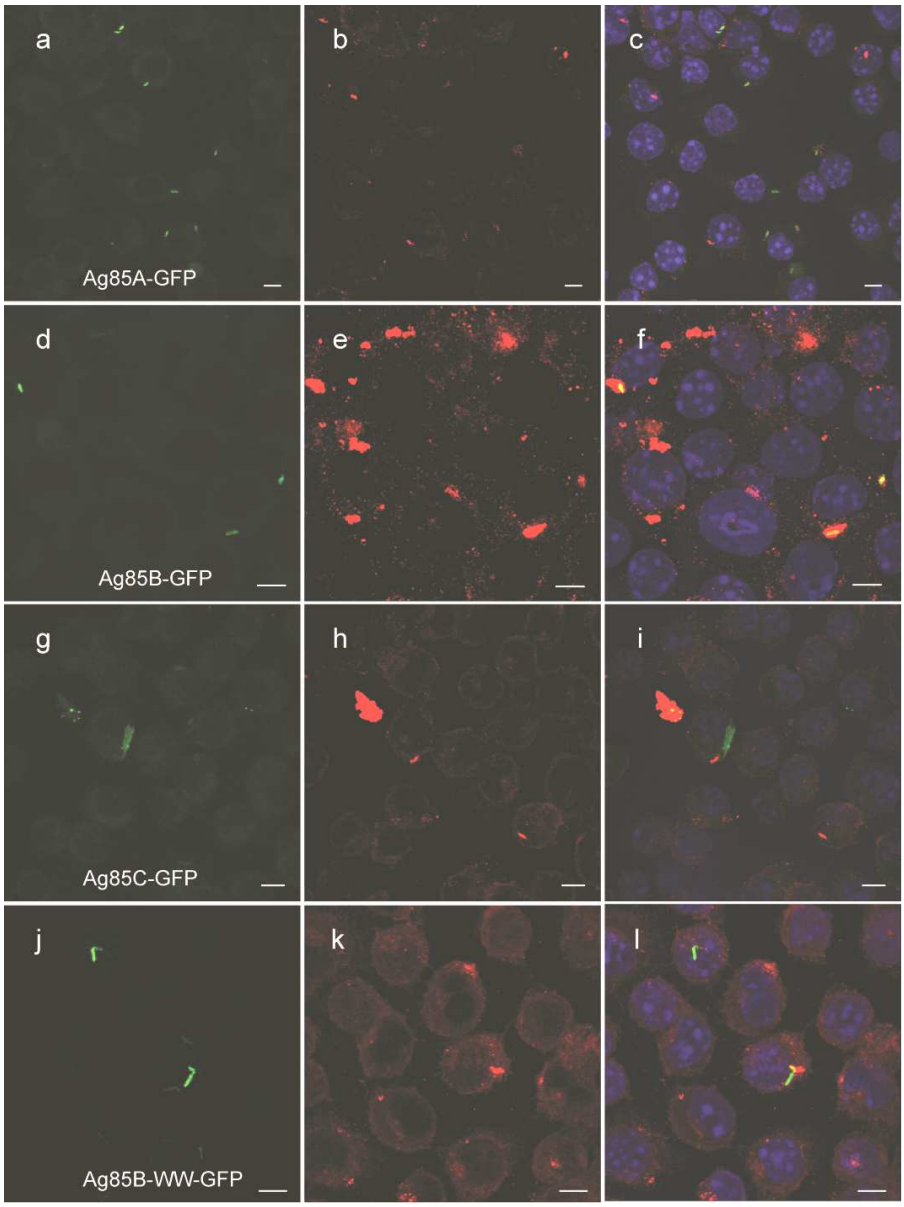


Figure 6.31. Overepxression of Ag85-GFP fusions in J774 macrophages demonstrates bacterial labelling. (a,d,g,j) Ag85-GFP expressing strains shows fluorescent regions of a size slightly smaller than entire *Mtb* (b,e,h,k) Anti-TB antibody (Axel Yv54601 (c,f,i,l) Overlay of 'a-b, d-e, g-h, j-k' with DAPI staining the nuclei shows reasonable colocalization of antibody and bacteria. Antibody staining has been previously shown to not label all bacteria with equal avidity.

It was hoped that the Anti-TB antibody would identify the *Mtb* localization of the Ag85. Consistent with the previous problems with the *Mtb* antibodies (**Chapter 5**), incomplete antibody labelling was observed with some bacteria staining intensely and others not detected by the antibody. This poor labelling precluded the detection of subcellular Ag85 localization. Qualitatively, the Ag85-GFP fluorescence of the Ag85 strains appeared much shorter than the

length of an entire bacterium, suggesting that the signal localized to only part of the *Mtb.* Strains that both express the Ag85-GFP Tet vectors and carry a constitutive RFP vector have been generated and future work with these dual labelled bacteria will shed light on the Ag85 localization.

6.2.9.1 MIC analysis of overexpression strains

To more fully evaluate the influence of Ag85 overexpression on cell wall biosynthesis, the minimum inhibitory concentrations (MICs) against the Ag85 Tet inducible strains of antimycobacterial drugs that target mycolate synthesis. It was hypothesized that the overexpression of the Ag85 proteins would alter the flux of mycolates to the cell wall, resulting in bacteria that were either more sensitive or more resistant to Isoniazid (INH) and SQ109. INH indirectly targets FASII of fatty acid synthesis, abolishing mycolate synthesis. SQ109, which targets TDM synthesis, through an unknown function (postulated to target lipid transporter MMPL3), was evaluated as another candidate compound that should show differential MIC.

MIC assays in *Mtb* Tet strains, including the Ag85B WW mutant and the catalytic serine-alanine mutants for Ag85A ,B and C did not reveal the expected rescue upon Tet induction. IN contrast, these experiments did show a modest growth inhibition under Ag85B and Ag85C induction and particularly under the WW mutant of Ag85B during treatment with both INH and SQ109 (See **Appendix 2** for MIC plates). These results are consistent with the INH and SQ109 induced downregulation of Ag85B observed through microarray analysis. The Ag85B-GFP WWL expressing strain has been generated and we propose that the greater catalytic activity of this form of Ag85B may cause a more significant growth inhibition in future MIC assays. The relatively slow induction of the pTE-2M1L strains is also under consideration as a limiting factor for the MIC assay. MIC assays of the constitutively expressing His-tagged Ag85 have been proposed, as these strains may produce the target proteins in larger quantities. However, one drawback of the constitutive strains is that they may have acquired other genetic changes to compensate for the protein overexpression. Overall, these MIC experiments are suggestive of differences in the Ag85 enzymes' abilities to rescue and/or sensitive mycobacteria from INH and SQ109 treatment and future work should shed additional light on precisely how the growth modulations occur.

6.2.9.2 TDM/TMM profiling of Tet inducible GFP fusion proteins in H37RV

A survey of the mycolate literature revealed that overexpression of mycolyl transferases in *Corynebacteria* results in a hyperincorporation of mycolates into the cell wall[30]. Complementation with *Corynebacteria*'s major mycolyl trasferase PS1 led to both an over-incorporation of cell wall mycolates and an increase in the trehalose-dicorynemycolates (TDCM) levels and corresponding decrease in the trehalose-monocorynemycolate (TMCM) levels. In contrast, overexpression of the Ag85 proteins in *Corynebacteria* afforded the over-incorporation of mycolates into the cell wall but did not affect the TDMCM and TMCM levels. In a related study the same authors found that deletion of Ag85A, B did not cause a cell wall mycolyl defect and that overexpression of Ag85A, B and C could incompletely complement for the mycolate loss observed in the *ag85C* deletion strains. This study considered that the defect might not be in the gross level of mycolates but in the linkage of the mycolate to the arabinan. However, they found that the Ag85 enzymes were not linkage specific and each enzyme transfers mycoloyl residues onto the four sites present in the arabinogalactan non-reducing end hexaarabinosides[2]. These results suggest that the Ag85 enzymes are only partially redundant and can differentially modulate the incorporation of the mycolates into the cell wall.

The over-incorporation of mycolates into *Corynebacteria* by the *Mtb* Ag85 proteins prompted us to differentiate between Ag85A, B and C through lipid profiling both in *M. smegmatis* and in *Mtb*. Tet inducible constructs were evaluated with ^{14}C-acetate labeling, under either a pre- or co-label with the radioisotope. During pre-label, culture was first exposed to the radiolabel, washed to remove label and then exposed to Tet induction. Under co-label, radiolabel and anhydrotetracyline were added simultaneously. Following a labelling period, the lipids were extracted through the standard protocol and evaluated through TLC(see **Appendix 2** for details). No dramatic differences between the strains were observed in the TDM and TMM profiles in *M. smegmatis*. Greater accumulation of radiolabel into TDM relative to TMM in the pre-label, compared with the co-label of all strains, supported that the flux of the pathway is from TMM to TDM. INH was utilized as a control to abolish mycolate production and to identify the TDM and TMM TLC bands (**Figure 6.32**).

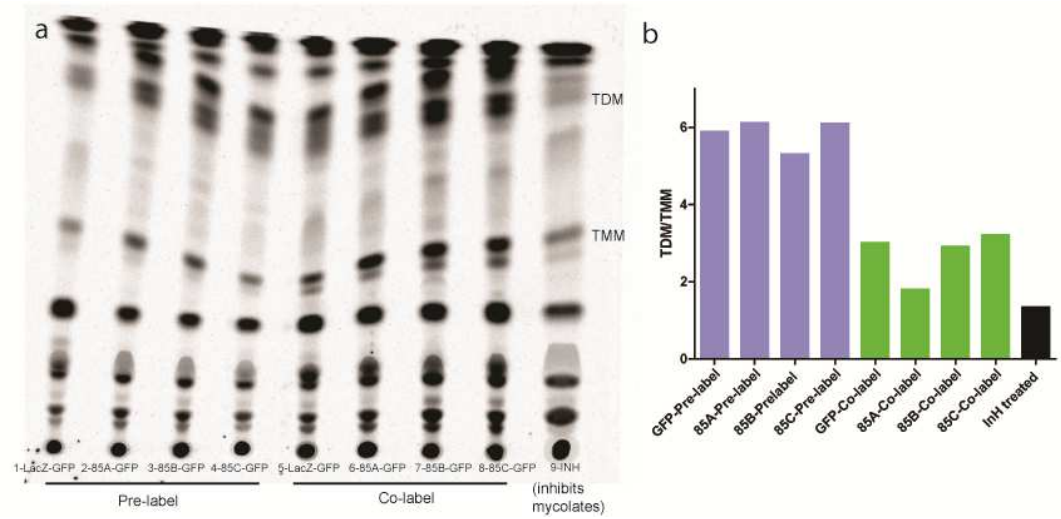


Figure 6.32. (a) TLC of TDM and TMM polar lipid extract from Tet inducible constructs in *M. smegmatis.* lanes 1-4 show extractable lipids that were pre-labeled with ^{14}C-acetate and subsequently exposed to ATc protein overexpression. Lanes 5-6 show extractable lipids from co-labeled bacteria treated with both ATc and ^{14}C-acetate concurrently. Lane 9-Isoniazid (INH) treatment reduces TMM and TDM levels. (b) Quantification of the ratio of TDM/TMM radioactivity shows an increase in the ratio under pre-label, compared with co-label.

Differences were observed in the covalent mycolate profile (mycolates that remained attached after extractable lipids were removed) in *M. smegmatis* (**Figure 6.33**). Under a pre-label, both Ag85C-GFP and Ag85A-GFP accumulated more radiolabel in the lipids than did Ag85B-GFP or LacZ-GFP. This result is consistent with Ag85A and C as responsible for the attachment of mycolates to the cell wall. In constrast with the pre-label, under a co-label, induction of Ag85B-GFP led to an increase in the covalent mycolates. However, it remains to be seen how Ag85B facilitates the transfer of mycolates. Possibly Ag85B is not directly catalyzing the transfer reaction, but is somehow involved in the flux of the pathway.

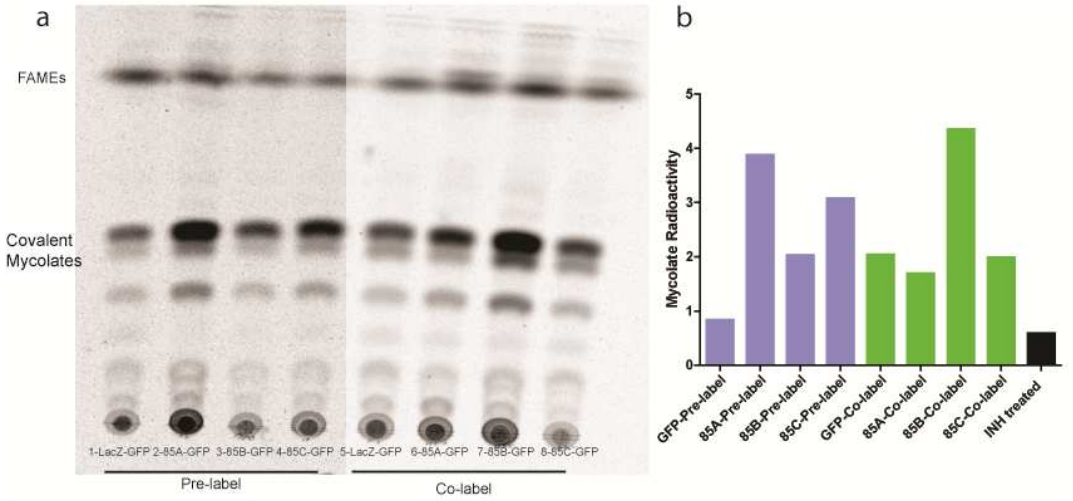


Figure 6.33. (a) Covalent mycolates from *M. smegmatis* labelling experiment were saponified and evaluated by TLC. Gain was adjusted for Co-label panel to increase contrast. (b) Quantification of mycolate TLC shows differential incorporation with Ag85B (Lane-7) overexpression resulting in mycolate increase under co-label, whereas, Ag85A and C (Lanes 2 and 4) lead to increased mycolation under pre-label conditions. INH treatment showed dramatic drease in radiolabel incorporation. Radioactivity was quantified in typhoon phosphorimager units and normalized by the fatty acid methyl esters (FAMEs).

Lipid profiling of the *Mtb* H37Rv Tet inducible strains was conducted in a manner analogous to the *M. smegmatis* experiments, albeit with appropriately longer induction and labeling times. Both pre-label and co-label experiments were conducted and the results are largely comparable to those reported in *M. smegmatis*. Induction of the specific Ag85A, B, or C GFP fusion, concurrent with ^{14}C-acetate co-label leds to a quantifiable increase in the cell wall linked mycolates (**Figure 6.** 34). The level of incorporation with Ag85B overexpression was slightly greater than A or C, consistent with, but more modest than, the *M. smegmatis* results (**Figure 6.33**). Under pre-label conditions Ag85C showed much less incorporation, which was inconsistent with the earlier results in *M. smegmatis*. This difference could be due to the longer radioisotope pre-label time (3 days compared with 3 hours). In the future we will conduct a shorter pulse labeling with ^{14}C-acetate, followed by protein induction.

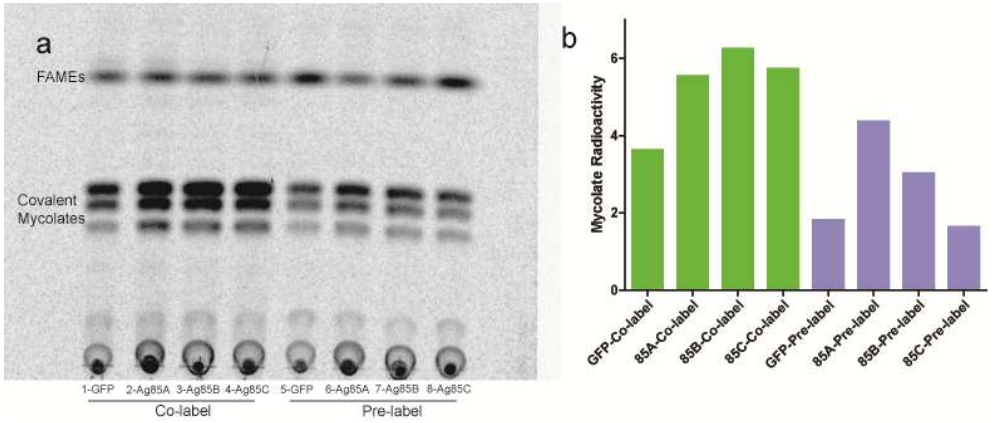


Figure 6. 34. (a) TLC of covalent mycolate saponification in *Mtb* reveals that the Ag85 enzymes can differentially incorportate the mycolates into the cell wall. (b) Quantitation of mycolates reveals that Ag85A, B and C overexpression all led to increased mycolation of the cell wall (lanes 2-4). Ag85A and B overexpression under pre-label conditions (Lanes 6,7) led to cell wall mycolation increase. Radioactivity was quantified in typhoon phosphorimager units and normalized by the fatty acid methyl esters (FAMEs).

Analysis of the TDM and TMM lipid extracts suggested that in *Mtb* the Ag85 enzymes differentially modulate the TDM and TMM levels. Overexpression of Ag85A led to an increase in the TDM/TMM ratio and Ag85B had the opposite affect, driving the formation of additional TMM (**Figure 6.35**). If these results prove consistent in future experiments, they could have implications in the TDM TMM pathway flux. An Ag85B driven flux towards TMM would lead to greater incorporation of mycolates into the cell wall, consistent with the results observed.

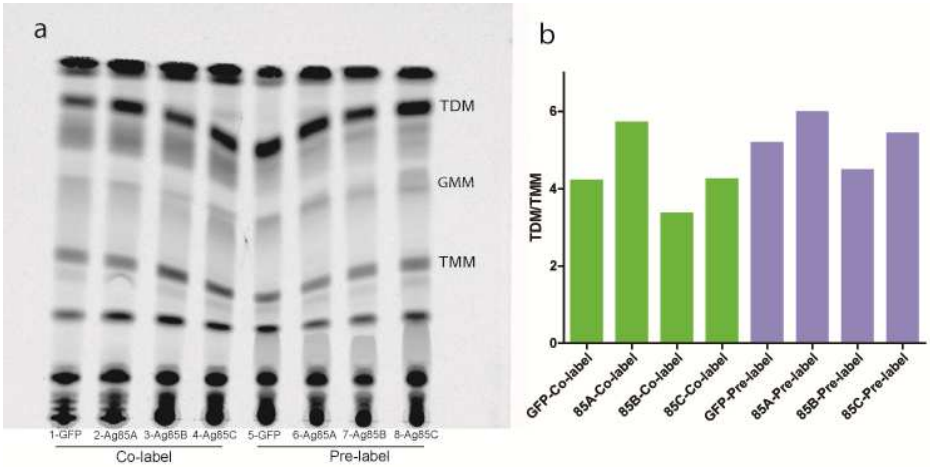


Figure 6.35. (a) TLC of extractable lipids from pre-label and co-label of H37Rv Tet inducible strains. GMM- glucose monomycolate. (b) Quantification of the ratio of TDM/TMM radioactivity shows a modest increase in the ratio under pre-label, compared with co-label. Ratio of TDM/TMM is decreased in Ag85B overexpression strain, which may have implications in lipid trafficking.

6.3 Conclusions

Chapter 6 has identified differences between the three isoforms of Ag85 both in enzyme assays and in cell culture. Phylogenetic studies revealed several highly conserved residues in a putative secondary site of the Ag85A, B and C proteins that differ between the protein isoforms. Molecular dynamics and mutagenesis studies demonstrated that this secondary site of the Ag85 proteins is critical for catalytic activity. Replacement of three key residues in Ag85B with their C amino acid afforded a large increase in Ag85B catalytic activity *in vitro*. Conversely, substitution of the Ag85B residues into the Ag85C proteins triggered a dramatic loss of activity. Ag85B has been identified as a dimer, whereas Ag85A and C are largely monomeric. The same secondary site that regulated activity also affected the Ag85B dimerization. Sequence analysis also led to the identification and purification of a novel trehalose acyl transferase, Rv0519c. Purification attempts at a second putative mycolyl transferase, Rv1288 are ongoing. Enzymatic assays of the Ag85 proteins have revealed that several key secondary site residues explain the dramatic differences in enzyme activity.

In cell culture, *Mtb* strains expressing tetracycline inducible Ag85-GFP fusion proteins have been generated and reveal differences between Ag85A, B and C. Preliminary results suggested that the Ag85 proteins can differentially inhibit *Mtb* growth during INH and SQ109 drug treatment. Lipid profiling revealed that Ag85-GFP protein overexpression results in hyperincorporation of th mycolates into the cell wall and that Ag85B may play an orthogonal role to both Ag85A and C. Microscopic subcellular localization studies of the Ag85 proteins have confirmed that proteins are highly secreted and these experiments were suggestive of greater Ag85A/C cell wall association. In summary, these results suggest that the Ag85 enzymes are not functionally redundant, but instead serve complementary roles in mycobacterial cell wall biosynthesis.

6.4 References

1. Sathyamoorthy, N. & Takayama, K. Purification and characterization of a novel mycolic acid exchange enzyme from *Mycobacterium smegmatis*. *J. Biol. Chem.* **262**, 13417-23 (1987).
2. Puech, V. et al. Evidence for a partial redundancy of the fibronectin-binding proteins for the transfer of mycoloyl residues onto the cell wall arabinogalactan termini of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **44**, 1109-1122 (2002).
3. Armitige, L.Y., Jagannath, C., Wanger, A.R. & Norris, S.J. Disruption of the Genes Encoding Antigen 85A and Antigen 85B of *Mycobacterium tuberculosis* H37Rv: Effect on Growth in Culture and in Macrophages. *Infect. Immun.* **68**, 767-778 (2000).
4. Sung, N., Takayama, K. & Collins, M.T. Possible association of GroES and antigen 85 proteins with heat resistance of *Mycobacterium paratuberculosis*. *Appl. Environ. Microbiol.* **70**, 1688-97 (2004).

5. Harth, G., Lee, B.Y. & Horwitz, M.A. High-level heterologous expression and secretion in rapidly growing nonpathogenic mycobacteria of four major *Mycobacterium tuberculosis* extracellular proteins considered to be leading vaccine candidates and drug targets. *Infect. Immun.* **65**, 2321-8 (1997).
6. Orme, I.M., Andersen, P. & Boom, W.H. T Cell Response to *Mycobacterium tuberculosis*. *J. Infect. Dis.* **167**, 1481-1497 (1993).
7. Comas, I. et al. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat. Genet.* **42**, 498-503 (2010).
8. De Sousa-D'Auria, C. et al. New insights into the biogenesis of the cell envelope of corynebacteria: identification and functional characterization of five new mycoloyltransferase genes in *Corynebacterium glutamicum*. *FEMS Microbiol. Lett.* **224**, 35-44 (2003).
9. Pessolani, M.C. & Brennan, P.J. *Mycobacterium leprae* produces extracellular homologs of the antigen 85 complex. *Infect. Immun.* **60**, 4452-4459 (1992).
10. Wixon, J. Featured Organism: Reductive evolution in bacteria: *Buchnera* sp., *Rickettsia prowazekii* and *Mycobacterium leprae*. *Comp. Funct. Genomics* **2**, 44-48 (2001).
11. Ronning, D.R. et al. Crystal structure of the secreted form of antigen 85C reveals potential targets for mycobacterial drugs and vaccines. *Nat. Struct. Biol.* **7**, 141-6 (2000).
12. Anderson, D.H., Harth, G., Horwitz, M.A. & Eisenberg, D. An interfacial mechanism and a class of inhibitors inferred from two crystal structures of the *Mycobacterium tuberculosis* 30 kDa major secretory protein (Antigen 85B), a mycolyl transferase. *J. Mol. Biol.* **307**, 671-81 (2001).
13. Lomize, A., Lomize, M. & Pogozheva, I. Orientations of proteins in membranes (OPM) database. Vol. 2011 (2005).
14. McPhie, P. Concentration-independent estimation of protein secondary structure by circular dichroism: a comparison of methods. *Anal. Biochem.* **375**, 379-381 (2008).
15. Barry, C.S., Backus, K.M., Barry, C.E. & Davis, B.G. ESI-MS Assay of *M. tuberculosis* Cell Wall Antigen 85 Enzymes Permits Substrate Profiling and Design of a Mechanism-Based Inhibitor. *J. Am. Chem. Soc.*, ASAP (2011).
16. Lim, J.-H. et al. Purification and Immunoreactivity of Three Components from the 30/32-Kilodalton Antigen 85 Complex in *Mycobacterium tuberculosis*. *Infect. Immun.* **67**, 6187-6190 (1999).
17. Brand, S., Niehaus, K., Pühler, A. & Kalinowski, J. Identification and functional analysis of six mycolyltransferase genes of *Corynebacterium glutamicum* ATCC 13032: the genes cop1, cmt1, and cmt2 can replace each other in the synthesis of trehalose dicorynomycolate, a component of the mycolic acid layer of the cell envelope. *Arch. Microbiol.* **180**, 33-44 (2003).
18. Takayama, K., Wang, C. & Besra, G.S. Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin. Microbiol. Rev.* **18**, 81-101 (2005).
19. Wu, S. & Zhang, Y. LOMETS: A local meta-threading-server for protein structure prediction. *Nucleic Acids Res.* **35**, 3375-3382 (2007).
20. Bateman, A. & Bycroft, M. The structure of a LysM domain from E. coli membrane-bound lytic murein transglycosylase D (MltD). *J. Mol. Biol.* **299**, 1113-1119 (2000).
21. Boshoff, H.I.M. et al. The Transcriptional Responses of *Mycobacterium tuberculosis* to Inhibitors of Metabolism. *J. Biol. Chem.* **279**, 40174-40184 (2004).
22. Kana, B.D. & Mizrahi, V. Resuscitation-promoting factors as lytic enzymes for bacterial growth and signaling. *FEMS Immunol. Med. Microbiol.* **58**, 39-50 (2010).
23. Kim, S.K. et al. Biochemical and structural characterization of the secreted chorismate mutase (Rv1885c) from *Mycobacterium tuberculosis* H37Rv: an *AroQ enzyme not regulated by the aromatic amino acids. *J. Bacteriol.* **188**, 8638-48 (2006).
24. Schneider, C.Z., Parish, T., Basso, L.A. & Santos, D.S. The two chorismate mutases from both Mycobacterium tuberculosis and *Mycobacterium smegmatis*: biochemical analysis and limited regulation of promoter activity by aromatic amino acids. *J. Bacteriol.* **190**, 122-34 (2008).
25. Rajagopalan, M. et al. *Mycobacterium tuberculosis* origin of replication and the promoter for immunodominant secreted antigen 85B are the targets of MtrA, the essential response regulator. *J. Biol. Chem.* **285**, 15816-27 (2010).
26. Cowley, S.C. & Av-Gay, Y. Monitoring promoter activity and protein localization in *Mycobacterium* spp. using green fluorescent protein. *Gene* **264**, 225-231 (2001).
27. Posey, J.E., Shinnick, T.M. & Quinn, F.D. Characterization of the Twin-Arginine Translocase Secretion System of *Mycobacterium smegmatis*. *J. Bacteriol.* **188**, 1332-1340 (2006).
28. Wiker, H.G., Harboe, M. & Nagai, S. A localization index for distinction between extracellular and intracellular antigens of *Mycobacterium tuberculosis*. *J. Gen. Microbiol.* **137**, 875-884 (1991).
29. Ehrt, S. et al. Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor. *Nucleic Acids Res.* **33**, e21 (2005).

30. Puech, V., Bayan, N., Salim, K., Leblon, G. & Daffé, M. Characterization of the *in vivo* acceptors of the mycoloyl residues transferred by the corynebacterial PS1 and the related mycobacterial antigens 85. *Mol. Microbiol.* **35**, 1026-41 (2000).

7 Chapter 7: Conclusions and further considerations

Despite decades of investigation into both their immunologic and enzymatic activities, the Antigen 85 enzymes (Ag85A,B and C) remain poorly understood. Little is known about the range of their substrates and the reason for their apparent functional redundancy in *Mtb*. This thesis sought to address this paucity of knowledge through an evaluation of AG85 enzyme functions both *in vitro* and *in vivo*. These studies resulted in a novel Ag85 activity assay, the development of a potent Ag85 inhibitor, the synthesis and evaluation of the first small molecule TB-selective probe and a better understanding of the differential functionalities of these important enzymes.

Synthesis and uptake studies of trehalose (Tre) and a library of Tre analogues were conducted. The uptake of ^{14}C-trehalose into *Mtb in vitro* and into infected macrophages suggested that Tre analogues may serve as novel probes of TB pathogenesis. A library of trehalose analogues, featuring modifications at every position of the sugar scaffold, has been synthesized. The optimization and use of a methyl ketoside based glycosylation afforded a rapid means to diversify the trehalose scaffold by modifications such as fluorine, fluorescein, biotin, quantum dots and other biologically relevant groups. This trehalose library should prove useful in future studies of mycobacterial trehalose catabolism/metabolism and Ag85 activity.

A novel, high throughput mass spectrometry-based Ag85 substrate assay was developed, and this assay revealed that Ag85A, B and C will tolerate extensive modifications on the trehalose scaffold, processing nearly all trehalose derivatives synthesized, including those with bulky groups and previously reported inhibitors. Substrate screening revealed that the antigen 85 enzymes have a remarkably broad tolerance towards modifications of the trehalose scaffold. Affinity for disaccharides, compared with monosaccharides, was confirmed, although some monosaccharides were transesterified at a modest rate. Glucopyranose and arabinofuranose configurations were preferred over galacto-, manno- or xylo-pyranose. These results support that the Ag85 complex is responsible for both trehalose-mycolate scrambling and mycolylation of the *Mycobacterial* arabinogalactan.

Mass spectrometry also provided an accurate and automated method to determine the full kinetic parameters of Ag85C and to screen for enzyme inhibitors. Kinetic analysis revealed that Ag85C displays a combination of acyltransferase and acylhydrolase activities. Consistent

with prior reports, Ag85C was found to be markedly more active than Ag85A or B. The substrate screening and kinetic data were used to design fluorophosphonate-trehalose, which to our knowledge is the first potent, selective, covalent inhibitor of Ag85. We anticipate that this compound may find use as an *in vitro* chemical probe to inhibit Ag85 activity, and it will possibly serve as a starting point for the design of novel anti-*Mycobacterial* drugs.

Enzyme assays revealed that fluorescein-trehalose (FITC-Tre) was an efficient substrate of all three Ag85 enzymes. This fluorescent compound was found to label the cell wall of mycobacteria and is the first small molecule *Mtb*-selective probe. Lipid profiling confirmed that the labelling process was specific and Ag85 mediated. FITC-Tre labels the trehalose dimycolate (TDM) lipids and labelling decreases in the *ag85C* knockout strain of *Mtb.* The ability to fluorescently label lipids attached to trehalose should be useful in identifying other pathways that may employ Ag85 transesterification activity and help clarify this poorly understood area of biochemistry.

FITC-Tre was also found to selectively label mycobacteria residing within macrophages. Prior imaging work with *Mtb* in infected macrophages has utilized several strategies[1,2], but none allow for non-toxic imaging of bacteria *in vivo*. Bacteria have been labeled with Texas-Red through oxidation with sodium periodate prior to infection. These harsh labeling conditions are nonselective and would be expected to oxidize structurally important carbohydrates and immunomodulatory glycolipids[1]. Fluorescently-labeled vancomycin has been used to track the cell division patterns of *M. smegmatis* and *M. bovis* BCG[2], but its toxicity towards mycobacteria[3] may color experimental results. Antibodies can be utilized to label fixed *Mtb* and for enzyme linked immunosorbant assays (ELISA)[4], but are not useful for live imaging. FITC-Tre does not perturb mycobacterial growth, and presumably natural functions, allowing microscopy of healthy, viable bacteria. Unlike other possible probes the FITC-Tre mode of action is based on the activity of enzymes unique mycobacteria. Given the highly conserved nature of the Ag85 proteins[5], in mycobacteria (and the complete absence of Tre in mammalian biology) it is likely this probe would also label non-tuberculous mycobacteria as well as other members of the tuberculosis complex.

Understanding of the macrophage infection process is incomplete[6]. The ability to follow viable *Mtb* during infection would help characterize the bacterial transit to the phagosome and other intracellular compartments. The co-localization studies between FITC-Tre-labelled *Mtb* and various markers of phagosome maturation suggest that FITC-Tre may preferentially

label those bacteria that resist acidification and phagosome-lysosome fusion. An alternative hypothesis that cannot be overlooked is that incorporation of FITC-Tre into the membrane of a subset of bacilli alters phagosome maturation, as has been seen with *Brucella pertusis* prelabelled with FITC before infection[7]. Because infection was established prior to FITC-Tre addition, the influence of the FITC-Tre-dimycolate should be slight. Other trehalose-tagged molecules, potentially labelled at different positions or different tags on the trehalose scaffold may also prove useful as additional *Mtb* probes. Future work with FITC-Tre and related compounds will facilitate evaluation of the probe's *in vivo* potential as a possible diagnostic tool to label *Mtb* in an infected host. Preliminary studies with ^{18}F-fluorobenzyl-trehalose suggest utility of trehalose as a mycobacteria-selective PET probe. The broad substrate tolerance of the Ag85 proteins suggests the possibility of trehalose probes for many imaging modalities.

The discovery of the breadth of substrates processed by Ag85 isoforms allowed for mycobacteria-selective labelling *in vivo.* These results may also have implications for additional, uncharacterized enzyme functions. Although the Ag85 proteins are highly conserved, both according to sequence[5] and structural homology[8], across all species of mycobacteria, the individual Ag85 isoforms vary widely in their catalytic efficiency and relative quantities secreted[9,10]. Numerous glycolipids within *Mtb* contain a trehalose core, including, TDM, TMM, pentaacyl trehalose (PAT), triacyl trehalose (TAT) and sulfolipid-1 (Sl-1); little is known about the final acylation steps of PAT, TAT, and SL-1 and the enzymes of the Ag85 complex have been postulated to be involved in the lipidation of these compounds[11,12]. The substrate plasticity, particularly at the C-2 position, may also have implications in the final acylation steps of sulfolipid-1, which are unknown, but have been postulated to potentially occur *via* antigen 85[11].

A full analysis of the relative functional differences of the mycobacterial Ag85 proteins is ongoing. Antigen 85 substrate screening revealed Ag85 isoform-specific differences in substrate preference, hinting at subtle differences between the Ag85A,B and C active site architecture. Mutagenesis studies revealed that mutation of three amino acids (WWL) from Ag85C into Ag85B at a secondary site, results in a dramatic increase in Ag85B catalytic activity. Ag85B was found to exist largely as a dimer, whereas Ag85A and C are mainly monomeric. Mutation of the same WWL amino acids, which exist at the putative dimer

interface, cause Ag85B dimer dissociation. A secondary site on Ag85B has been identified as critical for regulating enzyme activity and dimeric state.

The studies of the Ag85 enzyme's differential function are a work in progress and additional analysis of these enzymes should afford a more complete understanding of their roles in mycobacteria. Future work with the Ag85 secondary site will focus on a complete analysis of the kinetics and general function of this portion of the enzyme. Calculation of the Ag85B dimer dissociation constant (K_D), through the standardbis(sulfosuccinimidyl)suberate (BS3) crosslinking experiment[13], will yield a better understanding the nature and stability of this dimer. Binding studies with arabinan fragments from mycobacteria, alongside similar studies with synthesized arabinan docosaccharide moity from Ito et al[14] should dissect whether the Ag85B dimer is important for arabinan binding. Binding studies with Ag85C and the Ag85B secondary site mutants will help link the Ag85 secondary site to the *in vivo* enzyme activities. Full kinetic analysis of Ag85B and the Ag85B WWL mutant is in progress; these assays will unveil whether the increase in activity of Ag85B WWL, compared with wildtype is due to an increase in substrate binding affinity (K_M) or catalytic efficiency of the enzyme (K_{cat}). These experiments should link the discovery that Ag85B exists as a dimer and is regulated by a secondary site at the dimeric interface to a larger understanding of the differences between Ag85A,B and C.

Lipid profiling of mycobacterial strains that overexpress individual Ag85 enzymes has revealed isoform dependent, differential modulation of TDM, TMM and cell wall associated mycolates. Ag85B overexpression, concurrent with ^{14}C-acetate labelling, results in the greater increase in cell wall mycolation relative to Ag85A or C overexpression strains. Ag85B overexpression also appears to increase the levels of TMM relative to TDM in the cell wall. Increase in TMM may lead to an increased availability of mycolates for cell wall transesterification. Previous reports in *Corynebacteria* have documented incomplete redundancy of related corynmycolyltransferases[15], with non-redundant modulation of the TDM/TMM and mycolates. Knockouts of the individual *ag85A,B* and *C* have been generated in Mtb, but the double and triple knockouts have not been reported. Quantification of the mycolate levels in these deletion strains suggested that the *ag85C* deletion strain displays a 40% decrease in cell wall associated mycolates, whereas the Ag85A and B strains feature wildtype mycolate levels[16,17]. Preliminary lipid profiling of *Mtb* strains over expressing Ag85-GFP fusion proteins suggests that these enzymes can uniquely regulate the lipid

profiles. Future lipid profiling experiments of the Mtb over expression strains, including the catalytically inactive serine-alanine mutants and the Ag85B secondary site mutants, will further the dissection of the Ag85 differential activities. We anticipate that these lipid profiles will link the *in vitro* enzyme activity (dimerization and secondary site) with functions within cell wall synthesis.

The Ag85 proteins are the most abundant secreted mycobacterial proteins, are highly antigenic and serve important roles in cell wall biosynthesis. Production of trehalose mono and dimycolate (TMM and TDM), catalyzed by Ag85C, is one of the primary causes of mycobacterial survival, granuloma formation and, more generally, the persistence of tuberculosis infection within the human population. The Antigen 85 analysis presented in this thesis has resulted in the synthesis of a potent, covalent inhibitor of the Ag85 proteins, the development of FITC-Tre, a small molecule, mycobacteria-selective probe and the a more complete understanding of the differential activities of Ag85A,B and C. Because the Ag85 proteins and their substrates serve vital roles in *Mtb*, a full understanding of their enzymology and differential activities is important for the diagnosis and treatment of tuberculosis.

7.1 References

1. Beatty, W.L. et al. Trafficking and release of mycobacterial lipids from infected macrophages. *Traffic* **1**, 235-47 (2000).
2. Thanky, N.R., Young, D.B. & Robertson, B.D. Unusual features of the cell cycle in mycobacteria: polar-restricted growth and the snapping-model of cell division. *Tuberculosis (Edinb.)* **87**, 231-6 (2007).
3. Provvedi, R., Boldrin, F., Falciani, F., Palu, G. & Manganelli, R. Global transcriptional response to vancomycin in *Mycobacterium tuberculosis*. *Microbiology* **155**, 1093-1102 (2009).
4. Glatman-Freedman, A., Martin, J.M., Riska, P.F., Bloom, B.R. & Casadevall, A. Monoclonal antibodies to surface antigens of *Mycobacterium tuberculosis* and their use in a modified enzyme-linked immunosorbent spot assay for detection of mycobacteria. *J. Clin. Microbiol.* **34**, 2795-2802 (1996).
5. Cole, S.T. et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537-544 (1998).
6. Schorey, J.S., Carroll, M.C. & Brown, E.J. A Macrophage Invasion Mechanism of Pathogenic Mycobacteria. *Science* **277**, 1091-1093 (1997).
7. Weingart, C.L. et al. Fluorescent Labels Influence Phagocytosis of *Bordetella pertussis* by Human Neutrophils. *Infect. Immun.* **67**, 4264-4267 (1999).
8. Ronning, D.R., Vissa, V., Besra, G.S., Belisle, J.T. & Sacchettini, J.C. *Mycobacterium tuberculosis* antigen 85A and 85C structures confirm binding orientation and conserved substrate specificity. *J. Biol. Chem.* **279**, 36771-7 (2004).
9. Wiker, H.G., Harboe, M. & Nagai, S. A localization index for distinction between extracellular and intracellular antigens of *Mycobacterium tuberculosis*. *J. Gen. Microbiol.* **137**, 875-884 (1991).
10. Harth, G., Lee, B.Y., Wang, J., Clemens, D.L. & Horwitz, M.A. Novel insights into the genetics, biochemistry, and immunocytochemistry of the 30-kilodalton major extracellular protein of *Mycobacterium tuberculosis*. *Infect. Immun.* **64**, 3038-47 (1996).
11. Converse, S.E. et al. MmpL8 is required for sulfolipid-1 biosynthesis and *Mycobacterium tuberculosis* virulence. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 6121-6126 (2003).

12. Hatzios, S.K. et al. PapA3 Is an Acyltransferase Required for Polyacyltrehalose Biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **284**, 12745-12751 (2009).
13. Graziano, V., McGrath, W.J., Yang, L. & Mangel, W.F. SARS CoV Main Proteinase: The Monomer-Dimer Equilibrium Dissociation Constant. *Biochemistry (Mosc.)* **45**, 14632-14641 (2006).
14. Ishiwata, A. & Ito, Y. Synthesis of docosasaccharide arabinan motif of mycobacterial cell wall. *J. Am. Chem. Soc.* **133**, 2275-91.
15. De Sousa-D'Auria, C. et al. New insights into the biogenesis of the cell envelope of corynebacteria: identification and functional characterization of five new mycoloyltransferase genes in *Corynebacterium glutamicum*. *FEMS Microbiol. Lett.* **224**, 35-44 (2003).
16. Puech, V., Bayan, N., Salim, K., Leblon, G. & Daffé, M. Characterization of the *in vivo* acceptors of the mycoloyl residues transferred by the corynebacterial PS1 and the related mycobacterial antigens 85. *Mol. Microbiol.* **35**, 1026-41 (2000).
17. Puech, V. et al. Evidence for a partial redundancy of the fibronectin-binding proteins for the transfer of mycoloyl residues onto the cell wall arabinogalactan termini of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **44**, 1109-1122 (2002).

8 Appendix 1: Chemistry synthesis and characterization

8.1 General Experimental

Hydrogenations were either performed manually or with a Thales Nano H Cube®.

Optical rotations were measured on a Perkin-Elmer 241 polarimeter with a path length of 1.0 dm and are reported with implied units of 10^{-1} deg cm^2 g^{-1}. Concentrations (c) are given in g/100 mL.

Melting points (m.p.) were recorded on a Leica Galen III hot stage microscope equipped with a Testo 720 thermocouple probe and are uncorrected.

Proton nuclear magnetic resonance (^{1}H NMR) spectra were recorded on a Bruker DPX400 (400 MHz), a Bruker AV400 (400 MHz) or a Bruker AVII500 (500 MHz) spectrometer, as indicated. Carbon nuclear magnetic resonance (^{13}C NMR) spectra were recorded on a Bruker AV400 (100 MHz) spectrometer or on a Bruker AVII500 (125 MHz) spectrometer, as indicated. NMR Spectra were fully assigned using COSY, HSQC, HMBC and DEPT 135. All chemical shifts are quoted on the δ scale in ppm using residual solvent as the internal standard (^{1}H NMR: $CDCl_3$ = 7.26, CD_3OD = 4.87; DMSO-d_6 = 2.50 and ^{13}C NMR: $CDCl_3$ = 77.0; CD_3OD= 49.0; DMSO-d_6 = 39.5). Coupling constants (J) are reported in Hz with the following splitting abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, quin = quintet, and a = apparent.

Infrared (IR) spectra were recorded on a Bruker Tensor 27 Fourier Transform spectrophotometer using thin films on NaCl plates for liquids and oils and KBr discs for solids and crystals. Absorption maxima (ν_{max}) are reported in wavenumbers (cm^{-1}) and classified as strong (s) or broad (br).

Low resolution mass spectra (LRMS) were recorded on a Waters Micromass LCT Premier TOF spectrometer using electrospray ionization (ESI) and high resolution mass spectra (HRMS) were recorded on a Bruker MicroTOF ESI mass spectrometer. Nominal and exact m/z values are reported in Daltons.

Thin layer chromatography (TLC) was carried out using Merck aluminium backed sheets coated with $60F_{254}$ silica gel. Visualization of the silica plates was achieved using a UV lamp (λ_{max} = 254 nm), and/or acid dip (MeOH:H_2O, 1:1, 10% H_2SO_4) and/or ammonium molybdate 5% in 2 M H_2SO_4, and/or potassium permanganate (5% $KMnO_4$ in 1 M NaOH with 5% potassium carbonate). Column chromatography was carried out using BDH PROLAB® 40-63 mm silica gel (VWR). Mobile phases are reported in ratio of solvents (e.g. petrol:ethyl acetate, 4:1)

Anhydrous solvents were purchased from Fluka or Acros with the exception of dichloromethane and THF, which were dried over Alumina cartiges. All other solvents were used as supplied (Analytical or

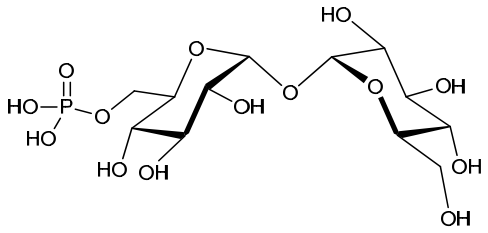
HPLC grade), without prior purification. Distilled water was used for chemical reactions and Milli-Q™ purified water for protein manipulations. Reagents were purchased from Sigma Aldrich and used as supplied, unless otherwise indicated. D-Trehalose (>99%) and D-arabinose (>99%) were purchased from Fluka. D-Glucose (>99.5%), UDP-glucose, 2-fluoro-2-deoxy-D-glucose and 3-fluoro-3-deoxy-D-glucose were purchased from Sigma. D-Glucosamine hydrochloride was purchased from Carbosynth. Qdot® 655 ITK™ amino (PEG) quantum dots (8 µM solution) were purchased from Invitrogen. 'Petrol' refers to the fraction of light petroleum ether boiling in the range 40–60 ºC. All reactions using anhydrous conditions were performed using flame-dried apparatus under an atmosphere of argon or nitrogen. 3 Å and 4 Å molecular sieves were activated by heating in a 400 °C furnace and were also employed for anhydrous reactions. Basic alumina refers to basic aluminum oxide and was utilized during some hydrogenation reactions. Brine refers to a saturated solution of sodium chloride. Anhydrous magnesium sulfate ($MgSO_4$) or sodium sulfate (Na_2SO_4) were used as drying agents after reaction workup, as indicated. DOWEX 50WX8 (H^+ form) was conditioned as follows: 100 g of the commercial resin was placed in a 500 mL sintered filter funnel and allowed to swell with 200 mL of acetone for 5 minutes. The solvent was removed by suction and the resin was washed successively with 800 mL of acetone, 500 mL methanol, 500 mL 5 M HCl, and then 1 L of water or until the pH of filtrate was ~ 7, as indicated by pH paper. The resin was partially dried on the filter and then stored and used as needed.

Protein Mass Spectrometry (LC–MS) was performed on a Micromass LCT (ESI TOF–MS) coupled to a Waters Alliance 2790 HPLC using a Phenomenex Jupiter C4 column (250x4.6 mm– 5µm). Water: acetonitrile, 95:5 (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid, were used as the mobile phase at a flow rate of 1.0 mL min^{-1}. The gradient was programmed as follows: 95% A (5 min isocratic) to 100% B after 15 min then isocratic for 5 min. The electrospray source was operated with a capillary voltage of 3.2 kV and a cone voltage of 25 V. Nitrogen was used as the nebulizer and desolvation gas at a total flow of 600L h^{-1}. Spectra were calibrated using a calibration curve constructed from a minimum of 17 matched peaks from the multiply charged ion series of equine myoglobin obtained at a cone voltage of 25 V. Total mass spectra were reconstructed from the ion series using the MaxEnt algorithm on MassLynx software (v. 4.0 from Waters) according to the manufacturer's instructions.

8.2 Synthetic Experimental Details

8.2.1 Trehalose-6 phosphate synthase (OtsA) synthesis of compounds 11 and 51.

α,α-D-trehalose 6-(dihydrogenphosphate) (11)[1,2]



Uridine-diphosphate-glucose **10** (82.86 mg, 0.135 mmol) and glucose-6-phosphate **9** (40mg, 0.11mmol) were added to a buffer of 500µl 100 mM TEA with 20 mM magnesium chloride. pH was equilibrated to 7.4 with HCl (5 M). 750 µl OTSA (31 mg/mL) was added at reaction was incubated at 27 ˚C overnight. Sample was then filtered using centrifugal concentrator Viva spin 500 (Sartorius) of nominal molecular weight cutoff 10,000 to remove proteins. Before purification, HPLC (Dionex Ultimate 3000) was purged with water and $NaHCO_3$ buffer and was calibrated with AMP (1 mg) in TEA $MgCl_2$ buffer. The desired product was purified by the HPLC using a strong anion exchange column Applied Biosystems, Poros® HQ (10 mm x 100 mm, 50 µm). The HPLC was eluted with gradient of 0 – 500 mM ammonium bicarbonate at a flow rate of 20 mL/min and eluants were detected with an evaporative light scattering detector (ELSD). The fractions from 3.5-5 minutes and 5.5-7 minutes were collected and dried separately on high vacuum. ESI showed produced in samples from 3.5-5 minutes. Sample was lyophilized overnight after freezing in liquid nitrogen. The desired product was obtained as a white, fluffy powder (8.7 mg, 15%).

TLC (water:isopropanol:ethyl acetate, 1:2:2, v/v): R_f=0; $[\alpha]_D^{18}$ + 150.3 (c = 1.0, H_2O), [lit. $[\alpha]_D^{21}$ + 151.2 (c = 0.8, H_2O)][2]; ^{1}H NMR (500 MHz, D_2O) δ ppm 3.35 (1 H, t, $J_{3,4}$ = 9.5Hz, $J_{4,5}$ = 9.5 Hz, H-4), 3.40 (1H, t, $J_{3',4'}$ = 9.3 Hz, $J_{4',5'}$ = 9.3 Hz, H-4'), 3.48 (1 H, dd, $J_{2',3'}$ = 10.7 Hz, $J_{1',2'}$ = 3.8 Hz, H-2'), 3.58 (1 H, dd, $J_{2,3}$ = 9.8 Hz, $J_{1,2}$ = 3.8 Hz, H-2), 3.66 (1 H, dd, $J_{6a',6b'}$ = 12.0 Hz, $J_{5',6b'}$ = 5.0 Hz H-6'b), 3.70-3.79 (4 H, m, H-3, H-3', H-5', H-6'a), 3.80 (1 H, m, H-5), 3.85 (1 H, ddd, $J_{6a,6b}$ = 12.3 Hz, $J_{6b,31P,}$ 5.7 Hz, $J_{5,6b}$ = 1.9 Hz, H-6b), 3.91 (1 H, ddd, $J_{6a,6b}$ = 13.4 Hz, $J_{6a,31P}$ = 6.2 Hz, $J_{5,6a}$ = 5.0 Hz, H-6a), 5.09 (1 H, d, $J_{1,2}$ = 3.8Hz, H-1). 5.12 (1 H, d, $J_{1'-2'}$ = 4.1 Hz, H-1'); ^{31}P NMR (162 MHz, DEUTERIUM OXIDE) δ ppm 3.11 (s, $ROP(O)(OH)_2$); MS *m/z* (ESI^-): 421.5 [$M-H^-$].

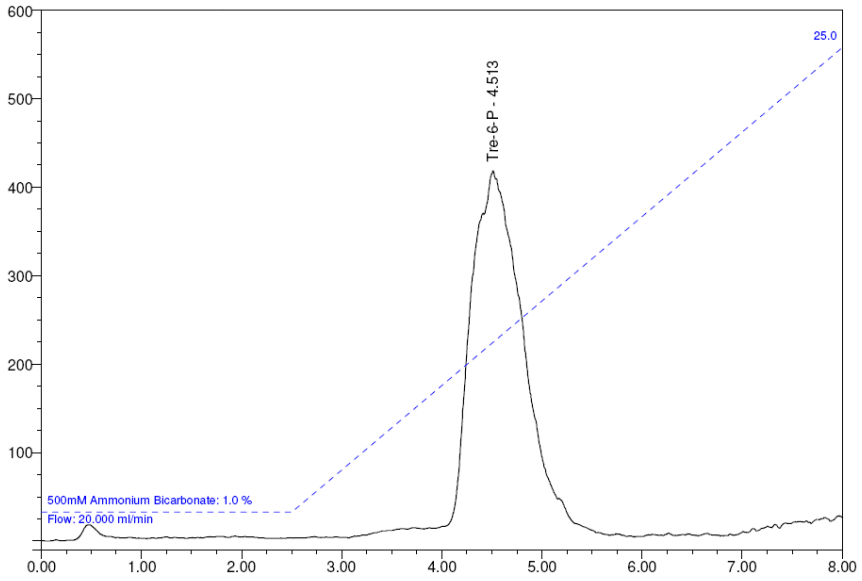
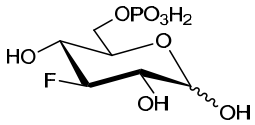


Figure 8.1. Preparative HPLC trace for the purification of 11. Conditions: Applied Biosystems, Poros® HQ strongly basic anion exchange column (10 mm x 100 mm, 50 µm). A gradient from 0 mM to 500 mM aqueous NH_4HCO_3 was used as the mobile phase at a flow rate of 20 mL/min. Compound 15 elutes at 4.5 min (Tre 6-P)

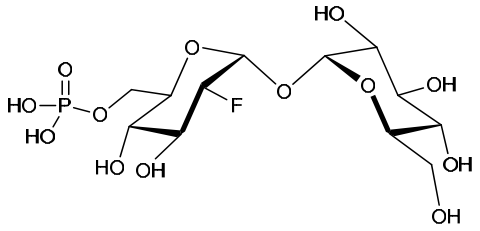
3-Fluoro-3-deoxy-D-glucose-6-phosphate (68)[3,4]



3-deoxy-3-fluoro-D-glucopyranose **67** (25 mg, 0.14 mmol) was mixed in an eppendorf tube with ATP (1 mg, 0.002 mmol), phosphoenolpyruvate (28.6 mg, 0.169 mmol) To this was added TEA buffer 20 mM with $MgCl_2$ (10 mM) and KCl (5 mM) in distilled water. pH was equilibrated to 7.6 with NaOH. 50 µl of pyruvate kinase (500 units/mg and 26mg/µl, 650 units) and hexokinase (1.5 mg at 363units/mg, 580.8 units). Reaction was incubated overnight at 30 °C. Following incubation, the mixture was filtered using centrifugal concentrator Viva spin 500 (Sartorius) of nominal molecular weight cutoff 10,000 to remove proteins. TLC (MeCN:H_2O, 4:1, with 10 mM tetrabutylammonium hydroxide) shows product with (R_f 0.29) and complete consumption of starting material. The desired product was purified by the HPLC (Dionex Ultimate 3000) using a strong anion exchange column Applied Biosystems, Poros® HQ (10 mm x 100 mm, 50 µm). The HPLC was eluted with gradient of 0 – 500 mM ammonium bicarbonate at a flow rate of 20 mL/min and eluants were detected with an evaporative light scattering detector (ELSD) and was obtained as a fluffy white solid and mixture of α and β anomers (5 mg, 20%).

TLC (MeCN:H_2O, 4:1, 10 mM ammonium hydroxide, v/v): R_f=0.29; $[\alpha]_D^{21}$ + -30 (c = 0.2, H_2O) [Lit. $[\alpha]_D^{24}$ + -58.1 (c = 1, H_2O)][3]; ^{1}H NMR (500 MHz, DEUTERIUM OXIDE) δ ppm 4.45-4.52 (2 H, m, H-4α, H-4β), 3.71 - 3.82 (4 H, m, H-2α, H-2β H-5α, H-5β), 3.87 (1 H, dd, $J_{2,3}$ = 9.1 Hz, $J_{1,2}$ = 3.2 Hz, H-2α), 3.91 - 4.07 (4 H, m, H-6α, H-6α'H-6β, H-6β'), 4.40 (1 H, dt, $J_{3,F}$ = 88.0 Hz, $J_{2,3}$ = 9.0 Hz, $J_{3,4}$ = 9.0 Hz H-3β), 4.40 (1 H, dt, $J_{3,F}$ = 89.1 Hz, $J_{2,3}$ 9.1 Hz, $J_{3,4}$ = 9.1 Hz H-3α), 4.36(1 H, dq, $J_{3,F}$ = 52.0 Hz, $J_{3,4}$ = 9.1 Hz, $J_{2,3}$ = 9.1 Hz, H-3β), 4.45 - 4.62 (1 H, m, H-3α), 4.64 (1 H, d, J = 7.9 Hz, H-1β), 5.19 (1 H, at, $J_{1,2}$ = 3.78 Hz, H-1α); ^{19}F NMR (377 MHz, DEUTERIUM OXIDE) δ ppm -76.4 (s);^{31}P NMR (162 MHz, DEUTERIUM OXIDE) δ ppm -23.02(s); MS m/z (ESI$^-$): 261.3 [M-H$^-$].

2-deoxy-2-fluoro-α,α-D-trehalose 6-(dihydrogenphosphate) **(71)**[i]



2-Fluoro-2-deoxy-D-glucose-6-phosphate **70** (prepared according to procedure [5]) was dissolved in 0.2 M, pH 7.4 HEPES buffer containing 10 mM $MgCl_2$ in the presence of 95 mg of UDP-glucose. To this mixture added trehalose-6-phosphate synthase (OtsA) (cloned and expressed from E. coli[6]) to a final concentration of 21.6 µM and pH was adjusted to 7.4. Total volume was 3 mL. The reaction mixture was incubated at 30 °C overnight. After confirming the reaction was completed by electrospray ionization mass spectrometer (ESI MS), the mixture was filtered using centrifugal concentrator Viva spin 500 (Sartorius) of nominal molecular weight cutoff 10,000 to remove proteins. Typically, overnight reaction was conducted to confirm the completion of the reaction. The filtrate was divided into three volumes and purified by the HPLC (Dionex Ultimate 3000) using a strong anion exchange column Applied Biosystems, Poros® HQ (10 mm x 100 mm, 50 µm). The HPLC was eluted with gradient of 0 – 500 mM ammonium bicarbonate at a flow rate of 20 mL/min and eluants were detected with an evaporative light scattering detector (ELSD). The product was eluted at retention time of 5.11 minutes. Collected fractions were pooled and concentrated under reduced pressure, yielding a white solid (25 mg, 63%).

^{1}H NMR (400 MHz, D_2O): δ 3.29 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.55 Hz, H-4'), 3.52 (1 H, dd, $J_{2,3}$ = 10.0 Hz, $J_{1,2}$ = 3.8 Hz, H-2'), 3.55 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.85 Hz, H-4) 3.61 – 3.90 (7 H, m, H-5, H-6a, H-6b, H-3', H-5', H-6a', H-6b'), 3.98 (1 H, dt, $J_{3,F}$ = 13.1 Hz, $J_{2,3}$ = $J_{3,4}$ = 9.3 Hz, H-3), 4.41 (1 H, ddd, $J_{2,F}$ = 49.0 Hz, $J_{1,2}$ = 4.2, $J_{2,3}$ = 9.5 Hz, H-2), 5.09 (1 H, d, $J_{1,2}$ = 3.8, H-1'), 5.28 (d, 1 H, $J_{1,2}$ = 4.2 Hz, H-1);

[i] Synthesis carried out by Dr. Seung Lee

^{13}C NMR (126 MHz, D_2O): δ 60.5 (C-6'), 62.2 (1 C, d, $J_{C-6,P}$ = 3.8 Hz, C-6), 68.4 (C-4), 69.5 (C-4'), 70.7 (1 C, d, $J_{C-3,F}$ = 16.1 Hz, C-3), 70.9 (C-2'), 71.9 (1 C, d, $J_{C-5,P}$ = 6.7 Hz, C-5), 72.2 (C-5'), 72.5 (C-3'), 89.6 (d, $J_{C-2,F}$ = 187.9 Hz, C-2), 91.4 (1 C, d, $J_{C-2,F}$ = 22.2 Hz, C-1), 94.0 (C-1'); ^{19}F(^{1}H) NMR (376 MHz, D_2O) δ -201.2; ^{31}P(^{1}H) NMR (162 MHz, D_2O): δ 4.5; HRMS *m/z* (ESI^-): [M-H$^-$] calcd. for $C_{12}H_{21}FO_{13}P^-$, 423.0709; found, 423.0709.

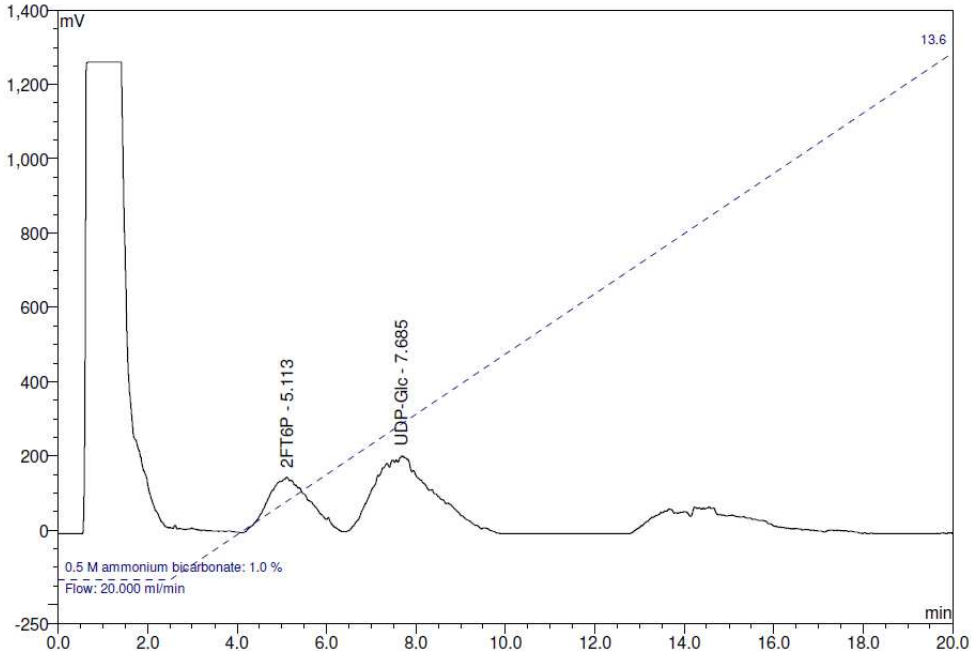
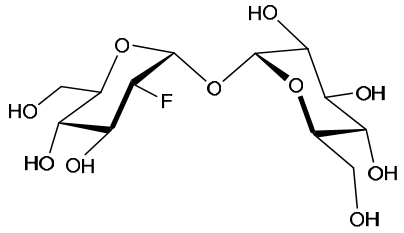


Figure 8.2. Preparative HPLC trace for the purification of 71. Conditions: HPLC (Dionex Ultimate 3000) using a strong anion exchange column Applied Biosystems, Poros® HQ (10 mm x 100 mm, 50 µm). The HPLC was eluted with gradient of 0 – 500 mM ammonium bicarbonate at a flow rate of 20 mL/min. Compound 71 elutes at 5.1 min (2FT6P)

2-deoxy-2-fluoro-α,α-D-trehalose **(52)**



71 (5 mg, 12 µmol) was dissolved in water and pH was adjusted to 8.0 with 1 M NaOH. 3 Units of alkaline phosphatase (Sigma) was added to initiate the reaction. The reaction mixture was incubated at 37 °C and the reaction progress was monitored by ESI MS. Typically, reaction was completed within 2 hours. After confirming the reaction was completed, the mixture was filtered using the centrifugal concentrator Viva spin 500 (Sartorius) of nominal molecular weight cutoff 10,000 to remove proteins.

The filtrate was loaded onto a column packed with anion exchange resin DEAE cellulose (Sigma) and the fraction eluted with water was collected and concentrated under reduced pressure. The resulting syrup was loaded onto the Phenomenex Luna NH_2 HPLC column (250 x 21.2 mm, 5 µm) on Dionex UltiMate 3000 system. Eluants were detected with an evaporative light scattering detector (ELSD). The product was eluted at 7.8 minutes of retention time by an isocratic elution with (water:acetonitrile, 30:70) at a flow rate of 18.0 mL/min. Fractions containing the product was pooled and concentrated under reduced pressure, yielding a colorless wax (4.0 mg, 99%).

^{1}H NMR (400 MHz, D_2O): δ 3.36 (1 H, at, $J_{3,4} = J_{4,5}$ = 9.0 Hz, H-4'), 3.42 (1 H, at, $J_{3,4} = J_{4,5}$ = 9.7 Hz, H-4), 3.56 (1 H, dd, $J_{1,2}$ = 3.7 Hz, $J_{2,3}$ = 9.9 Hz, H-2'), 3.58 – 3.81 (7 H, m, H-5, H-6a, H-6b, H-3', H-5', H-6a', H-6b'), 3.97 (1 H, dt, $J_{3,F}$ = 13.2 Hz, $J_{2,3} = J_{3,4}$ = 9.6 Hz, H-3), 4.36 (1 H, ddd, $J_{2,F}$ = 48.7 Hz, $J_{1,2}$ = 4.0 Hz, $J_{2,3}$ = 9.6 Hz, H-2), 5.07 (1 H, d, $J_{1,2}$ = 3.7 Hz, H-1'), 5.29 (1 H, d, $J_{1,2}$ = 4.0 Hz, H-1); ^{13}C NMR (126 MHz, D_2O): δ 60.3 (C-6 or C-6'), 60.5 (C-6 or C-6'), 69.0 (C-4), 69.1 (C-4'), 69.6 (C-3'), 70.9 (C-2'), 72.2 (C-5 or C-5'), 72.6 (C5 or C-5'), 71.1 (d, $J_{C-3,F}$ = 15.8 Hz, C-3), 89.5 (d, $J_{C-2,F}$ = 188.1 Hz, C-2), 91.2 (d, $J_{F2,C1}$ = 22.0 Hz, C-1), 94.0 (C-1'); ^{19}F(^{1}H) NMR (376 MHz, D_2O): δ -201.2; HRMS *m/z* (ESI$^-$): [M-H$^-$] calcd. for $C_{12}H_{20}FO_{10}^-$,343.1046; found, 343.1040.

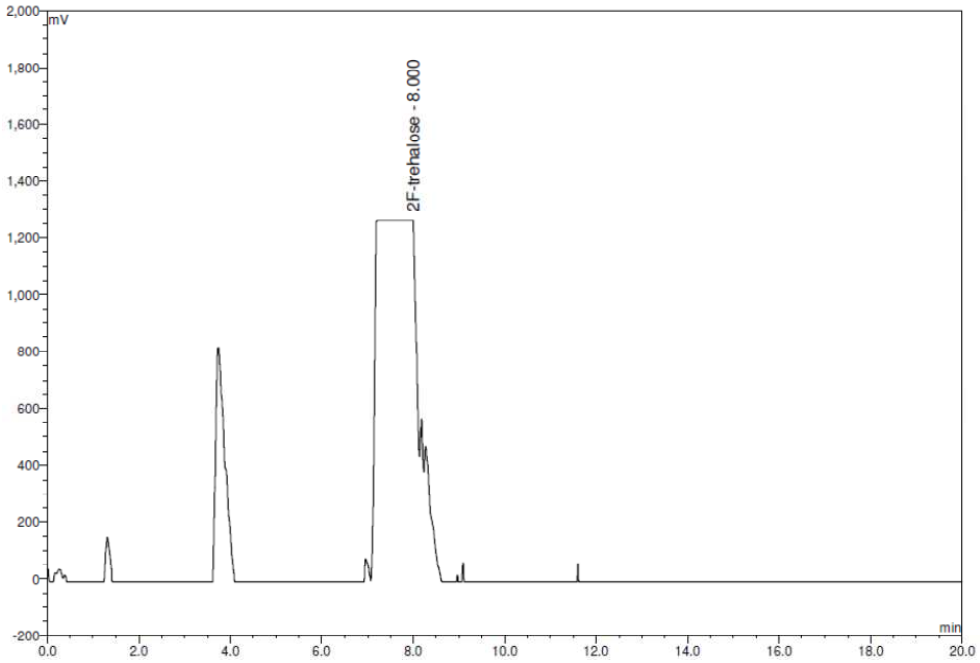
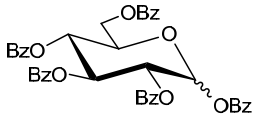


Figure 8.3. Preparative HPLC trace for the purification of 52. Conditions: the Phenomenex Luna NH2 HPLC column (250 x 21.2 mm, 5 µm) on Dionex UltiMate 3000 system. Eluants were detected with an evaporative light scattering detector (ELSD). The product 52 (2-F-trehalose) was eluted at 7.8 minutes of rentention time by an isocratic elution with 30/70 water/acetonitrile at a flow rate of 18.0 mL/min

8.2.2 Radical bromination: attempted synthesis of compounds 73, 74 and 76.

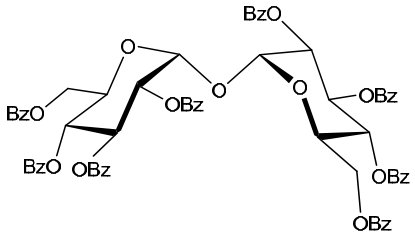
1,2,3,4,6-Penta-O-benzoyl-αß-D-glucopyranose (72)[7,8]



D-glucose **14** (3.6 g, 20.2 mmol) was added to a flame dried round bottom flask. Under argon, dry pyridine (20 mL) was added followed by benzoyl chloride (14 mL, 115.3 mmol). Reaction was stirred at room temperature for 3 hours until product had disappeared and TLC (petrol:EtOAc, 1:1) showed two close spots (R_f 0.4). Reaction was purified via flash column chromatography in (petrol:EtOAc, 3:2) ß and α anomers coeluted from the column as a 1:1 mixture of the desired compound as a pinkish, crystalline solid (11.3 g, 80%).

TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.40; $[\alpha]_D^{21}$ +108 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{20}$ +124 (c = 0.86, $CHCl_3$)][7]; M.p. = 146-148 °C (Lit. =156-158 °C)[9,10]; ^{1}H NMR (400 MHz CHLOROFORM-*d*) δ ppm 4.42 (1 H, ddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 5.2 Hz, $J_{5,6b}$= 2.7 Hz, H-5α), 4.46 - 4.57 (2 H, m, H-6α, H-5β), 4.61 - 4.72 (3 H, m, H-6α' H-6β, β'), 5.70 (1 H, dd, $J_{2,3}$ = 10.1 Hz, $J_{1,2}$ = 3.8 Hz, H-2α), 5.81 - 5.91 (3 H, m, H-4α, H-4β, β), 6.05 (1 H, at, *J* = 9.5 Hz, H-3β), 6.31 (1 H, d, $J_{1,2}$ = 8.1 Hz, H-1β), 6.34 (1 H, at, *J* = 10.0 Hz, H-3α), 6.87 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1α), 7.20 - 8.23 (50 H, m, Ar-H αβ); MS *m/z* (ESI^+): 718.3 $[M+NH_4]^+$ 723.2 $[M+Na]^+$.

***D-2,2',3,3',4,4', 6,6'-Octa-O-benzoyl-trehalose* (75)**[11]



D-Trehalose **1** (1.78 grams, 5.2 mmol, 1 equiv) was added to a flame dried round bottom flask. Under argon, dry pyridine (20 mL) was added, followed by benzoyl chloride (5 mL, 42.68 mmol). Reaction was stirred at room temperature for 3 hours until product had disappeared and TLC in (petrol:EtOAc, 1:1showed one spot (R_f 0.35). Reaction was purified via flash column chromatography in (petrol:EtOAc, 1:1) to afford the desired compound as a white solid (5.9 g, 92%).

TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.35 $[\alpha]_D^{21}$ +233 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{25}$ +224 (c = 0.86, $CHCl_3$)][11]; M.p. = 108-110 °C (Lit. = 118-120 °C)[12]; ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm

3.89 (2 H, dd, $J_{6a,6b}$ = 12.4 Hz, $J_{5,6a}$ = 4.80 Hz, H-6a, H-6a'), 4.01 (2 H, dd, $J_{6a,6b}$ = 9.7 Hz, $J_{5,6b}$ = 2.6 Hz, H-6b, H-6b'), 4.28 - 4.36 (2 H, m, H-5, H-5'), 5.51 (2 H, dd, $J_{2,3}$ = 10.2 Hz, $J_{1,2}$ = 4.0 Hz, H-2, H-2'), 5.67 (2 H, at, J = 9.9 Hz, H-4, H-4'), 5.75 (2 H, d, $J_{1,2}$ = 4.0 Hz, H-1, H-1'), 6.31 (2 H, at, J = 9.9 Hz, H-3, H-3'), 7.23 - 7.9 (20 H, m, Ar-H), 7.51 - 7.59 (2 H, m, Ar-H), 7.52 - 7.59 (2 H, m, Ar-H), 7.70 (4 H, t, J = 8.2, Ar-H), 7.89 - 7.97 (8 H, m, Ar-H), 8.07 (4 H, d, J = 7 Hz, Ar-H); MS m/z (ESI^+): 1254.4 $[M+Pyridinium]^+$.

Procedure for Radical Bromination at C-5

72 (740 mg, 1.06 mmol) was dissolved in CCl_4. To this was added NBS (392 mg, 2.17 mmol) and benzoyl peroxide (26.6 mg, 0.109 mmol) The reaction was exposed to UV light 300 watts for 3 hours at 60 ˚C, under argon and was monitored by TLC in (petrol:ethyl acetate, 2:1). TLC revealed the appearance of a new spot (R_f 0.22) and the starting material dissappearance at (R_f 0.27). However, by NMR the reaction is still primarily starting material, with about 1/5 conversion to multiple products.

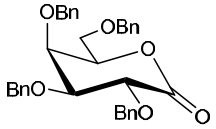
Table 8.1. Radical bromination of C-5 reaction attempts

Entry	Sugar	Bromine source	Benzoyl Peroxide	T ˚C	Time	Result
1	**72**	NBS	no	60 ˚C	6 h	NR
2	**130** (acetylglucopyranose)	NBS	no	60 ˚C	3 h	Product by ESI
3	**72**	NBS	no	80 ˚C	4 h	Hydrolysis product detected by ESI
4	**72**	NBS	yes	100 ˚C	5 h	Product by ESI
5	**75**	NBS	yes	60 ˚C	3 h	NR
6	**75**	Br_2	no	80 ˚C	3 h	Multiple inseparable products

8.2.3 Ketoside trehalose synthesis

8.2.3.1 Synthesis of methyl ketosides 96, 97 and 103

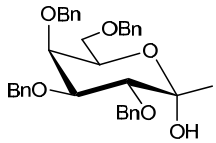
2,3,4,6-Tetra-O-benzyl-galactono-1,5-lactone (92)[13]



Acetic anhydride (9 mL) was added to a solution of **91** (2.17 g, 4.01 mmol) in dry DMSO (14 mL). After 16 hours reaction was complete as judged by TLC (petrol:EtOAc, 3:1), which indicated the formation of a product (R_f 0.5) and complete disappearance of starting material (R_f 0.3). Reaction mixture was poured onto brine (100 mL of saturated solution) and extracted into toluene (2x 100 mL). The organic layers were washed with $NaHCO_3$ (saturated aqueous solution) until pH neutral and then washed with brine (2x20 mL). After workup, the desired compound was obtained (2.43 g, 99%) as a clear oil.

TLC (petrol:EtOAc, 3:1, v/v): R_f = 0.5; $[\alpha]_D^{24}$ + 80.5 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{21}$ + 75.2 (c = 1, $CHCl_3$)]^{13}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 3.75 (1 H, dd, $J_{6a,6b}$ = 9.2 Hz, $J_{5,6a}$ = 4.9, H-6a), 3.76 (1 H, dd, $J_{6a,6b}$ = 8.0 Hz, $J_{5,6b}$ = 1.1, H-6b), 3.96 (1 H, dd, $J_{3,4}$ = 9.7 Hz, $J_{3,4}$ = 2.1 Hz, H-3) 4.24 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 1.8 Hz, H-4) 4.39 (1 H, ddd, $J_{5,6a}$ 7.7 Hz, $J_{5,6b}$ = 5.9 Hz, $J_{4,5}$ = 1.3 Hz, H-5) 4.49 - 4.60 (3 H, m H-2, 2 x OCH_2Ph) 4.69 (1 H, d, *J* = 11.4 Hz, 1 x OCH_2Ph) 4.75 (1 H, d, *J* = 11.9, 1 x 1 x OCH_2Ph) 4.81 (1 H, d, *J* = 11.9Hz, 1 x OCH_2Ph) 4.86 (2 H, d, *J* = 10.9 Hz, 1 x OCH_2Ph), 5.01 (2 H, d, *J* = 11.1 Hz, *J* = 1 x OCH_2Ph) 5.26 (2 H, d, *J* = 10.9 Hz, 1 x OCH_2Ph) 7.17 - 7.62 (m, 20 H, Ar-H); MS *m/z* (ESI^+): 556.3 $[M+NH_4]^+$.

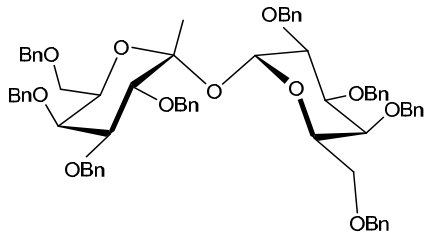
3,4,5,7-Tetra-O-benzyl-1-deoxy-D-galacto-hept-2-ulopyranose (92)[14]



To **92** (63.15 mg, 0.11 mmol, 1 equiv) in anhydrous THF (40 mL) was added methyllithium (1.2 mL, 1.38M in Et_2O, 15 equiv) at -78 °C. After 1.5 hours reaction was determined to be complete by TLC (petrol:EtOAc, 2:1) showed exclusively product (R_f 0.3) and dissappearance of starting lactone (R_f 0.55). Reaction was quenched with aqueous NH_4Cl and washed with saturated brine and $NaHCO_3$ and purification by column chromatography (petrol:EtOAc, 3:1) afforded the desired compound (54.27 mg, 85%) as a clear oil.

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.3; $[\alpha]_D^{24}$ +20.6 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{25}$ +22.1 (c = 1, $CHCl_3$)]***[14]***; ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.43 (3 H, s, CH_3), 3.57 (1 H, dd, $J_{7a,7b}$ = 6.6 Hz, $J_{6,7a}$ = 1.1 Hz, $J_{6,7b}$ = 1.1 Hz, H-7b), 3.58 (1 H, dd, $J_{7a,7b}$ = 6.6 Hz, $J_{6,7a}$ = 1.1 Hz, $J_{6,7b}$ = 1.1 Hz H-7a), 3.85 (1 H, dd $J_{3,4}$ = 9.6 Hz, $J_{4,5}$ = 1.1 Hz H-4), 3.93 (1 H, d, $J_{3,4}$ = 8.1 Hz, H-3), 4.03 (1 H, br. s., H-4), 4.08 - 4.15 (1 H, ddd, $J_{6,7a}$ = 5.6 Hz, $J_{6,7b}$ 4.1 Hz, $J_{5,6}$ = 1.2 Hz, H-6), 4.45 (1 H, d, *J* = 11.9 Hz, 1 x OCH_2Ph), 4.51 (1 H, d, *J* = 12.6 Hz, 1 x OCH_2Ph), 4.65 (1 H, d, *J* = 11.6 Hz, 1 x OCH_2Ph), 4.67 (1H, d, J = 10.1 Hz, 1 x OCH_2Ph), 4.73 - 4.81 (2 H, m, 2 x OCH_2Ph), 4.96 (1 H, d, *J* = 10.6 Hz 1 x OCH_2Ph), 5.01 (1 H, d, *J* = 11.4 Hz, 1 x OCH_2Ph), 7.13 - 7.51 (m, 20 H, Ar-H); MS *m*/*z* (ESI^+): 572.3 $[M+NH_4]^+$.

2,3,4,6-tetra-O-benzyl-1-deoxy-α-D-galacto-hept-2-ulopyranosyl-(2→1)-2,3,4,6-tetra-O-benzyl-α-D-galactopyranoside (96) [15]

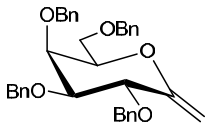


According to the procedure in Yamanoi, 2001[16], a stirred solution of $Bi(OTf)_3$ (5.6 mg, 0.0067 mmol) and **91** (48.3 mg, 0.0893 mmol) in dry CH_2Cl_2 (3.5 mL) was added to a dry flask in the presence of molecular sieves (ca. 100 mg). To this was added **93** (79.3 mg, 0.134 mmol) at 0 °C in the under an Ar atmosphere. The resulting mixture was stirred for 3 h. TLC in (petrol:EtOAc, 2:1) revealed the production of two new products (R_f 0.85 and 0.8.) The reaction was then quenched by the addition of a satd. $NaHCO_3$ solution (5 mL). The reaction mixture was extracted with ethyl acetate, and the organic layer was washed with water and a satd. brine solution. After the organic layer was dried over Na_2SO_4, the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (petrol:EtOAc, 4:1) to give as a colorless oil as a mixture of α and β products (47.1 mg, 46%) of about 70% α,α by nmr. A second column in (petrol:EtOAc, 6:1) produced exclusively the desired α,α product (16.4 mg, 16.3%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.85; $[\alpha]_D^{24}$ + 56.8 (c = 0.5, $CHCl_3$) [Lit. $[\alpha]_D^{20}$ +70.8 (c = 0.99, $CHCl_3$)] [15]; ^{1}H NMR (500 MHz, CHLOROFORM-*d*) δ ppm 1.48 (3H, s, CH_3), 3.32 - 3.46 (3 H, m, H-6_b, H-7_a', H-7_b'), 3.53-3.55 (1 H, m, H-6_a), 3.80 (1 H, d, $J_{3,4}$ = 9.8 Hz, H-3'),

3.88 - 3.90 (1 H, m, H-5'), 3.93 - 4.03 (3 H, m, H-3, H-4', H-4), 4.04 (1 H, dd, $J_{2,3}$ = 9.8, $J_{1,2}$ = 3.4 Hz, H-2), 4.29 (1 H, d, J = 11.7 Hz, 1 x OCH_2Ph), 4.33 (1 H, at, $J_{5,6a}$= $J_{5,6b}$ = 6.9 Hz, H-5), 4.38 (2 H, d, J = 12.9 Hz, 2 x OCH_2Ph), 4.42 - 4.48 (1 H, m, H-6'), 4.44 (1H d, J = 11.8 Hz, 1 x OCH_2Ph) 4.55 (1 H, d, J = 11.7 Hz, 1 x OCH_2Ph), 4.58 (1 H, d, J = 11.7 Hz, 1 x OCH_2Ph), 4.63 - 4.69 (2 H, m, 2 x OCH_2Ph), 4.69 - 4.77 (5 H, m, 5 x OCH_2Ph), 4.90 (1H, d, J = 11.6 Hz, 1 x OCH_2Ph), 4.91 (1H, d, J = 11.8 Hz, 1 x OCH_2Ph), 4.91 (1H, d, J = 11.3 Hz, 1 x OCH_2Ph), 5.35 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1), 7.36–7.16 (40H, m, Ar-*H*); MS *m*/*z* (ESI^+): 1094.5 [$M+NH_4$].

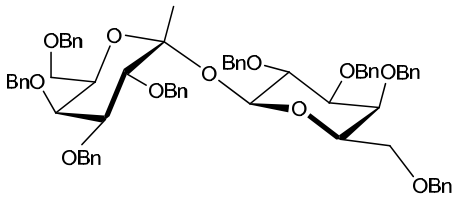
2,6-Anhydro-1-deoxy-3,4,5,7-tetra-O-benzyl-D-galacto-hept-1-enitol (102)[17]



According to the procedure in Ranjanbabu, 1986[18] **92** (988 mg, 1.829 mmol) in a mixture of anhydrous toluene (6 mL), tetrahydrofuran (2 mL), and pyridine (25 μL) was stirred with a solution of Tebbe's reagent (2 mL, 1 mmol) in the presence of molecular sieves (ca. 100 mg) at -40 °C for 1 h. The solution was then allowed to reach O °C and stirred for 30 min. The progress of the reaction was monitored by TLC (petrol:EtOAc, 3:1) and upon completion (R_f 0.8), the reaction was quenched with aqueous 15% sodium hydroxide, added dropwise (0.1 mL) and then diluted with ether (100 mL). The mixture was stirred for 30min, filtered through Celite® to remove the inorganic precipitate and then purified via column chromatography (petrol:EtOAc, 6:1) with triethylamine (1%). Evaporation of the solvents yielded the desired compound (295 mg, 30%) as a slightly yellow oil.

TLC (petrol:EtOAc, 3:1, v/v): R_f = 0.8; $[\alpha]_D^{24}$ + 48.1 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{20}$ + 54.2 (c = 0.88, $CHCl_3$)[17]; ^{1}H NMR (400 MHz, CHLOROFORM-d) δ ppm 3.62 (1 H, dd, $J_{6,7a}$ = 9.6 Hz, $J_{7a,7b}$ = 6.5 Hz, H-7_a), 3.66 (1 H, dd, $J_{3,4}$ = 9.0 $J_{4,5}$ = 2.7 Hz, H-4), 3.70 (1 H, dd, $J_{6,7b}$ = 9.6, $J_{7a,7b}$ = 6.0 Hz, H-7_b), 3.82 (1 H, ddd, $J_{5,6}$ = 6.5 Hz, $J_{6,7a}$ = 6.0 Hz, $J_{6,7b}$ = 2.0 Hz, H-6), 4.05 (1 H, at, J = 2.5 Hz, H-5), 4.36 (1 H, d, $J_{3,4}$ = 9.0 Hz, H-3), 4.45 (1 H, d, J = 11.9 Hz, 1x OCH_2Ph), 4.52 (1 H, d, J = 11.9 Hz, 1 x OCH_2Ph), 4.60 (, J = 11.5 Hz, 1x OCH_2Ph), 4.70 (1 H, d, J = 1.4 Hz, H-1_a), 4.71 (1 H, d, J = 11.8 Hz, 1x OCH_2Ph), 4.72 (1 H, d, J = 11.8 Hz, 1x OCH_2Ph), 4.73 (1 H, d, J = 11.9 Hz, 1x OCH_2Ph), 4.72 (d, 1 H, $J_{1a,1b}$ = 4.1 Hz, H-1_a), 4.74 (1 H, d, $J_{1a,1b}$ = 4.4 Hz, H-1_b), 4.77 (1 H, d, J = 11.9 Hz, 1x OCH_2Ph) 4.92 (1 H, d, J = 11.5 Hz, 1x OCH_2Ph), 7.25-7.37 (m, 20H, AR-H); MS *m/z* (ESI^+): 554.3 [$M+NH_4$]$^+$.

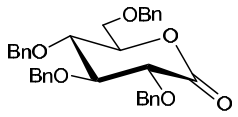
2,3,4,6-tetra-O-benzyl-1-deoxy-α-D-galacto-hept-2-ulopyranosyl-(2→1)-2,3,4,6-tetra-O-benzyl-β-D-galactopyranoside (97)[15]



Following the procedure in Li, 2001[19] a solution of **102** (55 mg, 0.102 mmol), **91** (51.6 mg, 0.955 mmol) and 4 Å molecular sieves (100 mg) in dry CH_2Cl_2 (5 mL) was stirred under nitrogen atmosphere at room temperature for 30 min. The solution was cooled to −78 °C, and 5 μL (0.05 equiv.) of TMSOTf was added. The reaction mixture was stirred at −78 °C for 30 minutes. The reaction was monitored by TLC (petrol:EtOAc, 6:1) and upon appearance of product (R_f 0.3) and disappearance of starting material (R_f 0.8) was quenched with one drop of triethylamine (Et_3N) and passed through celite. After the removal of the solvent, the residue was applied on a silica gel column chromatography (petrol:EtOAc, 7:1) as the eluent to afford the products **96** and **97** as clear oils (97.2 mg, 95%). The α,α and α,β products were fully separated by chromatography, producing a 3:7 ratio of α,α and α,β.

TLC (petrol:EtOAc, 6:1, v/v): R_f = 0.3; $[\alpha]_D^{24}$ + 33.4 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{25}$ + 38.2 (c = 1.07, $CHCl_3$)] [15]; ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.49 (3 H, s, CH_3), 3.41 (1H dd, $J_{6,7a}$ = 8.8 Hz, $J_{7a,\ 7b}$ = 5.5, Hz H-7$_a$'), 3.46 (1 H, dd, $J_{5,6}$ = 8.3 Hz, $J_{6a,6b}$ = 4.7 Hz H-6b), 3.50 (1 H, dd, *J* = 9.6 Hz, = 3.0 Hz, H-3), 3.51 (1 H, m, H-5), 3.56 (1 H, dd, $J_{6,7b}$ = 8.8 Hz, $J_{7a,7b}$ = 8.8 Hz, H-7$_b$'), 3.59 (1 H, dd, $J_{5,6a}$ = 8.3 Hz, $J_{6a,6b}$ = 8.3 Hz, H-6a), 3.81 (1 H, dd, $J_{2,3}$ = 9.8 Hz, $J_{1,2}$ = 7.9 Hz, H-2), 3.87 (1 H, d, $J_{3,4}$ = 10.1 Hz, H-3'), 3.92 (1 H, d, *J* = 2.5 Hz, H-4), 4.04 (dd, $J_{4,5}$ = 2.5 Hz, $J_{5,6}$ = 1.1 Hz, 1H, H-5'), 4.17 (1 H, dd, $J_{5,6}$ = 7.6 Hz, $J_{6,7a}$ = 2.8 Hz, H-4'), 4.27 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph), 4.29 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph), 4.32 (1 H, d, *J* = 11.8 Hz, 1 x OCH_2Ph), 4.38 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.51 (1 H, ddd, $J_{5,6}$ = 8.8 Hz, $J_{6,\ 7a}$ = 5.5 Hz, $J_{6,\ 7b}$ = 1.1 Hz, H-6'), 4.59 (1 H, d, *J* = 11.6 Hz, 1 x OCH_2Ph), 4.59 (1 H, d, *J* = 11.6 Hz, 1 x OCH_2Ph), 4.66 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph), 4.67 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.68 (1 H, d, $J_{1,2}$ = 7.7 Hz, H-1), 4.71 (1 H, d, *J* = 11.8 Hz, 1 x OCH_2Ph), 4.73 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph), 4.74 (1 H, d, *J* = 11.8 Hz, 1 x OCH_2Ph), 4.91 (1 H, d, *J* =11.6 Hz, 1 x OCH_2Ph), 4.92 (1 H, d, *J* = 11.6 Hz, 1 x OCH_2Ph), 4.96 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.99 (1 H, d, *J* = 11.6 Hz, 1 x OCH_2Ph), 7.35–7.12 (m, 40H, Ar-*H*). MS *m*/*z* (ESI^+): 1094.6 $[M+NH_4]^+$.

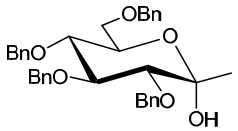
2,3,4,6-Tetra-O-benzyl-glucono-1,5-lactone (95)[13,20]



Acetic anhydride (0.7 mL) was added to a solution of **94** (190.1 mg, 0.31 mmol) in dry DMSO (1.2 mL). After 18 hours reaction was complete by TLC (petrol:EtOAc, 3:1) indicated the formation of a product (R_f 0.5) and complete disappearance of starting material (R_f 0.3). Reaction mixture was poured onto brine (5 mL of saturated solution) and extracted into toluene (2 x 10 mL). The organic layers were washed with satd $NaHCO_3$ until pH neutral and then washed with brine (2 x 10 mL). After workup, purification by column chromatography (petrol:EtOAc, 3:1) afforded the desired product (113.1 mg, 60%) as a clear oil.

TLC (petrol:EtOAc, 3:1, v/v): R_f = 0.5; $[\alpha]_D^{24}$ + 62.5 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{21}$ +54.2 (c = 1, $CHCl_3$)][13,20]; ^{1}H NMR (400 MHz, CHLOROFORM-d) δ ppm 3.72 (1 H, dd, $J_{6a,6b}$ = 10.8 Hz, $J_{5,6a}$ = 3.3, H-6a) 3.75 (1 H, dd, $J_{6a,6b}$ = 10.8 Hz, $J_{5,6b}$ = 2.3, H-6b), 3.91 - 4.03 (2 H, m, H-4, H-3), 4.16 (1 H, d, $J_{2,3}$ = 6.6 Hz, H-2), 4.48 (1 H, ddd, $J_{4,5}$ = 8.8 Hz, $J_{5,6b}$ = 3.3 Hz, $J_{5,6a}$ = 2.2 Hz, H-5) 4.50 (1 H, d, J = 12.1 Hz 1 x OCH_2Ph), 4.67 (1 H, d, J = 11.4 Hz, 1 x OCH_2Ph), 4.54 (1 H, d, J = 11.4 Hz, 1 x OCH_2Ph), 4.57 (1 H, d, J = 9.1 Hz, 1 x OCH_2Ph), 4.63 (1 H, d, J = 11.9 Hz, 1 x OCH_2Ph), 4.73 (1 H, d, J = 11.2 Hz, 1 x OCH_2Ph) 4.77 (2 H, d, J = 11.2 Hz, 2 x OCH_2Ph), 5.03 (1 H, d, J = 11.4 Hz, 1 x OCH_2P), 7.21-7.41 (20 H, m, ArH); MS *m/z* (ESI^+): 556.3 $[M+NH_4]^+$.

3,4,5,7-Tetra-O-benzyl-1-deoxy-D-gluco-hept-2-ulopyranose (85)[21]

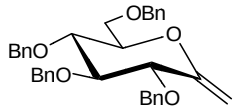


To **95** (1.068 g, 1.99 mmol, 1 equiv) in dry THF (10 mL) was added MeLi (1.5 mL, 1.38M in Et_2O, 1.2 equi) at -78 °C. After 1 hours reaction proceeded to completion with conversion to product (R_f 0.5) and disappearance of starting material (R_f 0.7) as monitored by TLC (petrol:EtOAc, 2:1)and was quenched with aqueous NH_4Cl and washed with saturated brine and $NaHCO_3$. Product was purified by column chromatography (petrol:EtOAc, 2:1) afforded the desired product (633 mg , 60 %) as a white crystalline solid as well as recovery of (115 mg, 10.7%) of the starting lactone.

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.5; $[\alpha]_D^{24}$ + 25.6 (C = 1, $CHCl_3$) [Lit. $[\alpha]_D^{22}$ +13 (c = 0.4, $CHCl_3$)][21]; ^{1}H NMR (400 MHz, CHLOROFORM-d) δ ppm 1.42 (3H, s, CH_3), 3.38 (1 H, d, $J_{3,4}$ = 9.3 Hz, H-3), 3.66 (1 H, at, J = 9.4 Hz, H-5) 3.69 (1 H, dd, $J_{7a,7b}$ = 10.1 Hz, $J_{6,7b}$ = 1.8 Hz, H-7b) 3.74

(1H, dd, $J_{7a,7b}$ = 10.8 Hz, $J_{6,7a}$ = 4.0 Hz, H-7a) 3.99 (1 H, at, *J* = 9.3Hz, H-4) 4.03 (1 H, ddd, $J_{5,6}$ = 8.4 Hz, $J_{6,7a}$ = 4.4 Hz, $J_{6,7b}$ = 2.2 Hz, H-6) 4.52 - 4.59 (2 H, m, 2 x OCH_2Ph) 4.61 - 4.66 (1 H, d, *J* = 10.8 Hz, 1x OCH_2Ph) 4.71 (1 H, d, *J* = 11.1 Hz, 1 x OCH_2Ph) 4.83 (1 H, d, *J* = 10.6 Hz, 1 x OCH_2Ph) 4.90 (2 H, d, *J* = 2.0 Hz, 1 x OCH_2Ph) 4.94 (1 H, d, *J* = 11.1 Hz, 1 x OCH_2Ph) 7.11 - 7.48 (20 H, m, Ar-H); MS *m/z* (ESI^+): 572.3 $[M+NH_4]^+$.

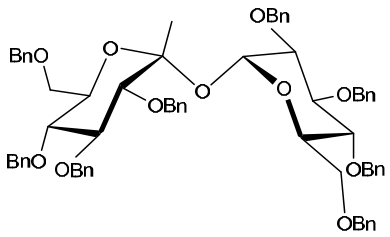
2,6-Anhydro-1-deoxy-3,4,5,7-tetra-O-benzyl-D-gluco-hept-1-enitol (99)[22]



According to the procedure in Ranjanbabu, 1986[18] **95** (308.8 mg, 0.573 mmol, 1 equiv) in a mixture of anhydrous toluene (5 mL), tetrahydrofuran (2 mL), and pyridine (2 µL) was stirred with a solution of Tebbe's reagent (2.5 mL, 1.25mmol, 2.2 equiv) in the presence of molecular sieves (ca. 100 mg) at -40 ˚C for 1 h. The solution was then allowed to reach 0 ˚C and stirred for 30 min. The progress of the reaction was monitored by TLC (petrol:EtOAc, 3:1) and upon completion (R_f 0.8), the reaction was quenched with aqueous 15% sodium hydroxide, added dropwise (0.1 mL) and then diluted with ether (100 mL). The mixture was stirred for 30min, filtered through Celite® to remove the inorganic precipitate and then purified via column chromatography (petrol:EtOAc, 3:1) with triethylamine (1%). Evaporation of the solvents yielded the desired compound (40 mg, 13%) as a slightly yellow oil.

TLC (petrol:EtOAc, 3:1, v/v): R_f = 0.8; $[\alpha]_D^{24}$ + 48.2 (c = 0.7, $CHCl_3$) [Lit. $[\alpha]_D^{20}$ +62.1 (c = 1, $CHCl_3$)] [22]; ^{1}H NMR (400 MHz, CHLOROFORM-d) δ 3.69-3.78 (5 H, m, H-4, H-5, H-6, H-7_a, H-7_b), 3.96 (1 H, d, $J_{2,3}$ = 7.2 Hz, H-3), 4.51 (1 H, d, *J* = 11.1 Hz, OCH_2Ph), 4.53 (1 H, d, *J* = 12.1 Hz, OCH_2Ph), 4.62 (1 H, d, $J_{1a,1b}$ = 6.2 Hz, H-1_a), 4.63 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph), 4.64 (1 H, d, *J* = 11.6 Hz, 1 x OCH_2Ph), 4.71 (1 H, d, *J* = 11.2 Hz, 1 x OCH_2Ph), 4.76 (1 H, d, $J_{1a,1b}$ = 2.8 Hz, H-1_b), 4.77 (1 H, d, *J* = 11.6 Hz, 1 x OCH_2Ph), 4.77 (1H, d, *J* = 11.2 Hz, 1 x OCH_2Ph), 4.86 (1 H, d, *J* = 11.1 Hz, 1 x OCH_2Ph), 7.12-7.15 (2 H, m, Ar-H), 7.24-7.37 (18 H, m, Ar-H); MS *m/z* (ESI^+): 554.3 $[M+NH_4]^+$.

2,3,4,6-Tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2,3,4,6-tetra-O-benzyl-α-D-glucopyranoside **(103)** [15]



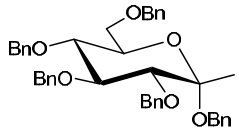
Following the procedure in Li, 2001[19] A solution of **99** (37.0 mg, 0.068 mmol), **94** (40.0 mg, 0.074 mmol) and molecular sieves 4 Å (100 mg) in dry CH_2Cl_2 (5 mL) was stirred under nitrogen atmosphere at room temperature for 30 min. The solution was cooled to −78 °C, and 5 μL (0.05 equiv.) of TMSOTf was added. The reaction mixture was stirred at −78 °C for 30 minutes. The reaction was monitored by TLC (petrol:EtOAc, 6:1) and upon completion (R_f 0.3) quenched with one drop of triethylamine (Et_3N) and passed through celite. After the removal of the solvent, the residue was applied on a silica gel column chromatography (petrol:EtOAc, 7:1) as the eluent to afford the desired products as a clear oil (50 mg, 68%). The α,α and α,β products were fully separated by chromatography, producing a 2:3 ratio of α,α and α,β.

*Alternate synthesis of **103:*** **85** (94 mg, 0.17 mmol) and **94** (87 mg, 0.16 mmol) were dried en vacuo for 1 hour then dissolved in anhydrous CH_2Cl_2 (8 mL) and added to a dry flask in the presence of molecular sieves (ca. 100 mg) and stirred with the sieves for 30 minutes. To this was added TMSOTf (ca. 5 μl, 0.027mmol) at -40 °C in the under an Ar atmosphere. The resulting mixture was stirred for 5 h. TLC (petrol:EtOAc, 4:1) revealed the production of one new product spot (R_f 0.5) that was UV active and complete consumption of starting material. The reaction was then quenched by the addition 0.01 mL NEt_3, filtered through celite and concentrated. The crude product was purified by column chromatography (petrol:EtOAc, 5:1) to yield the α product (47 mg, 25%).

TLC (petrol:EtOAc, 6:1, v/v): R_f = 0.3; $[\alpha]_D^{24}$ + 61.8 (c = 0.77, $CHCl_3$). [Lit $[\alpha]_D^{20}$ +75.0 (c = 0.97, $CHCl_3$)] [15]; ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ 1.49 (3H, s, CH_3), 3.29 (1 H, d, $J_{3,4}$ = 9.6 Hz, H-3'), 3.33–3.39 (3 H, m, H-6a, H-6b, H-7′a), 3.55–3.58 (1 H, m, H-7′b), 3.56 (1 H, dd, $J_{2,3}$ = 10.3 Hz, $J_{1,2}$ = 3.4 Hz, H-2), 3.64 (1 H, dd, $J_{4,5}$ = 10.3 Hz, $J_{3,4}$ = 9.6 Hz, H-4), 3.68 (1 H, at, *J* = 9.6 Hz, H-5′), 4.03 (1 H, t, *J* 10.3 Hz, H-3), 4.05 (1 H, at, $J_{3,4} = J_{4,5}$ = 9.6 Hz, H-4′), 4.17–4.19 (1 H, ddd $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 3.2 Hz, $J_{5,6b}$ = 2.3 Hz, H-5), 4.27 (1 H, ddd, $J_{5,6}$ = 10.2 Hz, $J_{6,7a}$ = 3.6 Hz, $J_{6,7b}$ = 2.9 Hz, H-6'), 4.36 (1 H, d, *J* = 12.4 Hz, 1 x OCH_2Ph), 4.37 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph), 4.47 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.48 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.51 (1 H, d, *J* = 12.4 Hz, 1 x OCH_2Ph), 4.57 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph), 4.59 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.64 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph), 4.69 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph),

4.80 (1 H, d, J = 11.0 Hz, 1 x OCH_2Ph), 4.80 (1 H, d, J = 11.0 Hz, 1 x OCH_2Ph), 4.84 (1 H, d, J = 10.7 Hz, 1 x OCH_2Ph), 4.87 (2 H, m, 2 × OCH_2Ph), 4.92 (1 H, d, J = 11.3 Hz, 1 x OCH_2Ph), 4.94 (1 H, d, J = 10.7 Hz, 1 x OCH_2Ph), 5.34 (1 H, d, $J_{1,2}$ = 3.4 Hz, H-1′); MS m/z (ESI^+): 1094.6 $[M+NH_4]$.

2,3,4,5,7-Penta-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranose (108) [19]

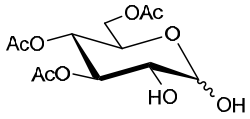


108 was isolated as a side project of the reaction between **85** and **94** as a clear oil in (25 mg, 29%)

TLC (petrol:EtOAc, 4:1, v/v): R_f = 0.8; $[\alpha]_D^{25}$ + 16.2 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{20}$ + 34.2 (c = 1, $CHCl_3$)[19]; ^1H NMR ($CDCl_3$): δ 1.39 (s, 3 H, CH_3), 3.39 (1 H, d, $J_{2,3}$ = 9.36 Hz, H-3), 3.61–3.76 (4 H, m H-7a, H-7b H-6 and H-5), 4.13 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.31 Hz, H-4), 4.50–4.62 (5 H, m, 5 x OCH_2Ph), 4.72 (1 H, d, J = 11.60 Hz, 1 x OCH_2Ph), 4.81–4.97 (4 H, m, 4 x OCH_2Ph), 7.13–7.16 (2 H, m, ArH), 7.22–7.37 (23 H, m, ArH); MS m/z (ESI^+): 662.3 $[M+NH_4]$.

8.2.4 Intramolecular aglycone delivery glycosylation

3,4,6-Tri-O-acetyl-α,β-D-glucopyranose (246) [23]

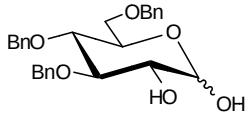


C*is*-Dihydroxylation of glycal or 2,3-unsaturated glycoside**.** [24] A mixture of $NaIO_4$ (296.4 mg, 1.4 mmol) and $CeCl_3 \cdot 7H_2O$ (37 mg, 0.1 mmol) in H_2O (2 mL) was stirred at room temperature for a few minutes. The reaction mixture was cooled to 0 °C, and to the cooled reaction mixture were added ethyl acetate (3 mL), CH_3CN (6 mL), and $RuCl_3 \cdot H_2O$ (5.2 mg, 0.025 mmol) successively. After the mixture stirred for 2.0 min, a solution of 3,4,6-tri-O-acetyl-glucal (270.02 mg, 1.0 mmol) in ethyl acetate (3 mL) was added, and the resulting heterogeneous mixture was stirred until the full consumption of the starting. After 15 minutes TLC (EtOAc:petrol, 1.5:1) indicated complete conversion to product (R_f 0.1) and disappearance of staring material (R_f 0.8). The reaction mixture was diluted with ethyl acetate (25 mL). The organic layer was washed with aqueous $NaHCO_3$ and water, dried (Na_2SO_4), passed through celite and concentrated under reduced pressure. The crude

product was purified via column chromatography (petrol:EtOAc, 3:1) yielding the desired compound (133.0 mg, 43.7%).

TLC (EtOAc:petrol, 3:2, v/v): R_f = 0.1; $[\alpha]_D^{24}$ + 93.0 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{20}$ +89.7 (c = 2.75, $CHCl_3$)][23]; ^{1}H NMR (400 MHz CHLOROFORM-*d* 3:1 mixture of anomers) δ ppm 2.03 - 2.06 (16 H, m), 2.08 - 2.11 (25 H, m), 3.55 (1 H, dd, $J_{2,3}$ = 9.2 Hz, $J_{1,2}$ = 7.9 Hz, H-2β), 3.66 - 3.74 (4 H, m, H-2α), 3.75 (1 H, ddd, $J_{4,5,}$ = 8.1, $J_{5,6}$ = 4.9, $J_{5,6}$ = 2.4 Hz, H-5β), 4.08 - 4.16 (7 H, m, H-6α, H-6β), 4.19 - 4.28 (8 H, m, H-6'α H-6'β, H-5α), 4.72 (1 H, d, $J_{1,2}$ = 7.7 Hz, H-1β), 5.03 (4 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.8 Hz, H-4α), 5.03 (3 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4β), 5.11 (1 H, at, $J_{2,3}$ = $J_{3,4}$= 9.4 Hz, H-3β), 5.28 (4 H, at, $J_{2,3}$ = $J_{3,4}$= 9.4 Hz, H-3α), 5.33 (1 H, d, $J_{1,2}$ = 3.9 Hz, H-1α); MS *m/z* (ESI$^+$): 324.1 $[M+NH_4]^+$.

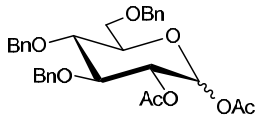
***3,4,6-tri-O-benzyl-α,β-D-glucopyranose (247)*[25]**



A mixture of $NaIO_4$ (907 mg, 4.2 mmol) and $CeCl_3 \cdot 7H_2O$ (114.7 mg, 0.3 mmol) in H_2O (4 mL) was stirred at room temperature for a few minutes. The reaction mixture was cooled to 0 °C, and to the cooled reaction mixture were added ethyl acetate (6 mL), CH_3CN (12 mL), and $RuCl_3 \cdot H_2O$ (18.4 mg, 0.075 mmol) successively. After the mixture stirred for 2.0 min, a solution of **109** (1.22 g, 2.92 mmol) in ethyl acetate (6 mL) was added, and the resulting heterogeneous mixture was stirred until the full consumption of the starting material (Table 1). After 15 minutes TLC (petrol:EtOAc, 1:1) indicated complete conversion to product (R_f 0.3) and disappearance of staring material (R_f 0.9). The reaction mixture was diluted with ethyl acetate (25 mL). The organic layer was washed with aqueous $NaHCO_3$ and water, dried (Na_2SO_4), passed through celite and concentrated under reduced pressure. The crude product was purified via column chromatography (petrol:EtOAc, 1:1). The desired compound was obtained as a (1.3:1) anomeric mixture (1.07 g, 81%).

TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.3; $[\alpha]_D^{24}$ + 48.1 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{23.5}$ +57.1 (c = 0.9, $CHCl_3$)][25]; ^{1}H NMR (400 MHz CHLOROFORM-*d*) δ ppm 3.48 (2 H, at, $J_{1,2}$ = $J_{2,3}$= 8.6 Hz, H-2β), 3.55 - 3.61 (3 H, m, H-6αβ, H-2α), 3.64 - 3.74 (4 H, m, H-4α, H-6αβ, H-5β, H-3β), 3.81 (1 H, at, $J_{2,3}$ = $J_{3,4}$= 9.0 Hz, H-3α), 4.06 (2 H, ddd, $J_{4,5}$ = 9.8 Hz, $J_{5,6a}$ = 4.4 Hz, $J_{5,6b}$ = 2.6 Hz, H-5α), 4.50 - 4.55 (3 H, m, H-1β, 2 x OCH_2Ph), 4.60 (1 H, d, *J* = 12.1 Hz, H-3β), 4.61 (1 H, d, *J* = 12.1 Hz, H-3β), 4.82 (1 H, d, *J* = 10.9 Hz, 1 x OCH_2Ph), 4.82 (1 H, d, *J* = 10.8 Hz, 1 x OCH_2Ph), 4.86 (1 H, d, *J* = 9.8 Hz, 1 x OCH_2Ph), 4.89 (1 H, d, *J* = 11.4 Hz, 1 x OCH_2Ph), 4.90 (1 H, d, *J* = 11.4 Hz, 1 x OCH_2Ph), 5.29 (1 H, at, *J* = 3.63 Hz H-1α), 7.07 - 7.41 (30 H, m, Ar-H). MS *m/z* (ESI$^+$): 468.2 $[M+NH_4]^+$, 473.2 $[M+Na]^+$.

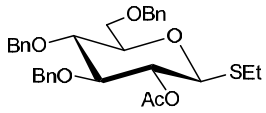
1,2-Di-O-acetyl-3,4,6-tri-O-benzyl-α/β-D-glucopyranose (110)[24]



247 (1.03g, 2.3 mmol) was dissolved in Ac_2O (30 mL) and pyridine (30 mL) and stirred at RT for 3h. TLC (petrol:EtOAc, 2:1) revealed complete conversion to product (R_f 0.7) and reaction was extracted into DCM (100 mL) washed with $NaHCO_3$ (3x100 mL) and satd brine (100 mL), concentrated under high vac and coevaporated with toluence yielding the desired compound a yellow oil (1.12 g, 90%) as an inseparable 3:1 mixture of α and β anomers.

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.7; $[\alpha]_D^{24}$ + 54.6 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{25}$ +164.1 (c = 1, $CHCl_3$)][24]; ^{1}H NMR (400 MHz CHLOROFORM-*d*) (α/β 3:1) δ ppm 1.92 (3 H, s, $2COCH_3\beta$), 1.95 (3 H, s, $OCOCH_3\alpha$), 2.07 (3 H, s, $OCOCH_3\beta$), 2.10 (3 H, s, $OCOCH_3\alpha$), 3.78-3.53 (7 H, m, H-4β,H-5α, H-5β, H-6abα,β), 3.83-3.80 (2 H, m H-3β, H-4α), 4.01-3.95 (1 H, dd, $J_{2,3}$ = 9.6 Hz, $J_{3,4}$ = 9.3 Hz, 1 H, H-3α), 4.72-4.51 (6 H, m, OCH_2Ph), 4.85-4.78 (6 H, m, OCH_2Ph), 5.07-5.02 (1 H, dd, $J_{2,3}$ = 10.2 Hz, $J_{1,2}$ = 3.6 Hz, H-2α), 5.11-5.03 (1 H, dd, $J_{1,2}$ = 8.7 Hz $J_{2,3}$ = 8.7 Hz, H-2β), 5.60 (1 H, d, $J_{1,2}$ = 8.4 Hz, H-1β), 6.29 (1 H, d, $J_{1,2}$ = 3.6 Hz, H-1α), 7.32-7.14 (m, 30 H, Ar-H); MS *m*/*z* (ESI^+): 552.2 $[M+NH_4]^+$.

Ethyl 2-O-acetyl-3,4,6-tri-O-benzyl-1-thio-β-D-glucopyranoside (248) [26]

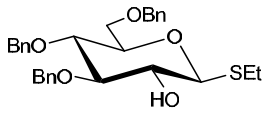


110 (871 mg, 1.63 mmol) was dried en vacuo for 1 hour. Sugar was then dissolved in 10 mL anhydrous DCM and to this was added 2 equivalents ethane thiol (2.4 mL, 3.2 mmol) and the mixture was cooled to O °C. BF_3 (5.2 mL) was added dropwise and the reaction was stirred for 1 hour until product (R_f 0.6) was detected in (petrol:EtOAc, 4:1) and starting material was almost completely consumed (R_f 0.4). After two rounds of flash chromatography (petrol:EtOAc, 6:1) followed by (petrol:EtOAc, 7:1), desired compound was obtained as an opaque oil (666 mg, 76%).

TLC (petrol:EtOAc, 4:1, v/v): R_f = 0.6; $[\alpha]_D^{24}$ + 25.4 (c = 1, $CHCl_3$) [Lit $[\alpha]_D^{22}$ +8.6 (c = 1, $CHCl_3$)[26]; ^{1}H NMR (400 MHz CHLOROFORM-*d*): δ 1.33 (3 H, t, J_{vic} 7.4 Hz, SCH_2CH_3), 2.05 (3 H, s, CH_3CO), 2.78 (2 H, dq, 1 H, J_{vic} = 7.5 Hz, Jg_{em} = 12.8 Hz, SCH_2CH_3), 3.57 (1 H, m, H-5), 3.74 (1 H, at, $\boldsymbol{J_{2,3} = J_{3,4}}$= 9.0 Hz, H-3), 3.77 (1 H, at, $J_{3,4} = J_{4,5}$ = 8.8 Hz, H-4), 3.78 (1 H, dd, $J_{6a,6b}$ = 11.3, $J_{5,6a}$ = 4.5 Hz, H-6a), 3.83 (1 H, dd, $J_{6a,6b}$ = 11.1, $J_{5,6b}$ = 1.9 Hz, H-6b), 4.43 (1 H, d, $J_{1,2}$ = 10.0 Hz, H-1β), 4.62 (1 H, d, J = 12.2 Hz, 1 x OCH_2Ph), 4.64 (1 H, d, J = 11.0 Hz, 1 x OCH_2Ph), 4.68 (1 H, d, J = 12.1 Hz, 1 x

OCH_2Ph), 4.76 (1 H, d, *J* = 11.4 Hz, 1 x OCH_2Ph), 4.86 (1 H, d, *J* = 10.6 Hz, 1 x OCH_2Ph), 4.88 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 5.10 (1 H, at, $\boldsymbol{J_{1,2} = J_{2,3}=}$ 9.2 Hz, H-2), 7.26–7.41 (m, 15H, Ar-H). MS *m*/*z* (ESI^+): 554.3 $[M+NH_4]^+$, 559.2 $[M+Na]^+$.

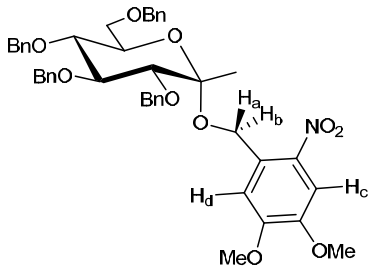
Ethyl -3,4,6-tri-O-benzyl-1-thio-β-D-glucopyranoside (24)[27]



248 (60 mg, 0.112 mmol) was dissolved in anhydrous methanol (10 mL). To this was added NaOMe (10 mg, 0.19 mmol) and the reaction was allowed to stir at RT for 3 h. Reaction was not complete by TLC (petrol:EtOAc, 7:1) showed product (R_f 0.3) and starting material (R_f 0.6). However due to concern about orthoester reformation reaction was quenched with dry ice to pH 7 and concentrated *in vacuo* and purified via column chromatography (petrol:EtOAc, 6:1) to give the desired product as an opaque oil (20 mg, 37%) as well as to recover starting material (36 mg, 60%).

TLC (petrol:EtOAc, 7:1, v/v): R_f = 0.3; $[\alpha]_D^{25}$ -11.9 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{24}$ -12.8 (c = 1, $CHCl_3$)][28]; ^{1}H NMR ($CDCl_3$): δ 1.33 (3 H, t, J_{vic} = 7.4 Hz, SCH_2CH_3), 2.75 (2H, dq, 1 H, J_{vic} 7.3 Hz, Jg_{em} 12.2 Hz, SCH_2CH_3), 3.51 (1 H, ddd, $J_{4,5}$ 9.2 Hz, $J_{5,6a}$ 4.8, $J_{5,6b}$ 2.0 Hz, H-5), 3.59 (1 H, at, *J* 9.1 Hz), 3.64 (1 H, at, *J* 9.1 Hz), 3.71 (1 H, dd, *J* 11.1, 4.8 Hz), 3.76 (3 H, dd, $J_{6a,6b}$ 10.8, $J_{5,6b}$ 2.0 Hz, H-6b), 4.32 (1 H, d, $J_{1,2}$ 9.3 Hz, H-1), 4.55 (1 H, d, *J* 11.8 Hz, 1 x OCH_2Ph), 4.57 (1 H, d, *J* 10.9 Hz, 1 x OCH_2Ph), 4.61 (1 H, d, *J* 12.1 Hz1 x OCH_2Ph), 4.85 (1 H, d, *J* = 10.9 Hz, 1 x OCH_2Ph), 4.87 (1 H, d, *J* 11.8 Hz, 1 x OCH_2Ph), 4.95 (1 H, d, *J* 11.4 Hz, 1 x OCH_2Ph), 7.13 - 7.42 (15 H, m, Ar-H)); MS *m*/*z* (ESI^+): 512.2 $[M+NH_4]^+$.

1-(2-nitro-4,5-dimethoxy-benzyl)-2,3,4,6-tetra-O-benzyl-1-deoxy-α-D-galacto-hept-2-ulopyranosyl (112)



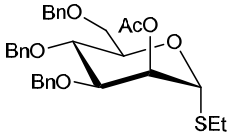
85 (162.4 mg, 0.29 mmol) and 2-nitro-4,5-dimethoxy-benzyl alcohol (57 mg, 0.27 mmol) were dried en vacuo for 1 hour then dissolved in anhydrous CH_2Cl_2 (7 mL) and added to a dry flask in the

presence of molecular sieves (ca. 100 mg) and stirred with the sieves for 30 minutes. To this was added TMSOTf (ca. 5 μl, 0.027mmol) at -40 °C in the under an Ar atmosphere. The resulting mixture was stirred for 1 h. TLC (petrol:EtOAc, 3:1) revealed the production of one new product spot (R_f 0.8) that was UV active and complete consumption of starting material. The reaction was then quenched by the addition 0.01 mL NEt_3, filtered through celite and concentrated. The crude product was purified by column chromatography (petrol:EtOAc, 5:1) to yield the product (150.3 mg, 76%).

TLC (petrol:EtOAc, 3:1, v/v): R_f = 0.8; $[\alpha]_D^{24}$ + 52.3 (c = 1, $CHCl_3$); ^{1}H NMR (500 MHz CHLOROFORM-*d*) δ ppm 1.59 (3 H, s, CH_3), 3.59 (1 H, dd, $J_{6a,6b}$ 9.1 Hz, $J_{5,6a}$ 5.7 Hz, H-6a), 3.66 (1 H, dd, $J_{6a,6b}$ = 8.9 Hz, $J_{5,6b}$ = 8.0 Hz, H-6b), 3.84 (1 H, ddd, $J_{5,6b}$ = 7.6 Hz, $J_{5,6a}$ = 5.7 Hz, $J_{4,5}$ = 1.3 Hz, H-5), 3.96 (2 H, d, $J_{2,3}$ = 10.1 Hz, H-2), 4.38 - 4.48 (2 H, m, H-4, H-3), 4.43 (1 H, d, *J* = 11.6 Hz, 1 x OCH_2Ph), 4.52 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.64 (2 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.71 (2 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.76 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.81 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.87 (1 H, d, *J* = 17.0 Hz, H_b), 4.96 (1 H, d, *J* = 11.7 Hz, 1 x OCH_2Ph), 5.07 (1 H, d, *J* = 10.9 Hz, 1 x OCH_2Ph), 5.09 (1 H, d, *J* = 17.3 Hz, H_a), 7.26 - 7.41 (32 H, m, Ar-H), 7.70 (1 H, s, Ar-H_c); ^{13}C NMR (126 MHz CHLOROFORM-*d*) *δ* ppm 21.0 (CH_3), 56.3 (C-b), 59.8 (C-a), 68.6 (C-6), 70.6 (C-5), 71.9 (OCH_2Ph), 73.5 (C-4), 74.3 (OCH_2Ph), 74.7 (OCH_2Ph), 76.0 (OCH_2Ph), 80.5 (C-3 or C-2), 80.6 (C-3 or C-2), 101.6 (C-1), 107.7 (C-d), 109.6 (C-c), 127.2, 127.5, 127.5, 127.7, 127.9, 128.0, 128.2, 128.3, 128.4, 128.5, 132.2, 137.8, 138.4, 138.5, 138.5, 138.7 (OCH_2Ph); IR (thin film): ν = br 3029.81 (OH), 2921.86, 1519.61 (N-O), 1453.86 (Ar C=C), 1325.95 (N-O), 1275.24, 1067.24, 1028.38, 920.44, 874.34, 794.75 cm^{-1}; HRMS *m/z* (ESI$^+$): $[M+Na]^+$ calcd. for $C_{44}H_{47}O_{10}Na^+$, 772.3098; found, 772.3089.

8.2.5 Mannose glycosylations: Synthesis of compounds 126-129

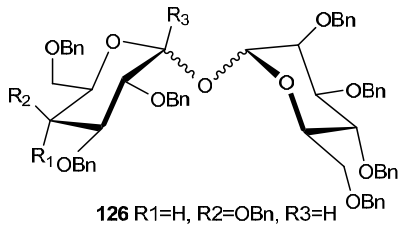
Ethyl 2-O-acetyl-3,4,6-tri-O-benzyl-l-thio-a-D-mannopyranoside (121)[28]



Ethanethiol (0.5 mL) was added to a solution of 3,4,6-tri-*O*-benzyl-β-D-manno-pyranose-1,2-methyl orthoacetate **120** (506 mg, 1 mmol) and $HgBr_2$, (36 mg, 0.1 mmol) in dry acetonitrile (3 mL). The mixture was heated to 60 °C under Ar for 4 h and concentrated. A solution of the residue in CH_2Cl_2 was washed with satd brine and $NaHCO_3$ and evaporated. The residue was eluted from a column of silica gel with (petrol:EtOAc, 7:1) to give the product as a clear oil (495mg, 77%).

TLC (petrol:EtOAc, 6:1, v/v): R_f = 0.4; $[\alpha]_D^{24}$ + 79.1 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{25}$ +84 (c = 1, $CHCl_3$)][29]; ^{1}H NMR (400 MHz CHLOROFORM-*d*) δ ppm 1.29 (3 H, t, J_{vic} = 7.5 Hz, CH_3), 2.18 (3 H, s, $OCOCH_3$), 2.65(1 H, dq, J_{vic} = 7.5 Hz, J_{gem} = 12.0 Hz, $CHCH_3$), 2.57 (1 H, dq, J_{vic} = 7.5 Hz, J_{gem} = 12.0 Hz, $CHCH_3$) 3.70 (1 H, dd, $J_{6a,6b}$ = 10.6 Hz, $J_{5,6a}$ = 2.0 Hz, H-6a), 3.86 (1 H, dd, $J_{6a,6b}$ = 10.9 Hz, $J_{5,6b}$ = 4.17 Hz, H-6b), (2 H, m, H-3 H-4), 4.16 - 4.20 (1 H, m, H-5), 4.49 (1 H, d, *J* = 10.9 Hz, 1 x OCH_2Ph), 4.50 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph), 4.54 (1 H, d, *J* = 11.1 Hz, 1 x OCH_2Ph), 4.70 (1 H, d, *J* = 11.6 Hz, 1 x OCH_2Ph), 4.71 (1 H, d, *J* = 11.9 Hz, 1 x OCH_2Ph), 4.87 (2 H, d, *J* = 10.6 Hz, 2 x OCH_2Ph), 5.34 (1 H, d, $J_{1,2}$ = 1.5 Hz, H-1), 5.46 (1 H, dd, $J_{2,3}$ = 2.8 Hz, $J_{1,2}$ = 1.8 Hz, H-2), 7.16-7.18 (2 H, m, Ar-H), 7.27 - 7.52 (13 H, m, Ar-H); MS *m/z* (ESI^+): 554.3 $[M+NH_4]^+$, 559.2 $[M+Na]^+$.

2-O-acetyl-3,4,6-tri-O-benzyl-1-deoxy-αβ-D-mannopyranosyl-2,3,4,6-tetra-O-benzyl-αβ-D-galactopyranoside (126) and 2,3,4,6-Tetra-O-benzyl-1-deoxy-α-D-gluco/galacto-hept-2-ulopyranosyl-(2→1)- 2-O-Acetyl-3,4,6-tri-O-benzyl-1-deoxy-αβ-D-mannopyranoside (127,128)



126 R1=H, R2=OBn, R3=H
127 R1=H, R2=OBn, R3=Me
128 R1=OBn, R2=H, R3=Me

Table 8.2. Manno thioglycoside glycosylation attempts

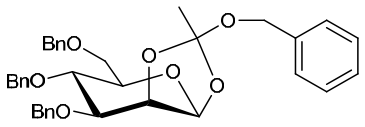
Entry	Acceptor	Time	Yield pdt as anomeric mixture	Yield side pdt	Recovered Donor
1	**94**	1 h	24%	10%	none
2	**93**	30 min	10%	48%	26%
3	**85**	1 h	26%	40%	53%

Procedure for glycosidation[30].

A solution of **121** (53.9 mg, 0.1 mmol), **94** (58.4 mg, 0.11 mmol), and N-iodosuccinimide (67.5 mg, 0.3 mmol) in dichloromethane (5 mL) was stirred for 0.5 h with 4 Å, molecular sieves (200 mg) under an Ar atmosphere at -10°C. A dilute solution of trifluoromethanesulfonic acid (0.02 mL in 3 mL CH_2Cl_2) was then added dropwise. Stirring was continued at the same temperature for another 0.5 h until product was detected (R_f 0.17) in (petrol:EtOAc, 6:1). The acid was neutralized with NEt_3.

The mixture was thoroughly washed with satd $NaHCO_3$ solution, water and satd brine solution and purified via flash chromatography (petrol:EtOAc, 7:1), yielding a mixture of the α,α, β,α and α,β products (indicated by coupling and HMBC) that could not be separated further (21 mg, 21%). HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{47}H_{53}O_{13}Na^+$, 1037.4452; found, 1037.4449.

4 3,4,6-tri-O-benzyl-1,2-O-(1-benzyloxyethylidene)-β-D-mannopyranose (129)[31]

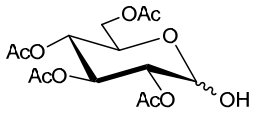


129 was isolated as a side product (37mg, 50%)

TLC (petrol:EtOAc, 6:1, v/v): R_f = 0.25; $[\alpha]_D^{25}$ + 30.9 (c = 1, $CHCl_3$); ^1H NMR (500 MHz CHLOROFORM-*d*) δ ppm 2.15 (3 H, s, CH_3), 3.70 (1 H, dd, $J_{6a,6b}$ = 10.7, $J_{5,6a}$ = 1.9 Hz, H-6a), 3.81 (1 H, dd, $J_{6a,6b}$ = 10.7 Hz, $J_{5,6b}$ = 4.7 Hz, H-6b), 3.86 (1 H, ddd, $J_{4,5}$ = 9.8 Hz, $J_{5,6b}$ = 4.7 Hz, $J_{5,6a}$ = 1.9 Hz, H-5), 3.91 (1 H, at, *J* = 9.5 Hz, H-4), 4.04 (1 H, dd, $J_{3,4}$ = 9.0 Hz, $J_{2,3}$ = 3.3 Hz, H-3), 4.48 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.51 (1 H, d, *J* = 11.6 Hz, 1 x OCH_2Ph), 4.54 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph)4.54 (2 H, d, *J* = 11.3 Hz, 2 x OCH_2Ph,), 4.71 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.71 (1 H, d, *J* = 11.3 Hz, 2 x OCH_2Ph), 4.86 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.95 (1 H, d, $J_{1,2}$ = 1.9 Hz, H-1), 5.43 (1 H, dd, $J_{2,3}$ = 3.5 Hz, $J_{1,2}$ = 1.9 Hz, H-2), 6.55 - 7.65 (20 H, m, Ar-H); MS *m/z* (ESI^+): 605.2 $[M+Na]^+$.

8.2.6 Synthesis of trehalose methyl ketosides

2,3,4,6-Tetra-O-acetyl-D-glucopyranose (84)[32]

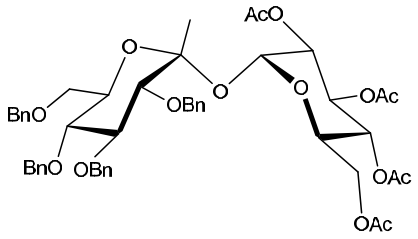


A solution of **130** (2.82 g, 9.804 mmol, 1 equiv) in DMF (20 mL) was treated with hydrazine acetate (720 mg, 7.8 mmol, 0.8 equiv) at RT overnight. TLC (petrol:EtOAc, 2:1, v/v) showed complete conversion to product (R_f 0.1) and disappearance of starting sugar (R_f 0.3) The reaction mixture was then treated with excess Amberlite 120 (H+) ion exchange resin to remove hydrazine derivatives, washed with brine and extracted into ether, concentrated and product was purified by silica gel column chromatography (petrol:EtOAc, 1:1, v/v) to yield the desired compound as a clear oil and a 3:1 mixture of anomers (1.92 g, 73.4 %).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.1; $[\alpha]_D^{24}$ +72.5 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{25}$ +75.3 (c = 1, $CHCl_3$)][32]; ^{1}H NMR (400 MHz, CHLOROFORM-d) δ ppm 1.96 - 2.32 (24 H, m, 4 x $OCOCH_3$-α, 4 x $OCOCH_3$-β), 3.75 (1 H, ddd, $J_{4,5}$ = 9.8 Hz, $J_{5,6a}$ = 4.8 Hz, $J_{5,6b}$ = 2.2 Hz, H-5β), 4.07 - 4.34 (4 H, m, H-6β,H-6'β, H-6α H-6'α), 4.27 (1 H, ddd, $J_{4,5}$ = 9.8 Hz, $J_{5,6a}$ = 4.2 Hz, $J_{5,6b}$ = 2.1 Hz H-5α), 4.74 (1 H, d, $J_{1,2}$ = 8.08 Hz, H-1β), 4.85 - 4.92 (3 H, m, H-3α, H-2αβ), 5.08 (2 H, at, J = 9.7 Hz, H-4αβ), 5.24 (1 H, at, J = 9.6 Hz, H-3β), 5.45 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1α), 5.53 (1H, at, J = 9.8 Hz, H-3α); MS *m/z* (ESI^+): 366.2 $[M+NH_4]^+$.

8.2.6.1 Synthesis of compound 32.

3,4,5,7-Tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2,3,4,6-tetra-O-acetyl-α-D-glucopyranoside **(32)**

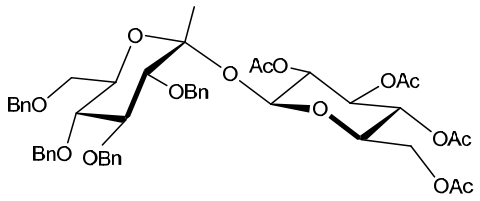


85 (134 mg, 0.24 mmol, 1 eq) and **84** (87 mg, 0.25 mmol, 1.05 eq) were dried *in vacuo* for 1 hour and then dissolved in anhydrous CH_2Cl_2 (7 mL) and added to a dry flask in the presence of molecular sieves (ca. 100 mg). To this was added TMSOTf (10 µl, 0.054 mmol, 0.2 eq) at – 40 °C in the under an Ar atmosphere. The resulting mixture was stirred for 3 h. (petrol:EtOAc, 5:2) revealed the production of two new product spots (R_f 0.3) and (R_f 0.25) and complete consumption of starting material. The reaction was then quenched by the addition 0.01 mL triethylamine, filtered through Celite® and concentrated. The crude product was purified by column chromatography (petrol:EtOAc, 5:2) to give colorless oil (189.5 mg, 93%) as a mixture of α,α and α,β products (7:1) that were separated during a second round of column chromatography (petrol:EtOAc, 3:1) to afford pure the desired compound (162 mg, 85.6 %) as a clear oil.

TLC (petrol:EtOAc, 5:2, v/v): R_f = 0.30; $[\alpha]_D^{24}$ + 80.9 (c = 1, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.46 (3 H, s, C-1'), 1.96 (3 H, s, 1 x $OCOC\underline{H}_3$), 2.01 (6 H, s, 2 x $OCOC\underline{H}_3$), 2.05 (3 H, s, 1 x $OCOC\underline{H}_3$), 3.32 (1 H, d, $J_{3',4'}$ = 9.6 Hz, H-3'), 3.56 - 3.64 (3 H, m, H-$7_{a'}$, H-7_b', H-5'), 3.82 (1 H, dd, $J_{6a,6b}$ = 12.4 Hz, $J_{5,6a}$ = 1.9 Hz, H-6_a), 3.88 (1 H, ddd, $J_{5,6}$ = 10.1 Hz, $J_{6,7a}$ = 3.8 Hz, $J_{6,7b}$ = 2.5 Hz H-6'), 4.05 - 4.11(2 H, m, H-4', H-6_b), 4.33 (1 H, ddd, $J_{4,5}$ = 10.4 Hz $J_{5,6a}$ = 4.3 Hz, $J_{5,6b}$ = 2.0 Hz, H-5), 4.48 (1 H, d, J = 12.3 Hz, 1x OCH_2Ph), 4.56 (1 H, d, J = 12.4 Hz, 1x OCH_2Ph), 4.56 (1 H, d, J = 11.0

Hz, 1x OCH_2Ph), 4.62 (1 H, d, J = 11.3 Hz 1 x OCH_2Ph), 4.85 (1 H, d, J = 11.1 Hz 1 x OCH_2Ph), 4.92 (1 H, d, J = 10.9 Hz, 1 x OCH_2Ph), 4.95-5.07 (4 H, m, H-4, 2 x OCH_2Ph, H-2), 5.37 (1 H, d, $J_{1,2}$ = 3.6 Hz, H-1), 5.55 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.7 Hz, H-3), 7.08 - 7.40 (20 H, m, Ar-H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 20.6, 20.6, 20.7, 20.7, (4 x OCO<u>C</u>H_3) 22.6 (C-1'), 61.6 (C-6), 67.1 (C-5), 68.1 (C-7'), 68.6 (C-4), 70.2 (C-2), 70.7 (C-3), 72.0 (C-6'), 73.5 (OCH_2Ph), 74.8 (OCH_2Ph), 75.5 (2 x OCH_2Ph), 78.3 (C-5'), 82.6 (C-4'), 84.5 (C-3'), 89.2 (C-1), 101.4 (C-2'), 127.5, 127.6, 127.7, 127.9, 128.3, 128.3, 128.4, 128.8, 130.8 (4 x OCH_2Ph), 138.0, 138.2, 138.3, 138.5 (4 x 1 C, 4 x OCH_2Ph), 169.6, 169.8, 170.2, 170.5 (4 x C=O); IR (thin film): 3063, 3030, 2925, 1753 (C=O), 1496, 1454, 1367, 1222, 1133, 1090, 1038, 956, 916, 873, 737 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{47}H_{53}O_{13}Na^+$, 907.3517; found, 907.3512; analysis (calcd., found for $C_{47}H_{53}O_{13}$): C (66.50,66.51), H (6.38,6.30).

3,4,5,7-Tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2,3,4,6-tetra-O-acetyl-β-D-glucopyranoside (131)

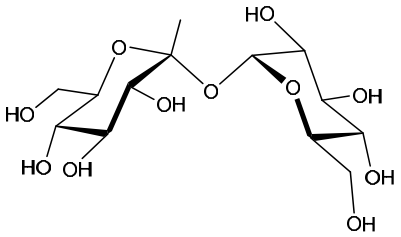


The titled compound was purified as the lower spot by TLC (petrol:EtOAc, 2:1) (R_f 0.25) from the reaction between **85** (134.4 mg, 0.242 mmol, 1 eq) and **84** (87.2 mg, 0.25 mmol, 1.05 eq) to produce **131** as a clear oil**.** (27 mg, 14.3 %).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.25; $[\alpha]_D^{24}$ + 17.7 (c = 1, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 1.43 (3 H, s, C-1'), 1.96 (1 H, s, 1 x OCOC<u>H</u>$_3$), 2.00 (5 H, s, 5 x OCOC<u>H</u>$_3$), 2.03 (5 H, m, 5 x OCOC<u>H</u>$_3$), 2.04 (1 H, s, 1 x OCOC<u>H</u>$_3$), 3.34 (1 H, d, $J_{3',4'}$ = 9.6 Hz, H-3'), 3.61 (1 H, ddd, $J_{4,5}$ = 12.4 Hz, $J_{5,6a}$ = 7.3 Hz, $J_{5,6b}$ = 2.5 Hz, H-5), 3.59-3.62 (1H, m, H-7$_b$'), 3.69-3.73 (2 H, m, H-5' H-7$_a$'), 4.03-4.07 (2 H, m, H-4', H-6$_a$), 4.12 (1 H, dd, $J_{6a,6b}$ = 12.4 Hz, $J_{5,6b}$ = 4.8 Hz, H-6$_b$), 4.18 (1 H, ddd, $J_{5',6'}$ = 6.6 Hz, $J_{6',7a'}$ = 3.0 Hz, $J_{6',7b'}$ = 2.0 Hz, H-6'), 4.48 (1 H, d, J = 11.9 Hz, 1 x OCH_2Ph), 4.53 (1 H, d, J = 10.9 Hz, 1 x OCH_2Ph), 4.59 (1 H, dd, J = 11.4 Hz, 1 x OCH_2Ph), 4.61 (1 H, dd, J = 12.3 Hz, 1 x OCH_2Ph), 4.84 (1 H, d, J = 10.86 Hz, 1 x OCH_2Ph), 4.89 (1 H, d, $J_{1,2}$ = 8.2 Hz , H-1), 4.89 (2 H, d, J = 10.4 Hz, 2 x OCH_2Ph), 4.94 (1 H, d, J = 11.1 Hz, 1 x OCH_2Ph), 5.00-5.06 (2H, m, H-4, H-2), 5.19 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.5 Hz, H-3), 7.16-7.18 (2 H, m, Ar-H), 7.27-7.35 (18 H, m, Ar-H); ^{13}C NMR (500 MHz, $CDCl_3$) δ 20.3, 20.6, 20.6, 20.7 (4 x OCO<u>C</u>H_3), 22.1 (C-1'), 62.0 (C-6), 68.4 (C-4), 68.5 (C-7'), 70.9 (C-2), 72.0 (C-6' or C-5), 72.4 (C-6' or C-5), 72.9 (C-3), 73.4 (OCH_2Ph), 74.9 (OCH_2Ph), 75.5 (OCH_2Ph), 75.5 (OCH_2Ph), 78.2 (C-5'), 82.5 (C-4'), 84.5 (C-3'), 94.6 (C-1), 102.4

(C-2'), 127.5, 127.5, 127.6, 127.6, 127.8, 127.8, 128.3, 128.3, 128.4(4 x OCH_2Ph), 138.1, 138.2, 138.4, 138.6 (4 x 1 C, 4 x OCH_2Ph), 169.0, 169.4, 170.3, 170.6 (4x C=O); IR (thin film): 3029, 2931, 1755 (C=O), 1496, 1454, 1366, 1230, 1211, 1038, 908, 737 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{47}H_{53}O_{13}Na^+$, 907.3517; found, 907.3512.

***1-Deoxy- α -D-gluco-hept-2-ulopyranosyl-(2→1)-α-D-glucopyranoside* (*32*)** [15]

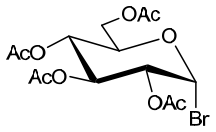


101 (120 mg, 0.13 mmol) was dissolved in 15 mL methanol with NaOMe (30 mg, 0.55 mmol, 4 eq) and stirred at room temperature for 1 hour until complete disappearance of starting material (R_f 0.55) and appearance of a new spot (R_f 0.05) was observed by TLC (petrol:EtOAc, 1:1) and deacetylated sugar was detected by ESI^+ *m/z* ($M+Na^+$): 741.3. Reaction was neutralized with DOWEX 50WX8 (H^+ form) ion exchange resin and concentrated *in vacuo*. The clear oil was redissolved in ethanol (15 mL) and to this was added basic alumina (70 mg) and 20% $Pd(OH)_2/C$ (80 mg) and palladium was activated by repeat purge flush cycles with hydrogen. Reaction was stirred under hydrogen atmosphere (balloon) at room temperature. After 72 h reaction was filtered through filter paper and filtrate was evaporated and purified by column chromatograph (ethyl acetate:methanol, 9:5) to give the desired, fully deprotected sugar, as a clear oil (40 mg, 83%).

TLC (MeOH:EtOAc 1:1 v/v) R_f=0.5; $[\alpha]_D^{25}$ +121.2 (0.32 in MeOH); [Lit. $[\alpha]_D^{27}$ +140.0 (c = 0.83 in MeOH)][15]; ^1H NMR (400 MHz, D_2O) δ 1.47 (3 H, s, C-1'), 3.24 (1 H, d, $J_{3,4}$ 9.9 Hz, H-3'), 3.35 (2 H, at, $J_{4,5}$ = $J_{5,6}$ = 9.6 Hz, H-5', H-4), 3.53 (1 H, dd, $J_{2,3}$ = 10.0 Hz, $J_{1,2}$ = 3.7 Hz, H-2), 3.59 (1 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6a}$ = 5.5 Hz), 3.61 (1 H, dd, $J_{7a,7b}$ = 11.8 Hz, $J_{6,7a}$ = 5.1 Hz, H-7a), 3.67 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 10.8 Hz, H-4), 3.73 - 3.83 (5 H, m, H-6b, H-5, H-3, H-4' H-7b'), 3.99 (1 H, ddd, $J_{5,6}$ = 10.1 Hz, $J_{6,7a}$ = 5.3 Hz, $J_{6,7b}$ = 2.3 Hz, H-6'), 5.22 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1); ^{13}C NMR (126 MHz, D_2O) δ 22.5 (C-1'), 60.5 (C-6), 60.5 (C-7), 69.8 (C-4), 69.9 (C-5'), 71.4 (C-5), 71.8 (C-2), 72.3 (C-6'), 72.5 (C-3), 72.8 (1 C-4'), 76.4 (C-3'), 91.5 (C-1), 101.0 (C-2'); MS *m/z* (ESI^+): 379.2$[M+Na]^+$; Spectroscopic data matches previously reported data[15].

8.2.6.2 Synthesis of compound 83

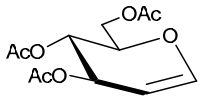
2,3,4,6-Tetra-O-acetyl-α-D-glucopyranosyl bromide (249)[33]



Hydrogen bromide (33% w/v in acetic acid, 10 mL) was added to **130** (2.98 g, 7.63 mmol) in DCM (20 mL). After 2 h TLC (petrol:EtOAc, 1:1), indicated formation of a product (R_f 0.7) with complete consumption of the starting material (R_f 0.6). Ice water (25 mL) was added and the reaction mixture stirred for 10 min. The phases were separated and the aqueous layer re-extracted with DCM (3 x 8 mL). The combined organic layers were washed with sodium hydrogen carbonate (saturated aqueous solution) until pH 8. The organic layer was filtered and concentrated *in vacuo* to afford the desired compound (3 g, 90 %) as a colorless oil.

TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.7; $[\alpha]_D^{21}$ +120 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{18}$ +122 (c = 1, $CHCl_3$)][33]; ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 2.03 (s, 3 H, 1 x $OCOCH_3$), 2.06 (3 H, s, 1 x $OCOCH_3$), 2.09 (3 H, s, 1 x $OCOCH_3$), 2.10 (3 H, s, 1 x $OCOCH_3$), 4.12 (1 H, d, $J_{6a,6b}$ = 10.8 Hz, H-6a), 4.35 (2 H, m, H-6', H-5), 4.84 (1 H, dd, $J_{2,3}$ = 9.8 Hz, $J_{1,2}$ = 4.1 Hz, H-2), 5.16 (1 H, at, $J_{3,4}$= $J_{4,5}$ = 9.8 Hz, H-4), 5.56 (1 H, at, $J_{2,3}$= $J_{3,4}$ = 9.6 Hz, H-3), 6.61 (1 H, d, $J_{1,2}$ = 4.1 Hz, H-1); MS *m/z* (ESI^+): 413.1 ($M+H^+$).

3,4,6-Tri-O-acetyl- D -glucal (132)[34]

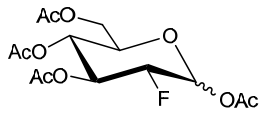


249 (3 g, 6.87 mmol) was dissolved in acetone (10 mL) water (10 mL). Zinc powder (6 g) was added and reaction was stirred at room temperature overnight. Upon completion of reaction by TLC, product is slightly less polar than starting material one spot (R_F 0.55) in (petrol:EtOAc, 1:1) and stains much darker in acid stain. Reaction was filtered through celite and product was concentrated *in vacuo* to produce the desired compound (2.31 g, 91%) as a yellow oil.

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.55; $[\alpha]_D^{21}$ -14 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{25}$ -21 (c = 1, $CHCl_3$)][34] ;^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 2.02 (3 H, s, 1 x $OCOCH_3$), 2.07 (3 H, s, 1 x $OCOCH_3$), 2.10 (3 H, s, 1 x $OCOCH_3$), 4.19 (1 H, dd, $J_{6a,6b}$ = 12.1, Hz, $J_{5,6a}$ = 3.1 Hz, H-6a), 4.22 - 4.30 (1 H, m, H-5), 4.41 (1 H, dd, $J_{6a,6b}$ = 12.0 Hz, $J_{5,6b}$ = 5.7 Hz, H-6b), 4.85 (1 H, dd, $J_{1,2}$ = 6.1 Hz, $J_{2,3}$ = 3.2 Hz,

H-2), 5.23 (1 H, dd, $J_{4,5}$ = 7.6 Hz, $J_{3,4}$ = 5.1 Hz, H-4), 5.32 - 5.40 (1 H, at, $J_{2,3}$= $J_{3,4}$= 4.9 Hz, H-3), 6.47 (1 H, dd, J = 6.1 Hz, 1.0 Hz, H-1); MS *m/z* (ESI^+): 295.2 $[M+Na]^+$.

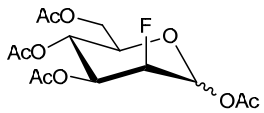
1,3,4,6-Tetra-O-acetyl-2-deoxy-2-fluoro-D-glucopyranose (134)[35]



132 (4.87g, 17.9 mmol) was dissolved in 5:1 $MeNO_2/H_2O$ (18mL). To this was added Selectfluor (7.24g, 20.33mmol) and the mixture was stirred at RT overnight according to the procedure by albert el al[36]. The mixture was then headed to reflux for 1 hour before the solvents were removed. Crude product was re-dissolved in DCM and passed through celite to remove the selectfluor salts. After concentration the residue was dissolved in Ac_2O and pyridine and stirred at RT for 1 hour. Upon completion TLC (petrol:EtOAc, 2:1) showed product (R_f 0.5) and reaction was concentrated *in vacuo* and purified by column chromatography (petrol:EtOAc, 1:1) to give the desired lactol as a 1:1 mixture of **133** and **134**. The manno and gluco derivatives were then partially separated by flash chromatography (petrol:EtOAc, 3:1) yielding pure the desired compound after two rounds of chromatography a clear oil (745 mg, 14%) as a 3:1 mixture of α and β anomers.

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.5; $[\alpha]_D^{24}$ +49.3 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{23}$ + 44.5 (c = 1, $CHCl_3$)][35]; ^{1}H NMR (CHLOROFORM-*d*) δ ppm 2.02-2.19 (30 H, 15 x $OCOCH_3$ -α 15 x $OCOCH_3$ -β), 3.85 (1 H, ddd, $J_{4,5}$ = 9.8 Hz, $J_{5,6a}$ = 2.3 Hz, $J_{5,6b}$ = 4.4 Hz, H-5β), 4.03-4.09 (3 H, m, H-6α, H-6β, H-5α), 4.25 (2 H, d, J = 4.3 Hz, H-6α, H-6β), 4.27-4.30 (2 H, m, H-6α, H-6β), 4.44 (1 H, ddd, $J_{2\beta,F}$ = 50.77 Hz, $J_{3\beta,2}$ = 8.7 Hz, $J_{2\beta,\beta}$ = 8.7 Hz, H-2β), 4,64 (2 H, ddd, $J_{2,F}$ = 48.51 Hz, $J_{1\alpha,2\alpha}$ = 9.6 Hz, $J_{2\alpha,3}$ = 4.0 Hz, H-2α), 5.07 (2 H, at, J = 9.7 Hz, H-4α, H-4β), 5.36 (1 H, dt, J= 14.2, 9.3 Hz H-3β), 5.53 (1 H, dt, J = 12.3, 9.5 Hz, H-3α), 5.77 (1 H, dd, $J_{1,2}$ = 8.2 Hz, $J_{1,F}$ = 3.2 Hz, H-1β), 6.4 (1H, d, $J_{1\alpha,2}$ = 4.04 Hz, H-1α); ^{19}F NMR (CHLOROFORM-*d*) δ ppm -201.47, -200.72; MS *m/z* (ESI^+): 368.2 $[M+NH_4]^+$, 723.1 $[2M+Na]^+$.

1,3,4,6-Tetra-O-acetyl-2-deoxy-2-fluoro-D-mannopyranose (133)[35]

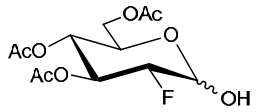


133 was isolated when 3,4,6-Tri-O-acetyl-D-glucal was treated with selectfluor and then acetylated with pyridine and Ac_2O during the synthesis of **83**. It was isolated as the lower spot (R_f 0.45) TLC

(petrol:EtOAc, 2:1) by flash chromatography (petrol:EtOAc, 3:1) (1.53g, 25%) By NMR product is a 6:1 mixture of α and β anomers.

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.45; $[\alpha]_D^{21}$ 6.7 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{23}$ +39.3 (c = 1, $CHCl_3$)][35]; ^{1}H NMR (500 MHz CHLOROFORM-*d*) δ ppm 2.03-2.22 (30 H, m, 15 x $OCOCH_3$ -α 15 x $OCOCH_3$ -β), 3.86 (1 H, ddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 4.7 Hz, $J_{5,6b}$ = 2.2, H-5β), 4.05 (1 H, ddd, $J_{4,5}$ = 10.2 Hz, $J_{5,6a}$ = 4.01 Hz, $J_{5,6b}$ = 2.1 Hz H-5α), 4.03-4.11 (2 H, m, H-6α, H-6β), 4.28 (2 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6}$ = 4.4 Hz, H-6α, H-6β), 4.45 (1 H, ddd, $J_{2,F}$ = 46.3 Hz, $J_{1,2}$ = 8.7 Hz, $J_{2,3}$ = 8.7 Hz, H-2β), 4.75 (3 H, ddd, $J_{2,F}$ = 46.3 Hz, $J_{1,2}$ = 2.2 Hz, $J_{2,3}$ = 2.2 Hz, H-2α), 5.04-5.11 (1 H, m, H-4β), 5.21-5.31 (2 H, m, H-3α, H-3β), 5.40 (3 H, at, $J_{3,4}$ = $J_{4,5}$ = 10.9 Hz, H-4α), 5.55 (1 H, dt, $J_{3,F}$ = 12.2 Hz, $J_{2,3}$ = $J_{3,4}$ = 9.5 Hz, H-3β), 5.78 (1 H, dd, $J_{1,2}$ = 5.4 Hz, $J_{1,F}$ = 3.2 Hz, H-1β), 6.27 (1 H, dd, $J_{1,2}$ = 6.5, $J_{1\alpha,F}$ = 2.4 Hz, H-1α); ^{19}F NMR (377 MHz, CHLOROFORM-*d*) δ ppm -203.82, -202.20 ; MS *m/z* (ESI$^+$): 368.2 $[M+NH_4]^+$.

3,4,6-Tri-O-acetyl-2-deoxy-2-fluoro- D -glucopyranoside (83)[37]

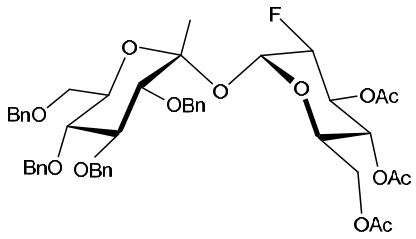


A solution of the **134**[35] (251.0mg, 0.716 mmol, 1 equiv) in DMF (20mL) was treated with hydrazine acetate (79.4 mg, 0.862 mmol, 1.2 equiv) at RT for 18 h. Upon completion by TLC (R_f 0.4) (petrol:EtOAc, 1:1), reaction was extracted into DCM, washed with satd. brine soln and satd $NaHCO_3$ soln and concentrated en vacuo. The residue was purified by column chromatography (petrol:EtOAc, 1:1) to yield the desired compound (157mg, 71%) as a clear oil. Product was found to be a 3:1 α/β mixture.

TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.4; $[\alpha]_D^{25}$ +87.2 (c = 1, $CHCl_3$) [Lit. $[\alpha]_D^{20}$ +100 (c = 0.8, $CHCl_3$)][37]; ^{1}H NMR (500 MHz CHLOROFORM-d) δ ppm 2.03-2.11 (18 H, 9 x $OCOCH_3$ -α 9 x $OCOCH_3$ -β), 3.78 (1 H, ddd, $J_{4,5}$ = 10.0 Hz, $J_{5,6a}$ = 4.7 Hz, $J_{5,6b}$ = 2.1 Hz, H-5β), 4.07-4.14 (3 H, m, H-6α, H-6β, H-5α), 4.20-4.34 (3 H, m, H-6α , H-6β, H-2β), 4.48 (1 H, ddd, $J_{2,f}$ = 49.5 Hz, $J_{2,3}$ = 9.6 Hz, $J_{1,2}$ = 3.8 Hz H-2α), 4.92 (1 H, dd, $J_{1,2}$ = 7.6 Hz, $J_{1,F}$ = 2.8 Hz, H-1β), 5.00-5.05 (2 H, m, H-4α, H-4β), 5.31-5.34 (1 H, m, H-3β), 5.49 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1α), 5.61 (1 H, adt, $J_{3,4}$ = 9.09 Hz, $J_{2,3}$ = 9.09 Hz, $J_{3,F}$ = 4.55 Hz, H-3α), ^{19}F NMR (1H) (377 MHz, CHLOROFORM-d) δ ppm -200.08, -199.4; MS *m/z* (ESI$^+$): 326.2 $[M+NH_4]^+$.

8.2.6.3 Synthesis of compound 33

3,4,5,7-Tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-3,4,6-tri-O-acetyl-2-deoxy-2-fluoro-α-D-glucopyranoside (100)

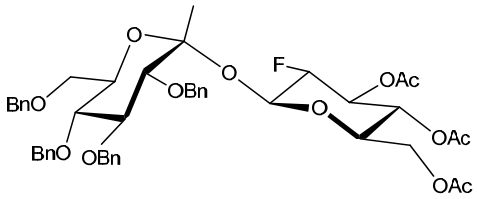


85 (81 mg, 0.15 mmol, 1 eq) and **83** (69 mg, 0.22 mmol, 1.5 eq) were dried *in vacuo* for 1 hour and then dissolved in anhydrous CH_2Cl_2 (7 mL) and added to a dry flask in the presence of molecular sieves (ca. 100 mg). To this was added TMSOTf (10 μl, 0.054 mmol, 0.3 eq) at –40 °C under an Ar atmosphere. The resulting mixture was stirred for 2 h upon which time TLC (petrol:EtOAc, 2.5:1) revealed the production of two very close new product spots (R_f 0.3) and (R_f 0.27) and complete consumption of starting material. The reaction was then quenched by the addition 0.01mL triethylamine, filtered through Celite® and concentrated to produce. The crude product was purified by column chromatography (petrol:EtOAc, 2:1) to give the desired compound as a colorless oil (104 mg, 82%) and a 6:1 mixture of α,α (89.1 mg) to α,β. None of the α,β product was obtained to purity.

TLC (petrol:EtOAc, 5:2, v/v): R_f = 0.30; $[\alpha]_D^{24}$ + 65.5 (c = 1.0, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.48 (3 H, s, C-1'), 2.01 (3 H, s, 1 x OCOC$\underline{H}_3$), 2.06 (3 H, s, 1 x OCOC$\underline{H}_3$), 2.08 (3 H, s, 1 x OCOC$\underline{H}_3$), 3.35 (1H, d, $J_{3',4'}$ = 9.4 Hz, H-3'), 3.62 (1 H, m, $J_{7a,7b}$ = 11.0 Hz, $J_{6,7a}$ =1.6 Hz, H-7b), 3.65 (1 H, at, $J_{5,6}$ = $J_{4,5}$ = 9.8 Hz, H-5'), 3.71 (1 H, d, $J_{7a,7b}$ = 11.3 Hz, $J_{6,7a}$ = 4.1 Hz, H-7a), 3.83 (1 H, dd, $J_{6b,6a}$ = 12.5 Hz, $J_{6b,5}$ = 2.0 Hz, H-6b), 3.97 (1 H, ddd, $J_{5',6'}$ = 9.9 Hz, $J_{6',7a'}$ = 2.0 Hz, $J_{6',7b'}$ = 1.7 Hz, H-6'), 4.06 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.4 Hz, H-4'), 4.08 (1 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{6a,6b}$ = 4.5 Hz, H-6a), 4.33 (1 H, ddd, $J_{4,5}$ = 10.4 Hz, $J_{5,6a}$ = 4.7 Hz, $J_{5,6b}$ = 2.2 Hz, H-5), 4.49 (1 H, d, J = 12.3 Hz, 1 x OCH_2Ph),), 4.54 (6 H, ddd, $J_{2,F}$ = 48.0 Hz, $J_{2,3}$ = 9.8 Hz, $J_{1,2}$ = 3.8 Hz, H-2), 4.57 (1 H, d, J = 11.1 Hz, 1 x OCH_2Ph), 4.60 (1 H, d, J = 12.3 Hz, 1 x OCH_2Ph), 4.62 (1 H, d, J = 10.4 Hz, 1 x OCH_2Ph), 4.83 (1 H, d, J = 10.9 Hz, 1 x OCH_2Ph), 4.93 (1 H, d, J = 11.4 Hz, 1 x OCH_2Ph), 4.96 (1 H, d, J = 12.3 Hz, 1 x OCH_2Ph), 4.99 (1 H, d, J = 11.4 Hz, 1 x OCH_2Ph) 4.99 (1 H, dt, $J_{4,5}$ = $J_{3,4}$ = 9.6 Hz, $J_{4,F}$ =3.0 Hz), 5.46 (1 H, d, $J_{1,2}$ = 3.9 Hz, H-1), 5.61 (1 H, dt, $J_{3,F}$ = 12.2 Hz, $J_{3,4}$ = 9.4 Hz, $J_{2,3}$ = 9.4 Hz H-3), 7.16 - 7.37 (20 H, m, Ar-H); ^{13}C NMR (126 MHz, $CHCl_3$): δ 20.6, 20.6, 20.8 (3 x OCO$\underline{C}H_3$), 22.9 (C-1'), 61.6 (C-6), 67.0 (C-5), 68.2 (C-4), 68.5 (C-7'), 71.3 (1 C, d, $J_{C-3,F}$ = 26.5 Hz, C-3) 71.3 (C-6'), 73.1 (OCH_2Ph), 74.6 (OCH_2Ph), 75.5 (OCH_2Ph), 75.5 (OCH_2Ph), 78.2 (C-5'), 82.6 (C-4'), 84.5 (C-3'), 87.2 (1 C, d, $J_{C-2,F}$ = 195.5 Hz, C-2), 88.8 (1 C, d, $J_{C-1,F}$ = 21 Hz, C-1), 101.5 (C-2'),

127.5, 127.5, 127.6, 127.7, 127.8, 127.9, 127.9, 128.2, 128.3, 128.4(4 x OCH_2Ph), 138.2, 138.4, 138.6 (4 x 1 C, 4 x OCH_2Ph), 169.7, 170.1, 170.5 (3 x C=O); ^{19}F NMR (1H) (377 MHz, $CHCl_3$): δ -197.5; IR(thin film) 3064, 3030, 2940, 1754 (C=O), 1605, 1540, 1497, 1454, 1366, 1223, 1132, 1089, 1064, 974, 949, 912, 872, 751, 737 $cm^{-}1$; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{47}H_{53}FO_{13}Na^+$, 867.3365; found, 867.3362.

3,4,5,7-Tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-3,4,6-tri-O-acetyl-2-deoxy-2-fluoro-β-D-glucopyranoside (135)

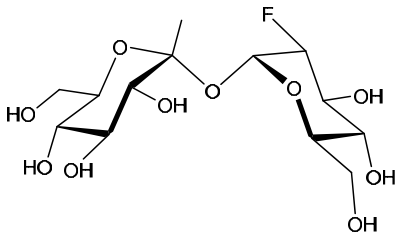


Following the procedure in Li, 2001[19] a solution of **99** (24 mg, 0.045 mmol, 1 eq) and **83** (29 mg, 0.094 mmol, 2 eq) and 4 Å molecular sieves (100 mg) in dry CH_2Cl_2 (5 mL) was stirred under nitrogen atmosphere at room temperature for 30 min. The solution was cooled to −78 °C and TMSOTf (5 μL, 0.027 mmol, 0.5 eq) was added. The reaction mixture was stirred at −78 °C for 30 minutes. The reaction was monitored by TLC (petrol:EtOAc, 2.5:1) and upon completion one broad product spot was visible (R_f 0.3) with disappearance of starting glycal **99** (R_f 0.8) and reaction was quenched with triethylamine (20 μl) and passed through Celite®. After the removal of the solvent under reduced pressure, the residue was purified by silica gel column chromatography (petrol:EtOAc, 5:2) as the eluent to afford the products (35 mg, 95%) 1:1 mixture of the α,α and α,β and gluco and manno sugars. From these only the titled compound could be isolated to purity as a clear oil (10 mg, 28%).

TLC (petrol:EtOAc, 5:2, v/v): R_f = 0.30; $[\alpha]_D^{24}$ + 39.7 (c = 0.37, $CHCl_3$); ^1H NMR (500 MHz,$CDCl_3$): *δ* 1.42 (3 H, C-1'), 2.02 (3 H, s, 1 x OCOC<u>H</u>$_3$), 2.05 (3 H, s, 1 x OCOC<u>H</u>$_3$), 2.09 (3 H, 1 x OCOC<u>H</u>$_3$), 3.40 (1 H, d, $J_{3',4'}$ = 9.8 Hz, H-3'), 3.59 (1 H, ddd, $J_{4,5}$ = 10.0 Hz, $J_{5,6a}$ = 5.1 Hz, $J_{5,6b}$ = 2.5 Hz, H-5), 3.63-3.70 (3H, m, H-7a', H-7b', H-5'), 4.99-4.11 (2H, m, H-6b, H-4'), 4.09 (1 H, dd, $J_{6a,6b}$ = 11.8 Hz, $J_{6a,6b}$ = 5.5 Hz, H-6a), 4.21 (1 H, ddd, $J_{5',6'}$ = 10.1 Hz, $J_{6',7a'}$ = 3.9 Hz, $J_{6',7b'}$ = 1.42 Hz, H-6'), 4.37 (1 H, ddd, $J_{F,2}$ = 50.0 Hz, $J_{2,3}$ = 9.1 Hz, $J_{1,2}$ = 8.2 Hz, H-2) 4.50 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.54 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.59 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.71 (1 H, d, *J* = 11.7 Hz, 1 x OCH_2Ph), 4.85 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.87 (1 H, d, *J* = 11.0 Hz 1 x OCH_2Ph), 4.91 (1 H, d, *J* = 13.3 Hz, 1 x OCH_2Ph) 4.94-4.98 (2 H, m, H-4, 1 x OCH_2Ph), 5.01 (1 H, dd, $J_{1,2}$ = 7.7 Hz $J_{1,F}$ = 2.4 Hz, H-1), 5.30 (1 H, dt, $J_{3,F}$ = 14.2 Hz, $J_{2,3}$ 9.2 Hz, $J_{3,4}$ 9.2 Hz, H-3), 7.16 - 7.34 (20 H, m, Ar-H); ^{13}C NMR (126 MHz, $CDCl_3$): *δ* 22.3, 22.6, 23.1 (3 x OCO<u>C</u>H$_3$), 28.8 (CH_3), 61.9 (C-6), 68.2

(1 C, d, $J_{C\text{-}4,F}$ = 5.0 Hz, C-4), 68.6 (C-7'), 72.1 (C-5), 72.6 (C-6'), 73.0 (1 C, d, $J_{C\text{-}3,F}$ = 18.9 Hz, C-3), 73.1 (OCH_2Ph), 73.5 (OCH_2Ph), 75.0 (OCH_2Ph), 75.5 (OCH_2Ph), 78.2 (C-5'), 82.7 (C-4'), 84.2 (C-3'), 88.9 (1 C, d, $J_{C\text{-}2,F}$ = 189 Hz, C-2), 94.0 (1 C, d, $J_{C\text{-}1,F}$ = 23.9 Hz, C-1), 102.6 (C-2'), 127.5, 127.69, 127.7, 127.7, 127.8, 127.8, 128.0, 128.3, (4 x OCH_2Ph), 138.1, 138.6 (4x 1 C. 4 x OCH_2Ph), 169.4, 170.13, 170.5 (3 x C=O); ^{19}F NMR (377 MHz, $CDCl_3$) δ -198.52; IR (thin film): 2921, 1754 (C=O), 1586, 1549, 1513, 1495, 1453, 1366, 1242, 1196, 1100, 1055, 1029, 735, 698 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{47}H_{53}FO_{13}Na^+$, 867.3365; found, 867.3362.

1-Deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2-deoxy-2-fluoro-α-D-glucopyranoside (33)



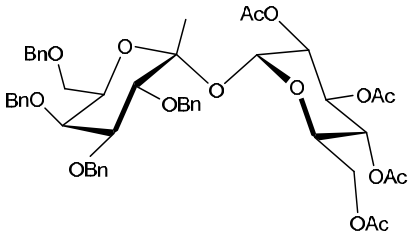
100 (66.6 mg, 0.078 mmol, 1 eq) was dissolved in 15 mL methanol with NaOMe (20 mg, 0.37 mmol, 4.8 eq.) and stirred at room temperature for 3 hours until complete disappearance of starting material (R_f 0.55) and appearance of a new spot (R_f 0.05) was observed by TLC (petrol:EtOAc, 1:1) and deacetylated sugar was detected by ESI^+ *m/z* ($M+Na^+$): 741.3. Reaction was neutralized with DOWEX 50WX8 (H^+ form) ion exchange resin and concentrated *in vacuo*. The clear oil was re-dissolved in methanol (10 mL) and to this was added basic alumina (5.7 mg) and 20% $Pd(OH)_2/C$ (112 mg) and reaction was stirred under hydrogen atmosphere (balloon) at room temperature. After 48 h. reaction was filtered through filter paper and filtrate was evaporated and purified by column chromatography (EtOAc:MeOH, 7:3) and Isolute SPE C_{18} cartridge to give the desired, fully deprotected sugar, as a clear oil (27 mg, 98%).

TLC (EtOAc:MeOH, 1:1, v/v): R_f = 0.55; $[\alpha]_D^{25}$ + 95.5 (c = 0.42, MeOH); ^1H NMR (500 MHz, D_2O): δ 1.45 (3 H, s, C-1'), 3.24 (1 H, d, $J_{3,4}$ = 9.8 Hz, H-3'), 3.33 (1 H, at, $J_{4,5}$ = $J_{5,6}$= 10.1 Hz, H-5'), 3.40 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4), 3.63 (1 H, dd, $J_{6a,6b}$ = 11.8 Hz, $J_{5,6a}$ = 5.8 Hz, H-6a), 3.66 (1 H, dd, $J_{7a,7b}$ = 11.9 Hz, $J_{6,7a}$ = 5.7 Hz, H-7a'), 3.67 (1 H, t, J = 9.5 Hz, H-4'), 3.72 - 3.79 (3 H, m, H-6b, H-7b', H-6'), 3.80 (1 H, ddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 5.0 Hz, $J_{5,6b}$ = 2.4 Hz, H-5), 4.02 (1 H, dt, $J_{3,F}$ = 12.9, $J_{3,4}$ = 9.4 Hz, $J_{2,3}$ = 9.4 Hz, H-3), 4.36 (1 H, ddd, $J_{2,F}$ = 49.5 Hz, $J_{2,3}$ = 9.8 Hz, $J_{1,2}$ = 3.8 Hz, H-2), 5.45 (1 H, d, $J_{1,2}$ = 4.1 Hz, H-1); ^{13}C NMR (126 MHz, D_2O): δ 22.27 (C-1'), 60.31 (C-6), 60.59 (C-7), 69.2 (1 C, d, $J_{C\text{-}4,F}$ = 7.5 Hz , C-4), 69.8 (C-5'), 71.2 (1 C, d, $J_{C\text{-}3,F}$ = 17.6 Hz, C-3), 71.9 (C-5), 72.5 (C-6'), 72.9 (C-4'), 76.1 (C-3'), 88.8 (C-1), 89.7 (1 C, d, $J_{C2\text{-}F}$ = 189 Hz, C-2), 101.1 (C-2'); ^{19}F

NMR (377 MHz, D_2O) δ -197.2 (dt, $J_{2,F}$ = 49.3 Hz, $J_{1,F}$ = $J_{3,F}$ = 12.6 Hz); HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{13}H_{23}F\ O_{10}Na^+$ 381.1173; found, 381.1167.

8.2.6.4 Synthesis of compound 34.

3,4,5,7-Tetra-O-benzyl-1-deoxy-α-D-galacto-hept-2-ulopyranosyl-(2→1)-2,3,4,6-tetra-O-acetyl-α-D-glucopyranoside (142)

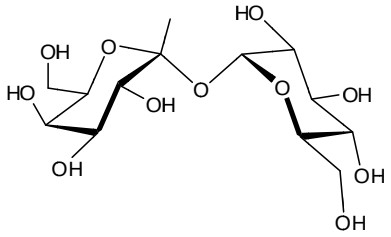


93 (176.1 mg, 0.317 mmol, 1 eq) and **84** (146.2 mg, 0.419 mmol, 1.3 eq) were dried *in vacuo* for 1 hour and then dissolved in anhydrous CH_2Cl_2 (8 mL) and added to a dry flask in the presence of molecular sieves (ca. 100 mg). To this was added TMSOTf (10µl, 0.54 mmol, 0.17 eq) at –40 °C under an Ar atmosphere. The resulting mixture was stirred for 30 min until TLC (petrol:EtOAc, 5:2) revealed the production of two new product spots (R_f 0.3) and (R_f 0.27) and complete consumption of starting material. The reaction was then quenched by the addition 0.1 mL triethylamine, filtered through Celite® and concentrated to produce. The crude product was purified by column chromatography (petrol:EtOAc, 5:2) to give as a colorless oil (246 mg, 88%) as a mixture of α,α and α,β products (5:1) that were separated during further chromatography (petrol:EtOAc, 3:1) to yield the desired compound as exclusively α,α (180.7 mg, 65%)**.**

TLC (petrol:EtOAc, 5:2, v/v): R_f = 0.30; $[\alpha]_D^{24}$ + 72.0 (c = 1.0, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.46 (3 H, s, C-1'), 1.94 (3H, s, 1x OCOC$\underline{H}_3$) 2.01 (6 H, s, 2 x OCOC$\underline{H}_3$), 2.04 (3 H, s, 1 x OCOC$\underline{H}_3$), 3.40 (1 H, dd, $J_{7a,7b}$ 8.9 Hz, $J_{6,7a}$ 5.3 Hz, H-7b'), 3.56 (1 H, dd, $J_{7a,7b}$ = 8.4, $J_{6,7a}$ = 8.4 Hz, H-7a'), 3.75 (1 H, dd, $J_{6a,6b}$ = 12.5 Hz, $J_{5,6b}$ = 2.2 Hz, H-6b), 3.83 (1 H, d, $J_{3,4}$ 9.9 Hz, H-3'), 4.01 (1 H, Br s, H-5'), 4.02 (1 H, dd, $J_{6a,6b}$ = 12.2 Hz, $J_{5,6a}$ = 4.2 Hz, H-6a), 4.07 (1 H, dd, $J_{4,5}$ = 9.9, $J_{5,6}$ = 2.6 Hz, H-5'), 4.10 (1 H, ddd, J = 7.2, 4.3, 1.2 Hz, H-6') 4.34 (1 H, d, J = 11.8 Hz, 1 x OCH_2Ph), 4.38 (3 H, ddd, $J_{4,5}$ = 10.4 Hz, $J_{5,6a}$ = 4.3 Hz, $J_{5,6b}$ = 2.2 Hz, H-5), 4.40 (1 H, d, J = 11.6 Hz, 1 x OCH_2Ph), 4.57 (1 H, d, J = 11.6 Hz, 1 x OCH_2Ph), 4.61 (1 H, d, J = 11.1 Hz, 1 x OCH_2Ph), 4.82 (1 H, Br. s, 1 x OCH_2Ph), 4.95 (1 H, d, J = 11.4 Hz, 1 x OCH_2Ph), 4.99 (1 H, dd, $J_{2,3}$ = 10.2 Hz, $J_{1,2}$ = 3.6 Hz, H-2), 5.02 (1 H, at, $J_{1,2}$ = $J_{2,3}$ = 10.1, H-2), 5.04 (2 H, d, J = 10.41 Hz, 1 x OCH_2Ph), 5.06 (1 H, d, J = 10.2 Hz, 1 x OCH_2Ph), 5.33 (1 H, d, $J_{1,2}$ = 3.6 Hz, H-1), 5.53 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.8 Hz, H-3), 6.85 - 7.64 (20 H, m, Ar-H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 20.3, 20.6, 20.6, 20.7 (4 x OCO$\underline{C}H_3$), 22.4 (C-1'), 61.6 (C-6), 67.0 (C-5), 68.4 (C-4), 68.4 (C-7) 70.5 (C-2), 70.5 (2 C, s, C-4 and C-6'), 72.6

(OCH_2Ph), 73.5 (OCH_2Ph), 74.2 (C-5'), 74.5 (OCH_2Ph), 75.6 (OCH_2Ph), 80.1 (C-4'), 80.86 (C-3'), 89.3 (C-1), 102.3 (C-2'), 127.5, 127.5, 127.6, 127.7, 127.8, 127.9, 128.1, 128.2, 128.2, 128.3, 128.4(4 x OCH_2Ph), 137.6, 138.4, 138.6, 138.7 (4 x 1 C, 4 x OCH_2Ph), 169.6, 169.9, 170.2, 170.6 (4 x C=O) IR (thin film): 3030, 2932, 2868, 1752 (C=O), 1497, 1454, 1368, 1221, 1139, 1100, 1038, 963, 922, 877, 845, 772.3, 753 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{47}H_{53}O_{13}Na^+$, 907.3517; found, 907.3511; analysis (calcd., found for $C_{47}H_{53}O_{13}$): C (66.81,66.59), H (6.37, 6.42).

1-deoxy-α-D-galacto-hept-2-ulopyranosyl-(2→1)-α-D-glucopyranoside **(34)** [15]

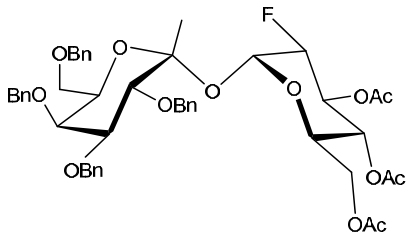


142 (163.2 mg, 0.184 mmol, 1 eq) was dissolved in 15 mL methanol with NaOMe (46.2 mg, 0.855 mmol) and stirred at room temperature for 1 hour until complete disappearance of starting material (R_f 0.55) and appearance of a new spot (R_f 0.05) was observed by TLC (petrol:EtOAc, 1:1) and deacetylated sugar was detected by ESI^+ *m/z* (M+Na): 739.2. Reaction was neutralized with DOWEX 50WX8 (H^+ form)ion exchange resin and concentrated *in vacuo*. The resulting clear oil was redissolved in ethanol (15mL) and solution was degassed with nitrogen. To this was added basic alumina (50 mg) and 20% $Pd(OH)_2/C$ (70 mg) and reaction was stirred under hydrogen atmosphere (balloon) at room temperature. After 72 h. reaction was filtered through filter paper and filtrate was evaporated and purified by column chromatography (9:5 ethyl acetate/methanol) to give the desired, fully deprotected sugar, as a clear oil.

(23 mg, 32%). TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.5; $[\alpha]_D^{25}$ + 156.5 (c = 0.46, MeOH); [Lit. $[\alpha]_D^{22}$ +164.9 (c = 0.64, MeOH)][15]; ^1H NMR (500 MHz, $CDCl_3$): δ 1.51 (3 H, s, C-1'), 3.38 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4), 3.55 (1 H, d, $J_{3,4}$ = 10.1 Hz, H-3'), 3.56 (1 H, dd, $J_{2,3}$ = 10.1 Hz, $J_{1,2}$ = 3.8 Hz, H-2), 3.66 (1 H, dd, $J_{7a,7b}$ = 12.0 Hz, $J_{6,7b}$ = 5.4 Hz, H-7b'), 3.69 (1 H, dd, $J_{7a,7b}$ = 12.0 Hz, $J_{6,7b}$ = 7.7 Hz, H-7a), 3.71 (1 H, dd, $J_{6a,6b}$ = 12.6 Hz, $J_{5,6a}$ = 5.4 Hz, H-6a), 3.77 - 3.85 (3 H, m, H-3, H-5, H-6b), 3.94 (1 H, dd, $J_{4,5}$ = 3.5 Hz, $J_{5,6}$ = 1.1 Hz, H-5'), 3.95 (1 H, dd, $J_{3,4}$ = 10.7 Hz, $J_{4,5}$ = 3.5 Hz, H-4'), 4.22 (1 H, ddd, $J_{6,7a}$ = 7.1 Hz, $J_{6,7b}$ = 5.0 Hz, $J_{5,6}$ = 1.4 Hz, H-6'), 5.26 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1); ^{13}C NMR (126 MHz, $CDCl_3$): δ 25.2 (C-1'), 63.0 (C-6), 63.6 (C-7'), 71.8 (C-5'), 72.0 (C-4'), 72.4 (C-4), 73.8 (C-6'), 73.9 (C-2), 74.2 (C-5), 75.0 (C-3), 75.9 (C-3'), 94.0 (C-1), 103.7 (C-2'); MS *m/z* (ESI^+): 379.2$[M+Na]^+$. Spectroscopic data matches with that previously reported [15]

8.2.6.5 Synthesis of compound 35.

3,4,5,7-Tetra-O-benzyl-1- deoxy-α-D-galacto-hept-2-ulopyranosyl-(2→1)-3,4,6-tri-O-acetyl-2-deoxy-2-fluoro-α-D-glucopyranoside (143)

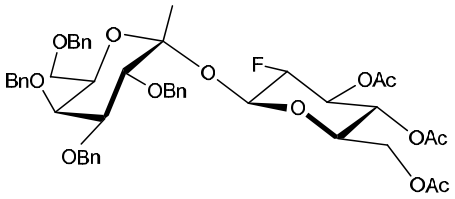


93 (88.7 mg, 0.16 mmol, 1.15 eq) and **83** (44.1 mg, 0.14 mmol, 1 eq) were dried *in vacuo* for 1 hour then dissolved in anhydrous CH_2Cl_2 (7 mL) and added to a dry flask in the presence of molecular sieves (ca. 100 mg). To this was added TMSOTf (5µl, 0.027 mmol, 0.18 eq) at –40 °C under an Ar atmosphere. The resulting mixture was stirred for 1 hour upon which time TLC (petrol:EtOAc, 5:2) revealed the production of two new product spots (R_f 0.3) and (R_f 0.27) and complete consumption of starting material. The reaction was then quenched by the addition 0.01mL triethylamine, filtered through Celite® and concentrated. The crude product was purified by column chromatography (petrol:EtOAc, 2:1) to give a 6:1 mixture of the α,α and α,β products (106 mg, 90%). Further purification in (petrol:EtOAc, 3:1) yielded the desired compound (70.7 mg, 60%) as a colorless oil. α,β product was not obtained to purity through this method.

TLC (petrol:EtOAc, 5:2, v/v): R_f = 0.30; $[\alpha]_D^{24}$ + 46.6 (c = 0.78, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.47 (3 H, s, C-1'), 2.01 (3 H, s, 3 x OCOC$\underline{H}_3$), 2.06 (3H, s, 3 x OCOC$\underline{H}_3$), 2.09 (3H, s, 3 x OCOC$\underline{H}_3$), 3.52 (2 H, ad, *J* 6.6 Hz, H-7a', H-7b'), 3.83 (2 H, dd, $J_{6a,6b}$ 12.5 Hz, $J_{5,6b}$ 2.4 Hz, H-6b), 3.84 (1H, d, $J_{3,4}$ = 9.8 Hz, H-3'), 3.98 (1 H, dd, $J_{4,5}$ = 2.7 Hz, $J_{5,6}$ = 1.4 Hz, H-5'), 4.05 (1 H, dd, $J_{3,4}$ = 9.9 Hz, $J_{4,5}$ = 2.7 Hz, H-4'), 4.08 - 4.15 (2 H, m, H-6', H-6a), 4.33 (1 H, ddd, $J_{4,5}$ = 10.3 Hz, $J_{5,6a}$ = 4.8 Hz, $J_{5,6b}$ = 2.2 Hz, H-5), 4.42 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.47 (1 H, d, *J* = 11.9 Hz, 1 x OCH_2Ph), 4.53 (1 H, ddd, $J_{2,F}$ = 50.8 Hz $J_{2,3}$ = 9.6 Hz, $J_{1,2}$ = 3.9 Hz, H-2), 4.58 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.63 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.76 (1 H, d, *J* = 11.9 Hz, 1 x OCH_2Ph), 4.79 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.93 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 5.00 (2 H, t, *J* = 9.7 Hz, H-4), 5.02 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 5.44 (1 H, d, $J_{1,2}$ = 4.1 Hz, H-1), 5.61 (1 H, adt, $J_{3,F}$ = 12.3 Hz, $J_{3,4}$ = $J_{2,3}$ = 9.5 Hz, H-3), 6.87 - 7.54 (20 H, m, Ar-H), 7.30; ^{13}C NMR (126 MHz, $CDCl_3$): δ 20.6, 20.6, 20.8 (3 x OCO$\underline{C}H_3$), 22.7 (C-1'), 61.63 (C-6), 67.0 (C-5) 68.0 (1 C, d, $J_{C-4,F}$ =7 Hz, C-4,), 68.7 (C-7'), 71.1 (1 C, d, $J_{C-3,F}$ = 18.9 Hz, C-3), 71.1 (C-6'), 72.6 (OCH_2Ph), 73.1 (OCH_2Ph), 74.4 (OCH_2Ph or C-5'), 74.5 (OCH_2Ph or C-5'), 75.5 (OCH_2Ph), 80.1 (C-4'), 80.6 (C-3') 87.5 (1 C, d, $J_{C-2,F}$ 194 Hz, C-2), 88.7 (1 C, d, $J_{C-1,F}$ = 21 Hz, C-1), 102.0 (C-2'), 127.5, 127.5, 127.6, 127.7, 127.7, 128.0, 128.2, 128.3, 128.4 (4 x OCH_2Ph), 138.1, 138.4, 138.7, 138.7 (4 x 1 C, 4 x OCH_2Ph), 169.7,

170.1, 170.5 (3 x C=O); ^{19}F NMR (1H) (377 MHz, $CDCl_3$): δ -197.6; IR (thin film): 2922, 2852, 2408, 1747 (C=O), 1496, 1453, 1367, 1220, 1055, 772 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{47}H_{53}FO_{13}Na^+$, 867.3365, found, 867.3362.

3,4,5,7-Tetra-O-benzyl-1-deoxy-α-D-galacto-hept-2-ulopyranosyl-(2→1)-3,4,6-tri-O-acetyl-2-deoxy-2-fluoro- β -D-glucopyranoside (152)

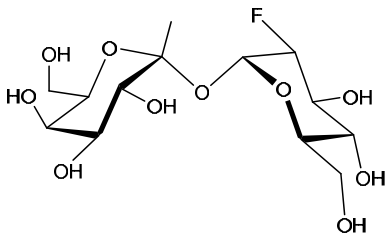


Following the procedure in Li, 2001[19]. A solution of **102** (168 mg, 0.31 mmol, 1 eq), and **83** (178 mg, 0.58 mmol, 1.6 eq) and 4 Å molecular sieves (ca.100 mg) in dry CH_2Cl_2 (5 mL) was stirred under nitrogen atmosphere at room temperature for 30 min. The solution was cooled to −78 °C and TMSOTf (5μL, 0.027 mmol, 0.05 eq.) was added. The reaction mixture was stirred at −78 °C for 30 minutes. The reaction was monitored by TLC (petrol:EtOAc, 5:2) and upon completion, the appearance of one broad spot (R_f 0.3) and disappearance of starting glycal (R_f 0.8) was detected. Reaction was quenched with triethylamine (20 μ1) and passed through Celite® and concentrated. The residue was applied on a silica gel column chromatography (petrol:EtOAc, 2:1) to afford the products as a 1:2 mixture of the α,α and α,β and gluco and manno sugars (75 mg, 30%). From these only the titled compound could be isolated to purity as a clear oil (57 mg, 22%).

TLC (petrol:EtOAc, 5:2, v/v): R_f=0.30; $[\alpha]_D^{24}$ + 18.7 (c = 1.0, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.41 (3 H, s, C-1'), 2.00 (6 H, s, 2 x OCOC$\underline{H}_3$) 2.05 (3 H, s, 1 x OCOC$\underline{H}_3$), 2.08 (3 H, s, 1 x OCOC$\underline{H}_3$), 3.41 (1 H, dt, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 3.2 Hz, $J_{5,6b}$ = 3.2 Hz H-5), 3.45 (1 H, dd, $J_{7a,7b}$ = 5.9 Hz, $J_{6,7b}$ = 3.1 Hz, H-7_b'), 3.50 - 3.55 (1 H, m, H-7_a'), 3.90 (1 H, d, $J_{3,4}$ = 9.8 Hz, H-3'), 3.96 (2 H, dd, J = 5.9 Hz, J = 3.0 Hz, H-6_a, H-6_b), 3.99 (1 H, at, $J_{4,5}$ = $J_{5,6}$ = 1.0 Hz, H-5'), 4.02 (1 H, dd, $J_{3,4}$ = 9.8 Hz, H-4'), 4.27 - 4.35 (1H, m, H-2, H-6'), 4.41 (1 H, d, J = 9.8 Hz, 1 x OCH_2Ph), 4.45 (1 H, d, J = 11.3 Hz, 1 x OCH_2Ph), 4.63 (1 H, d, J = 11.7 Hz, 1 x OCH_2Ph), 4.72 (1 H, d, J = 11.3 Hz, 1 x OCH_2Ph), 4.76 (2 H, s, 2 x OCH_2Ph), 4.93 - 4.97 (3 H, m, H-1, 2 x OCH_2Ph), 4.99 (1 H, dd, $J_{4,5}$ = 7.2 Hz $J_{3,4}$ = 4.9 Hz, H-4), 5.25 (1 H, dt, $J_{3,F}$ = 14.1 Hz, $J_{2,3}$ = $J_{3,4}$ = 9.2 Hz.H-3), 7.19 - 7.44 (20 H, m); ^{13}C NMR (126 MHz, $CDCl_3$) δ 20.4, 20.5, 20.6 (3 x OCO$\underline{C}H_3$), 22.3 (C-1'), 61.4 (C-6), 67.9 (1 C, d, $J_{C-4,F}$ = 7.6 Hz, C-4) 69.9 (C-7'), 71.8 (C-6'), 71.9 (C-5), 72.8 (OCH_2Ph), 73.3 (1 C, d, $J_{C-3,F}$ = 20.1 Hz, C-3), 73.7 (OCH_2Ph), 74.3 (C-5'), 74.6 (OCH_2Ph), 75.6 (OCH_2Ph), 80.1 (C-3'), 80.4 (C-4') 88.8 (C-2,d, $J_{C-2,F}$ = 190.8 Hz), 93.6 (C-1, d, $J_{C-1,F}$ = 22.9 Hz), 103.1 (C'-2), 127.4, 127.5, 127.5, 127.6, 127.7, 128.1, 128.1, 128.2, 128.2, 128.3, 128.4 (4 x OCH_2Ph) 138.0, 138.3, 138.5, 138.6 (4 x 1 C, 4 x

OCH_2Ph), 170.1, 170.5, 171.1 (3 x C=O); ^{19}F NMR (377 MHz, $CDCl_3$) δ -198.73; IR (thin film): 3063, 3030, 2923, 2857, 1752 (C=O), 1604, 1548, 1496, 1454, 1367, 1230, 1212, 1139, 1100, 1054, 919, 808, 736 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{47}H_{53}O_{13}Na^+$, 867.3365; found, 867.3383.

1-deoxy-α-D-galacto-hept-2-ulopyranosyl-(2→1)-2-deoxy-2-fluoro-α-D-glucopyranoside (35)

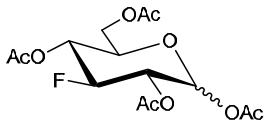


143 (42 mg, 0.05 mmol, 1 eq) was dissolved in 20 mL methanol with NaOMe (31 mg, 0.6 mmol, 12 eq) and stirred at room temperature for 3 hours until until complete disappearance of starting material (R_f 0.55) and appearance of a new spot (R_f 0.05) was observed by TLC (petrol:EtOAc, 1:1) and deacetylated sugar was detected by ESI^+ *m/z* (M+Na): 741.3. Reaction was neutralized with DOWEX 50WX8 (H^+ form)ion exchange resin and concentrated *in vacuo*. The resulting clear oil was re-dissolved in methanol (10mL) and solution was degassed with nitrogen. To this was added basic alumina (31.0 mg) and 20% $Pd(OH)_2/C$ (73 mg) and reaction was stirred under hydrogen atmosphere (balloon) at room temperature. After 96 h. reaction was filtered through filter paper and filtrate was evaporated. Further purification was obtained utilizing Isolute SPE C_{18} cartridge to give the desired, fully deprotected sugar, as a clear oil (20 mg, 100%).

TLC (EtOAc:MeOH, 1:1, v/v): R_f=0.55; $[\alpha]_D^{25}$ +122.8 (0.42, MeOH);^1H NMR (500 MHz, D_2O) δ 1.47 (3 H, s, C-1'), 3.39 (1 H, at, $J_{3,4} = J_{4,5}$ = 9.5 Hz, H-4), 3.52 (1 H, d, $J_{3,4}$ = 10.4 Hz, H-3'), 3.61 (1 H, dd, $J_{7a,7b}$ = 12.0 Hz, $J_{6,7a}$ = 5.4 Hz, H-7b'), 3.64 (1 H, dd, $J_{7a,7b}$ = 12.0 Hz, $J_{6,7a}$ = 7.3 Hz, H-7°'), 3.66 (1 H, dd, $J_{6a,6b}$ = 12.6 Hz, $J_{5,6a}$ = 5.4 Hz, H-6a), 3.76 (1 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6b}$ = 2.2 Hz, H-6b), 3.79 (1 H, dd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 4.8 Hz, $J_{5,6b}$ = 2.6 Hz, H-5), 3.83 (1 H, dd, $J_{3,4}$ = 10.1, $J_{4,5}$ = 3.2 Hz, H-4'), 3.90 (1 H, dd, $J_{4,5}$ = 3.1 Hz, $J_{5,6}$ = 1.0 Hz, H-5'), 3.97 (1 H, ddd, $J_{6,7a}$ = 7.1, $J_{6,7b}$ = 5.5 Hz, $J_{5,6}$ = 1.0 Hz, H-6'), 4.02 (1 H, dt, $J_{3,F}$ = 13.0 Hz, $J_{3,4}$ =$J_{2,3}$ = 9.4 Hz, H-3), 4.37 (1 H, m, $J_{2,F}$ = 49.5 Hz, $J_{2,3}$ = 9.5 Hz, $J_{1,2}$ = 3.8 Hz, H-2), 5.46 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1); ^{13}C NMR (126 MHz, D_2O): δ 22.44 (C-1'), 60.3 (C-6), 61.3 (C-7'), 69.3 (1 C, d, $J_{C-4,F}$ = 7.5 Hz, C-4), 69.3 (C-5'), 69.6 (C-4'), 71.2 (1 C, d, $J_{C-3,F}$ = 17.6 Hz, C-3), 71.6 (1 C, $J_{C-5,F}$ = 6.3 Hz, C-5), 71.8 (C-6'), 73.2 (C-3'), 88.8 (1 C, d, $J_{C-1,F}$ = 21.4 Hz, C-1), 89.9 (1 C, d, $J_{C2,F}$ = 189 Hz, C-2), 101.34 (C-2'); ^{19}F NMR (377 MHz, D_2O): δ -197.5 (dt, $J_{2,F}$ = 51.6, $J_{1,F}$ = $J_{3,F}$ =14.9 Hz); HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{13}H_{23}FO_{10}Na^+$, 381.1173; found, 381.1165.

8.2.6.6 Synthesis of compound 36.

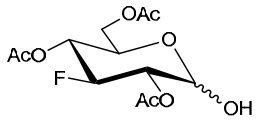
1,2,4,6-Tetra-O-acetyl-3-deoxy-3-fluoro-αβ-D-glucopyranoside (153)[38,39]



67 (100 mg, 0.55 mmo1) was dissolved in anhydrous pyridine (10 mL). To this was added acetic anhydride (10 mL) and reaction mixture was stirred at RT for 3 h until complete conversion was detected by ESI mass spectrometry and TLC (petrol:EtOAc, 1:1with conversion to product (R_f 0.65) and disappearance of starting material (R_f 0.0). Reaction was concentrated under reduced pressure and then coevaporated with toluene to yield the desired product (194 mg, 99%, 1:1 α,α:α,β) as a slightly yellow oil.

TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.65; $[\alpha]_D^{25}$ = 35.8 (c = 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 2.05 - 2.18 (24 H, m, 12 x OCOC$\underline{H}_3$α, 12 x OCOC$\underline{H}_3$β), 3.74 (1 H, dddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 4.4 Hz, $J_{5,6b}$ = 2.1 Hz, $J_{5,F}$ = 1.0 Hz, H-5β), 4.03 (1 H, ddd, $J_{4,5}$ = 10.1, $J_{5,6a}$ = 3.8 Hz, $J_{5,6b}$ = 2.8 Hz, H-5α), 4.07 - 4.15 (2 H, m, H-6bα, H-6bβ), 4.25 (1 H, dd, $J_{6a,6b}$ = 12.3, $J_{5,6a}$ = 4.4 Hz, H-6aα), 4.27 (1 H, dd, $J_{6a,6b}$ = 12.4 Hz, $J_{5,6a}$ = 4.5 Hz, H-6aβ), 4.60 (1 H, dt, $J_{3,F}$ = 51.9, $J_{2,3}$ = 9.2 Hz, $J_{3,4}$ = 9.2 Hz, H-3β), 4.82 (1 H, dt, $J_{3,F}$ = 53.4, $J_{2,3}$ = 9.4 Hz, $J_{3,4}$ = 9.4 Hz, H-3α), 5.17 (1 H, ddd, $J_{2,F}$ = 12.1, $J_{2,3}$ = 9.8 Hz, $J_{1,2}$ 3.9 Hz, H-2α), 5.21 - 5.34 (3 H, m, H-4α, H-4β, H-2β), 5.65 (1 H, d, $J_{1,2}$ = 8.3 Hz, H-1β), 6.34 (1 H, at, $J_{1,2}$ = 3.8 Hz, $J_{1,F}$ = 3.8 Hz H-1α); ^{19}F NMR (1H) (377 MHz, CHLOROFORM-*d*) δ ppm -195.92 (3F-α), -199.96 (3F-β); MS *m/z* (ESI^+): 392.2 [M+MeCN].

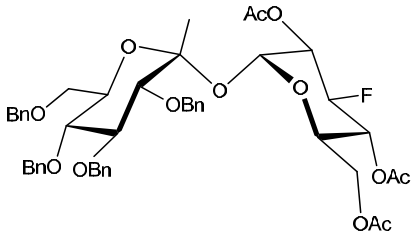
2,4,6-tri-O-acetyl-3-deoxy-3-fluoro-αβ-D-glucopyranoside (137)[38]



153 (173 mg, 0.49 mmol, 1 equi) was dissolved in anhydrous DMF (10 mL). To this was added hydrazine acetate (38 mg, 0.41 mmol, 0.85 equi) and reaction was stirred at RT for 2.5 h upon which time near-complete conversion was detected by TLC (petrol:EtOAc, 1:1) with disappearance of starting sugar (R_f 0.65) and appearance of product (R_f 0.5). Reaction was neturalized by stirring Dowex amberlite cation exchange resin 50WX8 for 10 minutes, upon which time Dowex was removed by filtration and reaction was concentrated under reduced pressure. Column chromatography (petrol:EtOAc, 1:1) yielded the desired product as a clear oil (70 mg, 48%) as a mixture of (α:β 4:1) as well as recovered starting material (6 mg, 4%).

TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.5; $[\alpha]_D^{25}$ = 44.1 (c = 0.95, $CHCl_3$); ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 2.08 - 2.18 (18 H, m, 9 x OCOC$\underline{H}_3$α, 9 x OCOC$\underline{H}_3$β), 3.66 (1 H, dddd, $J_{4,5}$ = 11.4, $J_{5,6a}$ = 3.6 Hz, $J_{5,6b}$ = 2.3, $J_{5,F}$ = 1.1 Hz, H-5β), 4.05 - 4.29 (5 H, m, H-6aα, H-6bα, H-6aβ, H-6bβ H-5α), 4.60 (1 H, dt, $J_{3,F}$ = 52.1 Hz, $J_{2,3}$ = 9.2 Hz, $J_{3,4}$ = 9.2 Hz H-3β), 4.66 (1 H, d, $J_{1,2}$ = 7.1 Hz, H-1β), 4.89 (1 H, dt, $J_{3,F}$ = 53.0 Hz, $J_{2,3}$ = 9.4 Hz, $J_{3,4}$ = 9.4 Hz, H-3α), 4.94 - 5.01 (2 H, m, H-2αβ), 5.18 - 5.27 (2 H, m, H-4αβ), 5.50 (1 H, at, $J_{1,2}$ = 3.0 Hz, $J_{1,F}$ = 3.0 Hz, H-1α); ^{19}F NMR (1H) (377 MHz, CHLOROFORM-*d*) δ ppm -195.43 (3F-β), -200.40 (3F-α); MS *m/z* (ESI^-): 367.2 [M+HAc-H].

3,4,5,7-tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2,4,6-tri-O-acetyl-3-deoxy-3-fluoro-α-D-glucopyranoside (144)

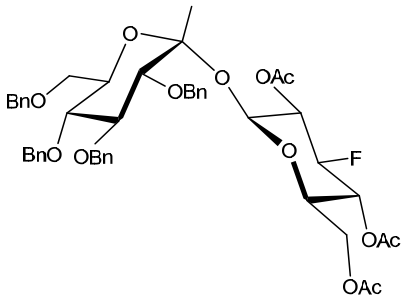


85 (116 mg, 0.21 mmol) and **137** (56 mg, 0.18 mmol) were dried under reduced pressure for one hour and then dissolved in anhydrous DCM (8 mL) and added to activated molecular sieves (ca 100 mg). Sugars were stirred with sieves at RT for 1 hour and were then cooled to –40 °C. TMSOTf (10 µl, 0.054 mmol, 0.3 eq) was added and 1 hour later additional TMSOTf (10 µl, 0.054 mmol, 0.3 eq) was added to the reaction. Reaction was stirred for 5.5 h, upon which time conversion was detected by

TLC (petrol:EtOAc, 2:1) with conversion to product α,α (R_f 0.5) and α,β (R_f 0.48) and disappearance of starting sugars (R_f 0.6) and (R_f 0.05). Reaction was then quenched with 1 drop triethylamine, filtered through Celite® to remove molecular sieves, concentrated under reduced pressure and purified by column chromatography (petrol:EtOAc, 5:2). The desired product was obtained as a clear oil (79 mg, 52%) α,α (63 mg) and α,β (16 mg) as well as recovered **137** (10 mg).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.5; $[\alpha]_D^{25}$ = 66.3 (c = 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.47 (3 H, s, C-1'), 2.02 (3 H, s, 1 x OCOC$\underline{H}_3$), 2.04 (3 H, s, 1 x OCOC$\underline{H}_3$), 2.13 (3 H, s, 1 x OCOC$\underline{H}_3$), 3.33 (1 H, d, $J_{3,4}$ = 9.6 Hz, H-3'), 3.58 - 3.70 (3 H, m, H-7a',H-7b', H-5'), 3.81 (1 H, dd, $J_{6a,6b}$ = 12.4 Hz, $J_{5,6b}$ = 1.6 Hz, H-6b), 3.90 (1 H, ddd, $J_{5,6}$ = 10.1 Hz, $J_{6,7a}$ = 4.0 Hz, $J_{6,7b}$ = 2.3 Hz, H-6'), 4.02 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.3 Hz, H-4'), 4.05 (1 H, dd, $J_{6a,6b}$ = 12.9 Hz, $J_{5,6a}$ = 4.5 Hz, H-6a), 4.25 (1 H, ddd, $J_{4,5}$ = 10.2 Hz, $J_{5,6a}$ = 4.2 Hz, $J_{5,6b}$ = 2.0 Hz, H-5), 4.50 (1 H, d, J = 12.1 Hz, 1 x OCH_2Ph), 4.58 (1 H, d, J = 9.3 Hz, 1 x OCH_2Ph), 4.60 (1 H, d, J = 12.1 Hz, 1 x OCH_2Ph), 4.64 (1 H, d, J = 11.4 Hz, 1 x OCH_2Ph), 4.87 (1 H, d, J = 10.9 Hz, 1 x OCH_2Ph), 4.90 (1 H, dt, $J_{3,F}$ = 54.0 Hz, $J_{2,3}$ = 9.4 Hz, $J_{3,4}$ = 9.4 Hz, H-3), 4.92 (2 H, m, 2 x OCH_2Ph), 4.99 (1 H, d, J = 11.1 Hz, 1 x OCH_2Ph), 5.08 (1 H, ddd, $J_{2,F}$ = 11.4 Hz, $J_{2,3}$ = 9.9 Hz, $J_{1,2}$ = 3.5 Hz, H-2), 5.21 (1 H, ddd, $J_{4,F}$ = 13.6 Hz, $J_{4,5}$ = 10.4 Hz, $J_{3,4}$ 9.1 Hz, H-4), 5.39 (1 H, at, $J_{1,2}$ = 3.4 Hz, H-1), 7.08 - 7.44 (20 H, m, Ar-H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 20.6, 20.6, 20.7 (3 x OCO$\underline{C}H_3$), 22.5 (C-1'), 61.5 (C-6), 67.1 (1 C, d, $J_{C-5,F}$ = 6.3 Hz, C-5), 68.3 (1 C, d, $J_{C-4,F}$ = 17.4 Hz, C-4), 68.5 (C-7'), 70.7 (1 C, d, J_{C-2F} = 16.4 Hz, C-2) 72.1(C'-6), 73.5 (OCH_2Ph), 74.8 (OCH_2Ph), 75.5 (OCH_2Ph), 75.5 (OCH_2Ph), 78.2 (C-5'), 82.74 (C-4'), 84.5 (C-3'), 89.6 (1C = d, $J_{C-1,F}$ = 8.8 Hz C-1), 89.6 (1 C, d, $J_{C-3,F}$ = 189 Hz, C-3), 101.4 (C-2'), 127.5, 127.6, 127.7, 127.7, 127.8, 127.9, 127.9, 128.2, 128.3, 128.4 (4 x OCH_2Ph), 137.8, 138.2, 138.4, 138.6 (4 x 1 C, 4 x OCH_2Ph), 169.3, 169.7, 170.5 (3 x C=O) ;^{19}F(^{1}H) NMR(377 MHz, $CDCl_3$): δ -199.9; IR (thin film): 3063, 3030, 2923, 2864 (C=CH), 1751 (C=O), 1496, 1453, 1367, 1219, 1131, 1089, 1067, 1038, 736; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{47}H_{53}FO_{13}$ NH_4^+,867.3362; found, 867.3338; analysis (calcd., found for $C_{47}H_{53}O_{13}$): C (66.81,66.82), H (6.32,6.37).

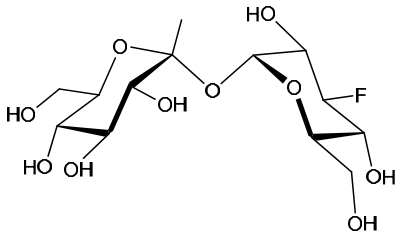
3,4,5,7-Tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2,4,6-tri-O-acetyl-3-deoxy-3-fluoro-β-D-glucopyranoside (154)



154 was isolated as the lower spot TLC (petrol:EtOAc, 2:1) (R_f 0.48) of the reaction between **137** and **85** as a clear oil (16 mg, 10%).

TLC (petrol:EtOAc, 2:1, v/v): R_f=0.48; $[\alpha]_D^{25}$ = 12.4 (c = 1.0, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.43 (3 H, s, C-1'), 2.04 - 2.20 (9 H, m, 3 x OCOC$\underline{H}_3$), 3.34 (1 H, d, $J_{3,4}$ = 9.8 Hz, H-3'), 3.52 (1 H, adt, $J_{4,5}$ = 9.9 Hz, $J_{5,6a}$ = $J_{5,6b}$ = 3.3 Hz, H-5), 3.59 (1 H, dd, $J_{7a,7b}$ = 10.7, $J_{6,7b}$ = 1.9 Hz, H-7b'), 3.69 (1 H, at, $J_{4,5}$ = $J_{4,6}$ = 9.5 Hz, H-5'), 3.70 (1 H, dd, $J_{7a,7b}$ = 10.4 Hz, $J_{6,7a}$ = 3.8 Hz, H-7a'), 4.06 (2 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.5 Hz, H-4'), 4.07 (2 H, dd, $J_{6a,6b}$ = 12.1 Hz, $J_{5,6b}$ = 2.2 Hz, H-6b), 4.12 (2 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6a}$ = 4.7 Hz, H-6a), 4.16 (1 H, ddd, $J_{5,6}$ = 10.2 Hz, $J_{6,7a}$ = 3.5 Hz, $J_{6,7b}$ = 2.0 Hz, H-6'), 4.51 (1 H, dt, $J_{3,F}$ = 53.0 Hz, $J_{2,3}$ = 9.1 Hz, $J_{3,4}$ = 9.1 Hz H-3), 4.47 (1 H, d, J = 12.0 Hz, 1 x OCH_2Ph), 4.53 (1 H, d, J = 11.0 Hz, 1 x OCH_2Ph), 4.59 (1 H, d, J = 12.3 Hz, 1 x OCH_2Ph), 4.60 (1 H, d, J = 11.3 Hz, 1 x OCH_2Ph), 4.81 (1 H, d, $J_{1,2}$ = 7.9 Hz, H-1), 4.84 (1 H, d, J = 10.7 Hz, 1 x OCH_2Ph), 4.86 - 4.90 (2 H, m, 2 x OCH_2Ph), 4.94 (1 H, d, J = 11.3 Hz, 1 x OCH_2Ph), 5.14 (2 H, m, H-4, H-2), 7.21 - 7.37 (20 H, m, Ar-H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 20.4, 20.5, 20.6 (3 x OCO$\underline{C}H_3$), 22.0 (C-1'), 61.9 (C-6), 68.4 (1 C, d, $J_{C-4,F}$ = 12.6 Hz, C-4), 68.5 (C-7') 70.9 (1 C, d, $J_{C-2,F}$ = 18.9 Hz, C-2), 71.2 (1 C, d, $J_{C-5,F}$ 7.5 Hz C-5), 72.4 (H-6'), 73.4 (O$\underline{C}H_2$Ph), 74.9 (O$\underline{C}H_2$Ph), 75.5 (O$\underline{C}H_2$Ph), 75.6 (O$\underline{C}H_2$Ph), 78.1 (H-5'), 82.5 (C-4'), 84.5 (C-3'), 91.7 (1 C, d, $J_{C-2,F}$ = 191.5 Hz, C-3), 94.2 (1 C, d, $J_{C-1,F}$ = 12.6 Hz, C-1), 102.4 (C-2'), 127.5, 127.6, 127.6, 127.6, 127.8, 128.3, 128.3, 128.4, 128.4 (4 x OCH_2Ph), 138.2, 138.2, 138.4, 138.6 (4x 1 C. 4 x OCH_2Ph), 168.8, 169.15, 170.6 (3 x C=O); IR (thin film): 3062, 3030, 2923, 2855 (C=CH), 1752 (C=O), 1496, 1453, 1368, 1218, 1151, 1126, 1063, 1043. 737; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{47}H_{53}FO_{13}Na^+$,867.3362; found, 867.3338.

1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-3-deoxy-3–fluoro-α-D-glucopyranoside (36)

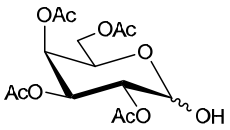


144 (42 mg, 0.049 mmol, 1 eq) was dissolved in anhydrous methanol (10 mL), to this was added sodium methoxide (20 mg, 0.37 mmol, 7.5 eq) and reaction was stirred under argon atmosphere for 1.5 hours until complete conversion to product was detected by TLC (petrol:EtOAc, 2:1) (R_f 0.0) with dissappearance of starting sugar (R_f 0.5). Reaction was neutralized with DOWEX 50WX8 (H^+ form) cation exchange resin. Resin was removed by filtration and filtrate was concentrated under reduced pressure and redissolved in ethanol (10 mL). Reaction mixture was circulated through a Thales Nano H Cube® 10% Pd/C cartridge at 70 bar, 25 °C for 1 hour. Near complete deprotection was detected by TLC (MeOH:EtOAc, 1:1) (R_f 0.45). Reaction mixture was partitioned between water and DCM and aqueous layer was lyophilized. Further purification was obtained utilizing Isolute SPE C18 cartridge to yield the desired product (15.6 mg, 87%) as a white, amorphous solid.

TLC (MeOH:EtOAc, 1:1, v/v): R_f=0.45; $[\alpha]_D^{25}$ = 108.2 (c = 0.22, MeOH); ^{1}H NMR (500 MHz, D_2O): δ 1.44 (3 H, s, C-1'), 3.23 (1 H, d, $J_{3,4}$ = 9.8 Hz, C-3'), 3.34 (1 H, at, $J_{4,5}$ = $J_{5,6}$ = 9.6 Hz, C-5'), 3.62 - 3.75 (6 H, m, H-6a, H-6b, H-7a, H-7b, H-4, H-4'), 3.79 (1 H, ddd, $J_{2,F}$ = 16.4 Hz, $J_{2,3}$ = 9.0 Hz, $J_{1,2}$ = 3.6 Hz, H-2), 3.80 (3 H, ddd, $J_{4,5}$ = 11.7 Hz, $J_{5,6a}$ = 6.3 Hz, $J_{5,6b}$ = 3.5 Hz, H-5'), 3.97 (2 H, ddd, $J_{5,6}$ = 10.1 Hz, $J_{6,7a}$ = 5.4 Hz, $J_{6,7b}$ = 2.2 Hz, H-6'), 4.68 (1 H, dt, $J_{3,F}$ = 55.0 Hz, $J_{3,4}$ = 9.1 Hz, $J_{2,3}$ = 9.1 Hz, H-3), 5.26 (2 H, at, J = 3.6 Hz, H-1);^{13}C NMR (126 MHz, D_2O): δ 22.4 (C-1'), 60.1 (C-6), 60.5 (C-7'), 68.1 (1 C, d, $J_{C4,F}$ = 16.4 Hz, C-4), 69.8 (C-5'), 69.9 (1 C, d, $J_{C2,F}$ = 16.4 Hz, C-2), 71.3 (1 C, d, $J_{C5,F}$ = 7.5 Hz, C-5), 72.3 (C-6') 72.8 (C-4'), 76.3 (C-3'), 91.7 (1 C, d, $J_{C3,F}$ = 11.3 Hz, C-1), 94.5 (1 C, d, $J_{C3,F}$ = 180.1 Hz, C-3), 101.1 (C-2'); ^{19}F NMR (377 MHz, D_2O): δ -199.2 (dtt, $J_{3,F}$ = 55.1, $J_{2,F}$ = $J_{4,F}$ = 14.9 Hz, $J_{1,F}$ = $J_{5,F}$ = 3.4 Hz); HRMS *m/z* (ESI^+): [M-H^-] calcd. for $C_{13}H_{22}F O_{10}^-$, 357.1202; found, 357.1197.

8.2.6.7 Synthesis of compound 145

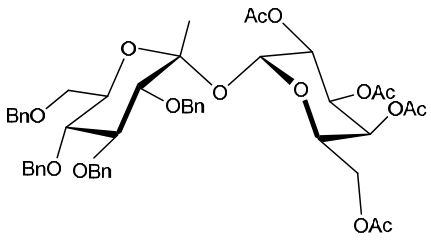
2,3,4,6-Tetra-acetyl-D-galactopyranose (139)[40]



A solution of the peracetylated galactose (6.25 g, 16.14 mmol, 1 equiv) in anhydrous DMF (15mL) was treated with hydrazine acetate (1.78 g, 19.3 mmol, 1.2 equiv) at RT for 4 hours. TLC (petrol:EtOAc, 2:1) showed complete conversion to product (R_f 0.1) and disappearance of starting sugar (R_f 0.3) . The reaction mixture was then treated with excess Amberlite 120 (H+) ion exchange resin to remove hydrazine derivatives, washed with brine and extracted into ether, concentrated and the product was by silica gel column chromatography (petrol:EtOAc, 1:1) to yield the desired compound as a clear oil and a 3:1 mixture of anomers (1.91g, 35%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.1; $[\alpha]_D^{24}$ + 45.0 (c = 1 in $CHCl_3$) [Lit. $[\alpha]_D$ + 43 (c = 1 in $CHCl_3$)][40]; ^{1}H NMR (400 MHz CHLOROFORM-*d*) δ ppm 1.99-2.15 (12H, m 4 x OCOC$\underline{H}_3$-α, 4 x OCOC$\underline{H}_3$-β), 3.96 (1 H, td, $J_{4,5}$ = 6.6 Hz, $J_{5,6ab}$ = 1.3 Hz, H-5β), 4.09 (1 H, dd, $J_{6a,6b}$ = 6.6 Hz, $J_{5,6}$ = 2.6 Hz, H-6bα), 4.09 (1 H, dd, $J_{6a,6b}$ = 6.8 Hz, $J_{5,6}$ = 3.2 Hz, H-6aα), 4.13 (1 H, dd, $J_{6a,6b}$ = 6.9, $J_{5,6a}$ = 2.2 Hz, H-6aβ) 4.15 (1 H, dd, $J_{6a,6b}$ = 6.6, $J_{5,6b}$ 1.0 Hz, H-6bβ), 4.47 (1 H, td, $J_{4,5}$ 6.6 Hz, $J_{5,6ab}$ 1.5 Hz, H-5α), 4.70 (1 H, dd, *J* = 6.6, *J* =1.3 Hz, H-1β), 5.05 - 5.09 (1 H, m, H-2β), 5.15 (1 H, dd, $J_{2,3}$ = 10.9 Hz, $J_{1,2}$ = 3.5 Hz, H-2α), 5.42 (3 H, m, H-3α, H-3β, H-4β), 5.47 (1 H, dd, *J* = 3.4, $J_{3,4}$ = 1.4 Hz, H-4α), 5.51 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1α); MS *m/z* (ESI$^+$): 366.2 $[M+NH_4]^+$

3,4,6,7-Tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2,3,4,6-tetra-O-acetyl-α-D-galactopyranoside (145)

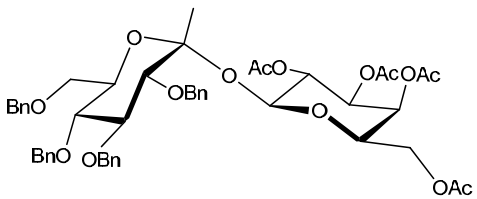


85 (49 mg, 0.088 mmol) and **139** (35.4 mg, 0.1 mmol) were dried *in vacuo* for 1 hour and then dissolved in anhydrous CH_2Cl_2 (8 mL) and added to a dry flask in the presence of molecular sieves (ca. 100 mg). To this was added TMSOTf (ca. 10 µl, 0.06 mmol) at -78 °C in the under an Ar atmosphere. The resulting mixture was stirred for hours until TLC (petrol:EtOAc, 3:1) revealed the

production of one new product (R_f 0.3) spot and complete consumption of starting material (R_f 0.35 and R_f 0.05). The reaction was then quenched by the addition 1 drop NEt_3, filtered through celite and concentrated. The crude product was purified by column chromatography (petrol:EtOAc, 3:1) to give as a colorless oil (59 mg, 76%) as a mixture of α,α and α,β products (1:6.6) that were separated during chromatography to yield the desired compound as exclusively α,α (7.8 mg, 10%) as well as the α,β (51.4 mg, 66%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.3; $[\alpha]_D^{25}$ + 141.1 (c = 0.11, $CHCl_3$) ^{1}H NMR (500 MHz, CHLOROFORM-*d*) δ ppm 1.46 (3 H, CH_3), 1.95 (3 H, s, 1 x $OCOCH_3$), 1.97 (3 H, s, 1 x $OCOCH_3$), 1.99 (3 H, s, 1 x $OCOCH_3$), 2.14 (3 H, s, 1 x $OCOCH_3$), 3.31 (1 H, d, $J_{3,4}$ 9.5 Hz, H-3'), 3.57 - 3.64 (3 H, m, H-5', H-7a', H-7b'), 3.87 (1 H, ddd, $J_{5,6}$ = 10.1 Hz, $J_{6,7a}$ = 4.3 Hz, $J_{6,7b}$ = 2.0 Hz, H-6'), 3.98 (2 H, d, $J_{6a,6b}$ = 6.6 Hz, H-6a, H-6b), 4.03 (1 H, t, *J* = 9.5 Hz, H-4'), 4.48 (1 H, d, *J* = 11.9 Hz, 1 x OCH_2Ph), 4.51 - 4.59 (3 H, m, H-5, 2 x OCH_2Ph), 4.65 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.85 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.90 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.92 (1 H, d, *J* = 10.4 Hz, 1 x OCH_2Ph), 4.94 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 5.29 (1 H, dd, $J_{2,3}$ = 10.7 Hz, $J_{1,2}$ = 3.5 Hz, H-2), 5.41 (2 H, dd, $J_{2,3}$ = 9.8 Hz, $J_{3,4}$ = 3.8 Hz, H-3), 5.42 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1α), 5.48 (1 H, dd, $J_{3,4}$ = 3.8 Hz, $J_{4,5}$ = 1.6 Hz, H-4), 7.05 - 7.37 (20 H, m, Ar-H); ^{13}C NMR (126 MHz, CHLOROFORM-*d*) δ ppm 20.6, 20.6, 20.6 (4 x $OCOCH_3$) 22.7 (CH_3), 61.6 (C-6), 66.2 (C-5), 67.6 (C-2), 68.0 (C-3), 68.1 (C-4), 68.6 (C-7'), 71.9 (C-6') 73.5 (OCH_2Ph), 74.8 (OCH_2Ph), 75.5 (OCH_2Ph), 75.7 (OCH_2Ph), 82.7 (C-5'), 84.5 (C-4'), 89.8 (C-3'), 101.4 (C-2'), 127.5, 127.6, 127.6, 127.7, 127.8, 127.9, 128.2, 128.3, 128.4 (4 x OCH_2Ph), 137.9, 138.2,138.2,138.5 (4 x 1C, 4 x OCH_2Ph), 170.0, 170.0,170.2, 170.2 (4 x C=O); IR (thin film): ν = 3029.97 (C-H), 2926.76, (C=CH) 1750.39 (C=O), 1497.14, 1454.93, 1370.63 (C=C), 1231.10, 1212.80, 1130.00, 1065.57, 947.25, 737.34 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{49}H_{56}O_{15}Na^+$, 907.3963; found, 902.3965.

3,4,6,7-Tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2,3,4,6-tetra-O-acetyl-β-D-galactopyranoside (155)

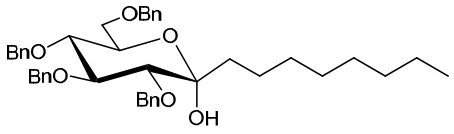


TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.27; $[\alpha]_D^{25}$ + 26.1 (c = 1, $CHCl_3$); ^{1}H NMR (500 MHz, CHLOROFORM-*d*) δ ppm 1.45 (3 H, s, CH_3), 1.70 (3 H, s, 1 x $OCOCH_3$), 1.95 - 2.00 (6 H, m, 2 x $OCOCH_3$), 2.17 (3 H, s, 1 x $OCOCH_3$), 3.34 (1 H, d, $J_{3,4}$ = 9.5 Hz, H-3'), 3.60 (1 H, dd, $J_{7a,7b}$ = 10.7 Hz, $J_{6,7b}$ = 1.9 Hz, H-7b'), 3.72 (1 H, at, *J* = 8.5 Hz, H-5'), 3.74 (1 H, dd, $J_{7a,7b}$ = 10.6, $J_{6,7a}$ = 3.5 Hz,

H-7a'), 3.84 (1 H, at, *J* = 6.3 Hz, H-5), 4.03 (1 H, dd, $J_{6a,6b}$ = 11.3 Hz, $J_{5,6a}$ = 6.3 Hz, H-6a), 4.08 (2 H, dd, $J_{6a,6b}$ = 11.3, $J_{5,6b}$ = 3.4 Hz, H-6b), 4.10 (2 H, dd, $J_{3,4}$ = 9.6 Hz, $J_{4,5}$ = 8.8 Hz, H-4'), 4.20 (1 H, ddd, $J_{5,6}$ = 10.2 Hz, $J_{6,7a}$ = 3.1 Hz, $J_{6,7b}$ = 1.9 Hz, H-6'), 4.49 (1 H, d, *J* = 12.3 Hz, 1 x OCH_2Ph), 4.53 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.60 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.63 (1 H, d, *J* = 12.3 Hz, 1 x OCH_2Ph), 4.83 (1 H, d, *J* = 10.0 Hz, 1 x OCH_2Ph), 4.85 (1 H, d, $J_{1,2}$ = 7.9 Hz, H-1β), 4.89 (2 H, d, *J* = 10.1 Hz, 2 x OCH_2Ph), 4.93 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 5.01 (1 H, dd, $J_{2,3}$ 10.7 Hz, $J_{3,4}$ 3.5 Hz, H-3), 5.22 (1 H, dd, $J_{2,3}$ 10.7 Hz, $J_{1,2}$ 7.9 Hz, H-2), 5.35 (1 H, dd, $J_{3,4}$ = 3.5 Hz, $J_{4,5}$ = 0.9 Hz, H-4), 7.03 - 7.42 (20 H, m, Ar-H); ^{13}C NMR (126 MHz, CHLOROFORM-*d*) δ ppm 20.4, 20.6, 20.6, 20.74 x $OCOCH_3$, 21.9 (CH_3), 61.4 (C-6), 67.0 (C-4), 68.5 (C-7'), 68.5 (C-2), 70.8 (C-3 or C-5), 70.9 (C-3 or C-5), 72.3 (C-6'), 73.4 (OCH_2Ph), 75.1 (OCH_2Ph), 75.5 (OCH_2Ph), 75.6 (OCH_2Ph), 78.2 (C-5'), 82.6 (C-4'), 84.6 (C-3'), 95.1 (C-1), 102.3 (C-2'), 127.5, 127.6, 127.6, 127.7, 127.7, 127.8, 128.0, 128.3, 128.3, 128.4 (4 x OCH_2Ph), 138.2, 138.4, 138.7 (4 x 1C, 4 x OCH_2Ph), 169.2, 170.2, 170.3, 170.4 (4 x C=O); IR (thin film): ν = 3063.02, 3031.00, 28.65.22 (C=CH) 1755.63 (C=O), 1497.04, 1454.27 (C=C), 1220.06, 1124.32, 1089.40, 1061.39, 1038.61, 909.11, 772.54 cm^{-1}; HRMS *m/z* (ESI^+): $[M+NH_4]^+$ calcd. for $C_{49}H_{56}O_{15}NH_4^+$, 907.3963; found, 902.3965.

8.2.6.8 Synthesis of compounds 146 and 147

2,3,4,6-Tetra-O-benzyl-1-C-octyl-α-D-glucopyranose (140)

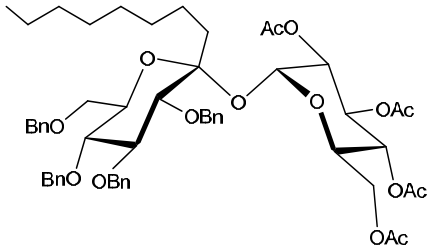


To **95** (116.3mg, 0.2 mmol) in dry THF (10 mL) was added octylmagnesium chloride (1.25 mL, 2M, 0.25mmol in EtO_2) dropwise at -78 ˚C. Reaction was monitored by TLC (petrol:EtOAc, 3:1). Product (R_f 0.5) coruns with starting lactone (R_f 0.5), but lactone is more black in acid stain. Upon completion (30 min) reaction was quenched with aqueous NH_4Cl and washed with saturated brine and $NaHCO_3$. Product was purified by flash chromatography (petrol:EtOAc, 4:1) afforded the desired product (78.2 mg, 60%) as a white amorphous solid.

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.5; $[\alpha]_D^{24}$ + 34.6 (c = 0.8, $CHCl_3$); ^{1}H NMR (500 MHz CHLOROFORM-*d*) δ ppm 1.18 - 1.33 (16 H, m $(CH_2)_7$ CH_3), 1.61 - 1.69 (2 H, m, $\underline{CH_2}(CH_2)_6$ CH_3), 3.44 (1 H, d, $J_{2,3}$ = 9.1 Hz, H-2), 3.63 - 3.70 (2 H, m, H-6a, H-3), 3.77 (1 H, dd, $J_{6a,6b}$ = 11.0 Hz, $J_{5,6b}$ = 3.8 Hz, H-6b), 3.96 - 4.04 (2 H, m, H-5, H-4), 4.56 (1 H, d, *J* = 12.3 Hz, 1 x OCH_2Ph), 4.62 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.64 (1 H, d, *J* = 12.3 Hz, 1 x OCH_2Ph), 4.69 (1 H, d, *J* = 11.0 Hz, 1 x

OCH_2Ph), 4.84 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.88 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.93 (2 H, d, *J* = 11.1 Hz, 1 x OCH_2Ph), 4.94 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 7.17 - 7.38 (20 H, m, Ar-H); ^{13}C NMR (126 MHz CHLOROFORM-*d*) δ ppm 22.5, 22.6, 29.2, 29.2, 29.5, 29.7, 31.8, 38.6 (8 C, $(CH_2)_7CH_3$)), 68.8 (C-6), 71.6 (C-5), 73.3 (OCH_2Ph), 74.9 (OCH_2Ph), 75.4 (OCH_2Ph), 75.5 (OCH_2Ph), 78.4 (C-3), 81.4 (C-2), 83.9 (C-4), 98.4 (C-2), 127.4 (1 C, s), 127.5, 127.6, 127.7, 127.8, 127.8, 127.9, 128.0, 128.3, 128.3, 128.4, 138.28, 138.47, 138.6 (OCH_2Ph); IR (thin film): ν = 2926.66, 2855.40, (C=CH) 1496.73, 1453.80, 1366.91 (C=C), 1220.46, 1039.82, 978.19, 772.61 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{42}H_{52}O_6Na^+$, 675.3662, found, 675.3647.

2,3,4,6-Tetra-O-benzyl-1-C-octyl-α-D-glucopyranosyl-2,3,4,6-tetra-O-acetyl-α-D-glucopyranoside (140)

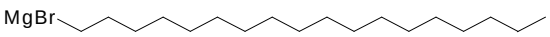


140 (36.5 mg, 0.055 mmol) and **84** (19.0 mg, 0.055 mmol) were dried en vacuo for 1 hour, then dissolved in anhydrous CH_2Cl_2 (8 mL) and added to a flame dried RB in the presence of molecular sieves (ca. 100 mg). To this was added TMSOTf (ca. 10 μl, 0.06 mmol) at -40 °C in the under an Ar atmosphere. The resulting mixture was stirred for 30 min until TLC in (petrol:EtOAc, 2:1) revealed the production of one new product spot (R_f 0.6) and complete consumption of starting material (R_f 0.7). The reaction was then quenched by the addition 1 mL NEt_3, filtered through celite and concentrated to produce. The crude product was purified by column chromatography (petrol:EtOAc, 5:2) to give as a colorless oil (246.0 mg, 88%) as a mixture of α,α and α,β products (6:1) that were separated during chromatography to yield the desired product (39.5 mg, 73%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.6; $[\alpha]_D^{24}$ + 80.5 (c = 0.1, $CHCl_3$); ^1H NMR (500 MHz CHLOROFORM-*d*) δ ppm 1.22 - 1.29 (10 H, m, $(CH_2)_7$ CH_3), 1.58 (5 H, br. s. $(CH_2)_7$ CH_3), 1.83 (2 H, t, $J_{CH2-CH2}$ 8.2 Hz, $(CH_2)_7$ CH_3), 1.98 (3 H, s, 1 x $OCOCH_3$), 2.01 (3 H, s, 1 x $OCOCH_3$), 2.03 (3 H, s, 1 x $OCOCH_3$), 2.05 (3 H, s, 1 x $OCOCH_3$), 3.57 - 3.60 (3 H, m, H-6a', H-2', H-4'), 3.63 (1 H, dd, $J_{6a,6b}$ = 11.0 Hz, $J_{5,6b}$ = 4.4 Hz, H-6b'), 3.83 (1 H, dd, $J_{6a,6b}$ = 12.5 Hz, $J_{5,6a}$ = 2.4 Hz, H-6a), 3.90 (1 H, ddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6b}$ = 4.4 Hz, $J_{5,6a}$ = 1.9 Hz, H-5'), 4.10 (1 H, dd, $J_{6a,6b}$ = 12.4 Hz, $J_{5,6b}$ = 4.8 Hz, H-6b), 4.12 - 4.16 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 7.25 Hz, H-3'), 4.36 (1 H, ddd, $J_{4,5}$ = 10.2 Hz, $J_{5,6b}$ = 4.7 Hz, $J_{5,6b}$ = 2.4 Hz, H-5), 4.47 - 4.52 (1 H, d, *J* = 12.3 Hz, 1 x OCH_2Ph), 4.53 - 4.58 (1 H, d, *J* = 12.3 Hz, 1 x OCH_2Ph), 4.61 (2 H, d, *J* = 11.0 Hz, 2 x OCH_2Ph), 4.86 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.93 (1

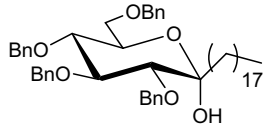
H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.96 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.99 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 5.03 - 5.08 (2 H, m, H-4, H-2), 5.39 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1), 5.58 (1 H, at, $J_{2,3}$= $J_{3,4}$= 9.8 = Hz, H-3); ^{13}C NMR (126 MHz CHLOROFORM-*d*) δ ppm 20.6, 20.7, 22.6, 24.4, 29.2, 29.57, 31.8, 34.2, (8 C, $(CH_2)_7$ CH_3) 61.9 (C-6), 67.1 (C-5), 68.6 (C-6'), 68.7 (C-4), 70.4 (C-2), 70.7 (C-3), 72.2 (C-5'), 73.3 (OCH_2Ph), 74.9 (OCH_2Ph), 74.9 (OCH_2Ph), 75.5 (OCH_2Ph), 78.4 (C-4' or C-2'), 79.6 (C-4' or C-2'), 82.9 (C-3'), 88.7 (C-1), 103.6 (C-1'), 127.3, 127.5, 127.5, 127.6, 127.6, 127.7, 127.9, 128.26, 128.3, 128.4, 138.3, 138.3, 138.6, 138.6 (OCH_2Ph), 169.7, 169.8, 170.2, 170.6, (4 x C=O); IR (thin film): ν= 2925.73, 2854.58 (C=CH), 1753.37 (C=O), 1452.93, 1367.07 (C=C), 1219.86, 1039.87, 978.20, 772.70 cm^{-1} HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{56}H_{70}O_{15}Na^+$, 1005.4612; found, 1005.4607.

Octadecylmagnesium bromide (157)



Magnesium turnings (2.04g, 840mmol) were added to a flamed dried RB and to this was added 1 iodine crystal and attached to a spiral condenser. 1-bromo-octadecane (20.11 g, 60 mmol) was dissolved in 30mL dry THF to a concentration of 2 M. 5 mL of the lipid solution was added via syringe to the magnesium turnings dropwise, very slowly and the solution was treated with heat gun until it began to reflux. After solution had begun to reflux of its own accord, the remaining lipid solution was added dropwise. Upon completion of reflux, solution was stored under argon and titrated against menthol (157.8 mg) with 1,10-phenanthroline as indicator. Grignard had a concentration of 0.77 M.

2,3,4,6-Tetra-O-benzyl-1-C-octadecyl-α-D-glucopyranose (141)

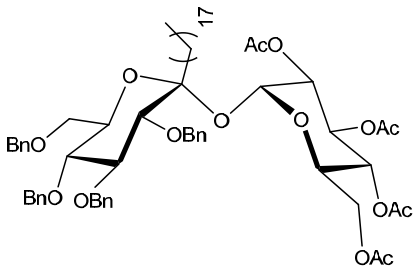


95 (360 mg, 0.66 mmol) in dry THF (10 mL) was added **157** (3 mL, 0.77 M, 2.31 mmol in THF) at -78 ˚C. After 3 hours and addition of more **157**, reaction did not proceed to completion. TLC (petrol:EtOAc, 2:1) revealed both product (R_f 0.8) and starting material (R_f 0.6). Reaction was quenched with aqueous NH_4Cl and washed with saturated brine and $NaHCO_3$. Product was purified by flash chromatography (petrol:EtOAc, 5:1) afforded the desired product as a clear oil (212 mg, 39%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.8; $[\alpha]_D^{24}$ + 10.2 (c = 1, $CHCl_3$); ^1H NMR (500 MHz CHLOROFORM-*d*) δ ppm 1.19 - 1.31 (31 H, m, $(CH_2)_{17}$ CH_3), 1.57 (4H, m, $(CH_2)_{17}$ CH_3), 1.62 - 1.67

(2 H, m, $(CH_2)_{17}CH_3$), 3.44 (1 H, d, $J_{2,3}$ = 9.1 Hz, H-2), 3.65 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4), 3.67 (1 H, dd, $J_{6a,6b}$ =11.0 Hz, $J_{5,6b}$ = 2.2 Hz, H-6b) 3.77 (1 H, dd, $J_{6a,6b}$ = 11.0 Hz, $J_{5,6a}$ = 3.8 Hz, H-6a), 3.99 (1 H, ddd, $J_{4,5}$ = 10.3 Hz, $J_{5,6b}$ = 3.8 Hz, $J_{5,6a}$ = 2.0 Hz, H-5), 4.01 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.1 Hz, H-3), 4.56 (1 H, d, *J* = 11.9 Hz, 1 x OCH_2Ph), 4.62 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.64 (1 H, d, *J* = 11.9 Hz, 1 x OCH_2Ph), 4.69 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.84 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.88 (1 H, d, *J* = 11.2 Hz, 1 x OCH_2Ph), 4.93 (1 H, d, *J* = 11.1 Hz, 1 x OCH_2Ph), 4.93 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 7.21 (2 H, dd, Ar-H), 7.25 - 7.37 (18 H, m, Ar-H); ^{13}C NMR (126 MHz CHLOROFORM-*d*) δ ppm 14.1, 22.5, 22.6, 29.3, 29.5, 29.6, 29.6 31.9, 38.6 (18 C $(CH_2)_{17}CH_3$), 68.8 (C-6), 71.6 (C-5), 73.3 (OCH_2Ph), 74.8 (OCH_2Ph), 75.4 (OCH_2Ph), 75.6 (OCH_2Ph), 78.4 (C4), 81.4 (C-2), 83.8 (C-3), 98.9 (C-1), 127.4, 127.6, 127.6, 127.7, 127.8, 127.9, 128.2, 128.3, 128.4, 137.9, 138.2, 138.4, 138.6 (OCH_2Ph); IR (thin film): ν = 3063.93, 3031.80, 2850.13, 1604.73, 1496.44, 1453.71, 1361.73, 1271.74, 1089.23, 772.54 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{52}H_{72}O_6Na^+$, 815.5227; found, 815.5221.

2,3,4,6-Tetra-O-benzyl-1-C-octadecyl-α-D-glucopyranosyl-2,3,4,6-tetra-O-acetyl-α-D-glucopyranoside (147)



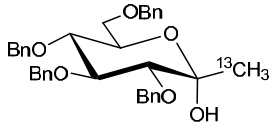
141 (150 mg, 0.19 mmol) and **84** (99.6 mg, 0.285 mmol) were dried en vacuo for 1 hour and then dissolved in anhydrous CH_2Cl_2 (12 mL) and added to a dry flask in the presence of molecular sieves (ca. 100 mg). To this was added TMSOTf (ca. 5 µl, 0.027 mmol) at -40 °C in the under an Ar atmosphere. The resulting mixture was stirred for 30min until TLC (petrol:EtOAc, 4:1) revealed the production of one new product spot (R_f 0.2) and complete consumption of starting material (R_f 0.5). The reaction was then quenched by the addition 1mL Net_3, filtered through celite and concentrated to produce. The crude product was purified by column chromatography (petrol:EtOAc, 3:1) to give the desired product as a colorless oil (132 mg, 61%) as a clear oil. **158** was not isolated to purity.

TLC (petrol:EtOAc, 4:1, v/v): R_f = 0.2; $[\alpha]_D^{25}$ + 34.7 (c = 1, $CHCl_3$) ^{1}H NMR (500 MHz CHLOROFORM-*d*) δ ppm 1.19 - 1.38 (31 H, m, $CH_2C_{16}H_{33}$), 1.51 - 1.62 (4 H, m, $(CH_2)_{17}CH_3$), 1.83 (1 H, t, $J_{CH2-CH2}$ = 8.2 Hz, $(CH_2)_{17}CH_3$), 1.98 (3 H, s, 1 x $OCOCH_3$), 2.01 (3 H, s, 1 x $OCOCH_3$), 2.03 (3 H, s, 1 x $OCOCH_3$), 2.05 (3 H, s, 1 x $OCOCH_3$), 3.53 - 3.65 (4 H, m, H-2', H-4', H-6a', H-6b'),

3.83 (1 H, dd, $J_{6a,6b}$ = 12.5 Hz, $J_{5,6a}$ = 2.4 Hz, H-6_a), 3.90 (1 H, ddd, $J_{4,5}$ = 10.0 Hz, $J_{5,6a}$ = 4.3 Hz, $J_{5,6b}$ = 2.0 Hz, H-5'), 4.10 (1 H, dd, $J_{6a,6b}$ = 12.8 Hz, $J_{5,6b}$ = 4.9 Hz, H-6_b), 4.14 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.5 Hz, H-3'), 4.36 (1 H, ddd, $J_{4,5}$ = 10.3 Hz, $J_{5,6b}$ = 4.7 Hz, $J_{5,6a}$ = 2.4 Hz, H-5), 4.49 (1 H, d, J = 11.98 Hz, 1 x OCH_2Ph), 4.56 (1 H, d, J = 12.3 Hz, 1 x OCH_2Ph), 4.61 (2 H, d, J = 10.7 Hz, 2 x OCH_2Ph), 4.86 (1 H, d, J = 11.3 Hz, 1 x OCH_2Ph), 4.94 (1 H, d, J = 11.3 Hz, 1 x OCH_2Ph), 4.96 (1 H, d, J = 11.0 Hz, 1 x OCH_2Ph) 4.99 (1 H, d, J = 11.3 Hz, 1 x OCH_2Ph), 5.05 (2 H, m, H-2, H-4), 5.38 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1), 5.58 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.6 Hz, H-3); ^{13}C NMR (126 MHz CHLOROFORM-*d*) δ ppm 14.1, 20.6, 20.7, 20.7, 22.6, 24.4, 25.7, 29.3, 29.4, 29.6, 29.6, 29.7, 30.0, 31.9, 32.8, 34.2, (18 C $(CH_2)_{17}CH_3$, 61.9 (C-6), 67.1 (C-5), 68.6 (C-4), 68.6 (C-6'), 70.4 (C-2), 70.7 (C-3), 72.2 (C-5'), 73.3 (OCH_2Ph), 74.9 (OCH_2Ph), 74.9 (OCH_2Ph), 75.5 (OCH_2Ph), 78.4 (C-4'), 79.6 (C-2'), 82.9 (C-3'), 88.7 (C-1), 103.6 (C-1'), 127.3, 127.5, 127.5, 127.6, 127.6, 127.7, 127.9, 128.2, 128.3, 128.4, 128.4, 138.3, 138.3, 138.6, 138.6 (OCH_2Ph), 169.7, 169.8, 170.2, 170.6 (4 x C=O); IR (thin film): ν = 3028.42, 2955.0, 1750.66 (C=O), 1496.74, 1472.20, 1462.39, 1366.28, 1211.18, 1122.77, 1038.69, 771.31 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{66}H_{90}O_{15}Na^+$, 1145.62; found 1145.5916. Isotopic distribution: Species observed $[M+Na]^+$, peaks observed 1145.5916 (100%), 2246.5955 (74%), 1147.5984 (28.5%), 1148.6009 (8.1%), 1149.6042 (1.9%); peaks predicted 1145.6172 (100%), 1146.6206 (73.0%), 1147.6237 (29.3%), 1148.6267 (8.2%), 1149.6295 (1.8%).

8.2.7 Isotopic labeling of trehalose: Synthesis of compounds 148 and 149

3,4,6,7-tetra-O-benzyl-1-$^{13}CH_3$-deoxy-α-D-gluco-hept-2-ulopyranosyl (86)

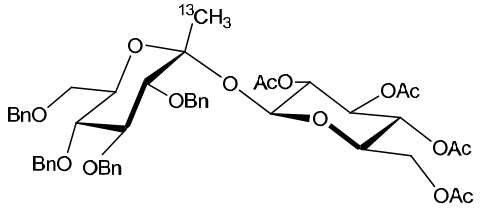


Under schlenk conditions, in anhydrous THF at -78 ˚C, ^{13}C-MeI **162** (75 µl, 1.2 mmol) was added to a solution of (0.9 mL, 1.53 mmol) t-buli in 2mL THF. The reaction turns white from the yellow color characteristic of *t*-BuLi dissolved in THF. To this was added **95** (384.6 mg, 0.73 mmol) in 5 mL anhydrous THF. Sugar was added via syringe dropwise over a period of 20 minutes. Reaction was stirred at -78 ˚C for an additional 20 minutes, upon which time product was detected (R_f 0.4) by TLC (petrol:EtOAc, 2:1). Column chromatography afforded the desired product (151.2 mg, 40%) as a white, crystalline solid as well as recovered lactone (57.7 mg, 15%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.4; M.p. = 80-82 °C; $[\alpha]_D^{25}$ + 37.3 (c = 1.66, $CHCl_3$) ;^{1}H NMR (500 MHz, CHLOROFORM-*d*) δ ppm 1.30 (1 H, s, 1 x ***^{13}C<u>H</u>$_3$***), 1.56 (1 H, s, 1 x ***^{13}C<u>H</u>$_3$***), 2.06 (1 H, s, 1 x ***^{13}C<u>H</u>$_3$***), 3.39 (1 H, d, $J_{3,4}$ = 9.8 Hz, H-3), 3.67 (1 H, at, $J_{4,5}$ = $J_{5,6}$ = 9.7 Hz, H-5), 3.69 (1 H, dd,

$J_{7a,7b}$ = 12.0 Hz, $J_{6,7b}$ = 2.2 Hz, H-7b), 3.75 (1 H, dd, $J_{7a,7b}$ = 11.7 Hz, $J_{6,7a}$ = 4.4 Hz, H-7a), 3.99 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.9 Hz, H-4), 4.02 (1 H, ddd, $J_{5,6}$ = 9.9 Hz, $J_{6,7a}$ = 4.4 Hz, $J_{6,7b}$ = 2.2 Hz, H-6), 4.55 (1 H, d, *J* = 12.3 Hz, 1 x OCH_2Ph), 4.58 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.64 (1 H, d, *J* = 12.3 Hz, 1 x OCH_2Ph), 4.72 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.84 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.90 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.93 (1 H, d, *J* = 11.9 Hz, 1 x OCH_2Ph), 4.95 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 6.82 - 7.64 (20 H, m, Ar-H); ^{13}C NMR (126 MHz, CHLOROFORM-*d*) δ ppm 26.2-26.7 (1 C, m, $^{13}CH_3$), 68.8 (C-7), 71.5 (C-5), 73.4 (OCH_2Ph), 74.9 (OCH_2Ph), 75.6 (OCH_2Ph), 75.7 (OCH_2Ph), 78.3 (C-5) 83.2 (C-3), 83.6 (C-4), 97.4 (1 C, d, *J* 47.6 Hz, C-2), 127.6, 127.7, 127.8, 127.8, 127.9, 127.9, 128.3, 128.3, 128.4 (1 C, s), 128.4 (4 x OCH_2Ph), 137.9, 138.2, 138.2, 138.6 (4 x 1C, 4 x OCH_2Ph.; IR (thin film): ν = 3029.65, 2864.62 (C=CH) 1496.66 1454.07, 1362.18 (C=C), 1072.93, 1027.58, 870.48, 735.36 cm^{-1}. HRMS *m/z* (ESI^+): $[M+Li]^+$ calcd. for $C_{34}{}^{13}CH_{38}O_6Li^+$, 562.2862; found 562.2858.

3,4,6,7-Tetra-O-benzyl-1-$^{13}CH_3$-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2,3,4,6-tetra-O-acetyl-β-D-glucopyranoside (164)

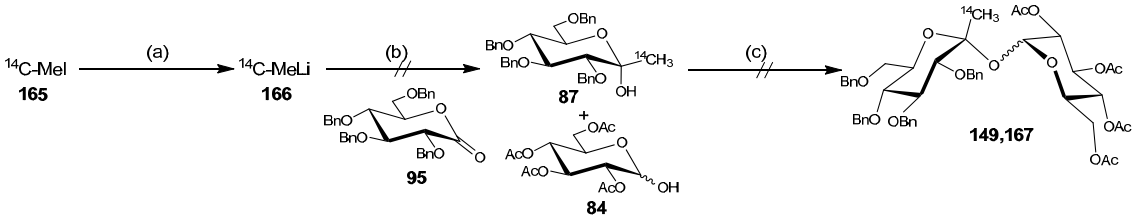


86 (39.7 mg, 0.071 mmol) and **84** (28.1 mg, 0.419 mmol) were dried *in vacuo* for 1 hour and then dissolved in anhydrous CH_2Cl_2 (8 mL) and added to a dry flask in the presence of molecular sieves (ca. 100 mg). To this was added TMSOTf (ca. 5µl, 0.027mmol) at -78 °C in the under an Ar atmosphere. The resulting mixture was stirred for 2 hours until TLC (petrol:EtOAc, 2:1) revealed the production of one new product (R_f 0.3) spot and complete consumption of starting material (R_f 0.35 and R_f 0.05). The reaction was then quenched by the addition 1 drop NEt_3, filtered through celite and concentrated. Column chromatography (petrol:EtOAc, 5:2) yielded the product (61.3mg, 92%) as a 25:1 ratio of α,β: α,α as detected by NMR. **148** was not isolated to purity.

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.3; $[\alpha]_D^{25}$ + 28.5 (c, 0.46 in $CHCl_3$); ^{1}H NMR (500 MHz, CHLOROFORM-*d*) δ ppm 1.60 (1 H, s, 1 x ***$^{13}C\underline{H}_3$***), 1.71 (2 H, s, 2 x ***$^{13}C\underline{H}_3$***), 2.00 (6 H, s, 2 x $OCOCH_3$), 2.02 (3 H, s, 1 x $OCOCH_3$), 2.05 (3 H, s, 1 x $OCOCH_3$), 3.33 (1 H, add, $J_{3,4}$ = 9.6, $J_{13C,3}$ = 2.0 Hz, H-3'), 3.60 (1 H, dd, $J_{7a,7b}$ = 11.0 Hz, $J_{6,7b}$ = 1.6 Hz, H-7b'), 3.62 (1 H, ddd, $J_{4,5}$ = 9.7 Hz, $J_{5,6a}$ = 4.7 Hz, $J_{5,6b}$ = 2.2 Hz, H-5), 3.69 (1 H, at, *J* = 10.9 Hz, H-5'), 3.71 (1 H, dd, $J_{7a,7b}$ = 11.0 Hz, $J_{6,7a}$ = 5.8 Hz, H-7a'), 4.05 (1 H, dd, $J_{6a,6b}$ = 12.1 Hz, $J_{5,6b}$ = 2.4 Hz, H-6b), 4.05 (1 H, at, *J* = 9.5 Hz, H-4'),

4.11 (1 H, dd, $J_{6a,6b}$ = 12.1 Hz, $J_{5,6a}$ = 5.4 Hz, H-6a), 4.17 (1 H, ddd, $J_{5,6}$ = 10.3 Hz, $J_{6,7a}$ = 3.4 Hz, $J_{6,7b}$ = 1.7 Hz, H-6'), 4.48 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.53 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.60 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.59 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.83 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.86 (1 H, d, *J* = 11.9 Hz, 1 x OCH_2Ph), 4.89 (1 H, d, *J* = 11.9 Hz, 1 x OCH_2Ph), 4.89 (1 H, d, $J_{1,2}$ = 8.2 Hz, H-1β), 4.93 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 5.03 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 10.1 Hz, H-4), 5.03 (1 H, dd, $J_{2,3}$ = 9.8 Hz, $J_{1,2}$ = 7.3 Hz, H-2), 5.20 (1 H, at, *J* = 9.5, H-3), 6.95 - 7.42 (20 H, m, Ar-H);^{13}C NMR (126 MHz, CHLOROFORM-*d*) δ ppm 22.1 ($^{13}CH_3$), 22.3 (1 C, s), 62.1 (C-6), 68.4 (1 C, s), 68.5 (C-4), 70.9 (C-7'), 72.1 (C-5), 72.4 (C-6'), 72.9 (C-3), 73.4 (OCH_2Ph), 74.9 (OCH_2Ph), 75.5 (OCH_2Ph), 75.5 (OCH_2Ph), 78.2 (H-5'), 82.5 (H-4'), 84.5 (H-3'), 94.6 (C-1), 102.4 (1 C, d, *J* = 47.2 Hz, C-2'), 127.5, 127.6, 127.6, 127.6, 127.8, 127.9, 128.3, 128.3, 128.4 (4 x OCH_2Ph), 138.2, 138.2, 138.4, 138.6 (4 x 1C, 4 x OCH_2Ph), 169.1, 169.4,170.4, 170.6 (4 x C=O); IR (thin film): ν = 3062.99, 3030.24, 2923.54, 2853.98 (C=CH) 1755.76 (C=O), 1497.22, 1454.70 (C=C), 1265.56, 1226.64, 1090.08, 1063.34, 1039.17, 909.33, 737.34 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{48}{}^{13}CH_{56}O_{15}Na^+$, 908.3550; found, 908.3533.

8.2.7.1 Attempted synthesis of ^{14}C-Trehalose 149



Scheme 8.1. Unsuccessful attempted synthesis of ^{14}C ketoside based trehalose analogues. (a) t-BuLi 2.2 eq, THF, -78 ˚C, 15 min, (b) 95, THF, -78 ˚C (c) TMSOTf, DCM, -40 ˚C.

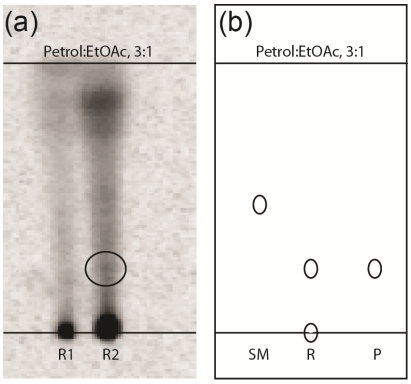
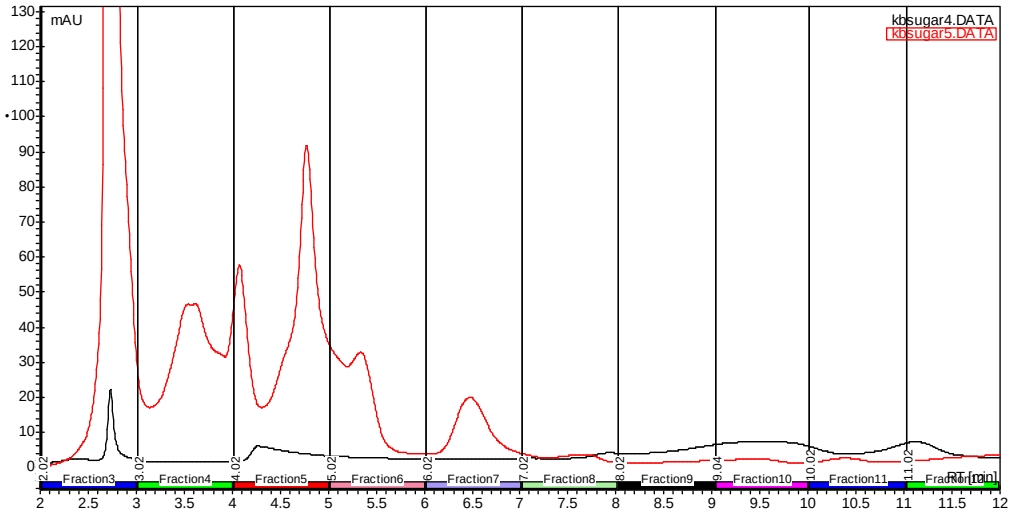


Figure 8.4. (a) Radiographic images of the TLC plates of both reactions (Scheme 8.1) conducted (R1 and R2). Nearly all radioactivity is concentrated in the baseline spot. The circle indicates where the product should have been. Little radioactivity is detected in that region (b) Example of the expected TLC of the radio reaction, based on TLCs from non-reactive test reactions.



C

Fraction	3	4	5	6	7	8	9	10	11
DPM	7263	1081	763	1015	612	235	182	144	580
Mass				607.2	515.4				650.3
				515.4	537.5				556.2
				477.3					531.2
				567.3					515.4

Figure 8.5. (a) HPLC trace of the methyl lithium reactions shows no expected product. (b) inset shows expansion of fractions where the product was expected. (c) Scintillation counting and detected masses are indicated for fractions of interest. Product was expected in fraction 11 ($M+NH_4$ 579.2), but no correct mass was detected. DPM-disintegrations per minute.

α,α **α,β**

Figure 8.6. HPLC trace of optimized separation of an anomeric mixture of unlabeled ketoside trehalose, utilizing Hexanes:EtOAc, 3:1 as mobile phase for normal phase chromatography on an Agilent silica column. Compounds on the chromatograph indicate where the α,α and α,β product elute, showing good separation.

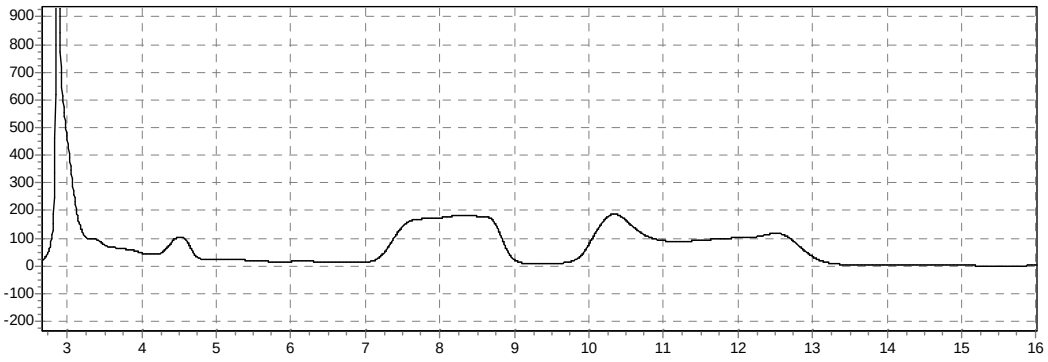
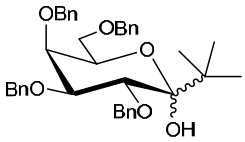


Figure 8.7. HPLC trace of reaction depicted in scheme 2.41 shows incomplete separation of the glycosylation of 85 with ^{14}C-tetra-acetyl-galactose 168. Separation was not obtained due to water and acetone on the column. The broad peak between fractions 10 and 12 is a mixture of starting material and product.

Optimized lithium halogen exchange reaction conditions.

Under schlenk conditions, In THF at -78 °C, MeI (3 µl or 0.05 mmol) was added to a solution of 0.17mmol t-buli in 100 µl THF, 3.4 equivalents of *t*-BuLi relative the the MeI. The reaction turns white from the yellow color characteristic of *t*-BuLi dissolved in THF. To this was added galactonolactone **92** (0.067 mmol, 1.1 equi) TLC (petrol:EtOAc, 3:1) relealed disappearance of starting material (R_f 0.4) and appearance of either product (R_f 0.2) or side product **172** (R_f 0 .3) **171** (R_f 0.5) or **173** (R_f 0.0).

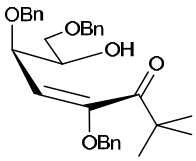
2,3,4,6-Tetra-O-benzyl-1-C-1-tert-butyl-α-D-galactopyranose (172)



Product **172** was obtained as a mixture of stereoisomers, with one major and one minor set of peaks in the NMR. Only the major peaks were assigned, but without NOE analysis identification of the particular stereoisomer was not possible. By TLC (petrol:EtOAc, 3:1) product was detected at (R_f 0.3). Purification by column chromatography (petrol:EtOAc, 3:1) yielded the title compound (7.0 mg, 15%) as a clear oil.

TLC (petrol:EtOAc, 3:1, v/v): R_f = 0.3; $[\alpha]_D^{25}$ +14.34 (c = 0.76, $CHCl_3$); ^{1}H NMR (CHLOROFORM-*d*) δ ppm 3.55 (1 H, dd, $J_{6a,6b}$ = 8.8 Hz, $J_{5,6a}$ = 5.4 Hz, H-6a), 3.65 (1 H, dd, $J_{6a,6b}$ = 9.8 Hz, $J_{5,6b}$ = 2.8 Hz, H-6b), 3.92 (1 H, dd, $J_{2,3}$ = 9.0 Hz, $J_{3,4}$ = 2.7 Hz, H-3), 4.04 (1 H, dd, $J_{3,4}$ = 2.7 Hz, $J_{4,5}$ = 0.8 Hz, H-4), 4.06 - 4.10 (1 H, m, H-5), 4.15 (1 H, d, $J_{2,3}$ = 9.14 Hz H-2), 4.46 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.47 - 4.50 (1 H, m, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.60 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.61 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.67 (1 H, d, *J* = 7.6 Hz, 1 x OCH_2Ph), 4.69 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.76 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 5.12 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph); IR (thin film): υ = br 3030.42 (OH), 2924.82, 1496.77, 1454.12 (Ar C=C), 1364.30, 10.93.92, 734.11 cm^{-1}; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{38}H_{44}O_6Na^+$, 511.2460; found, 511.2445.

4,6,8-Tris-benzyloxy-7-hydroxy-2,2-dimethyl-oct-4-en-3-one (171)

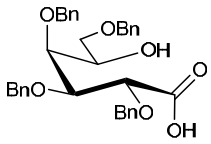


During the reaction between methyllithium and **92** TLC (petrol:EtOAc, 3:1) revealed a new spot (R_f 0.5) that was not the desired product **93** (R_f 0.2). This product was isolated by column chromatography (petrol:EtOAc, 3:1) and was found to be the titled compound (7.0 mg, 18%) as a clear oil.

TLC (petrol:EtOAc, 3:1, v/v): R_f = 0.2; $[\alpha]_D^{25}$ -57.7 (c = 0.7, $CHCl_3$); ^{1}H NMR (500 MHz CHLOROFORM-*d*) δ ppm 1.23 (9 H, s, 3 x CH_3), 2.63 (1 H, d, *J* = 4.7 Hz, OH), 3.47-3.50 (2 H, m, H-6a, H-6b), 3.68 - 3.76 (1 H, quint, *J* = 5.1 Hz, H-5), 4.24 (2 H, d, *J* = 11.7 Hz, 1 x OCH_2Ph), 4.45 (3 H, dd, $J_{3,4}$ = 9.3 Hz, $J_{4,5}$ = 5.8 Hz, H-4), 4.48 (2 H, d, *J* = 11.0 Hz, 2 x OCH_2Ph), 4.53 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.69 (1 H, d, *J* = 11.5 Hz, 1 x OCH_2Ph), 4.71 (1 H, d, *J* = 11.3 Hz, 1 x

OCH_2Ph), 5.53 (1 H, d, $J_{3,4}$ = 9.1 Hz, H-3), 7.01 - 7.47 (20 H, m, ArH); ^{13}C NMR (500 MHz CHLOROFORM-*d*) δ ppm 27.4 (3 x CH_3), 70.7 (C-6), 70.8 (OCH_2Ph), 72.9 (C-5), 73.2 (C-4), 73.4 (OCH_2Ph), 73.6 (OCH_2Ph), 118.2, 127.6, 127.7, 127.75, 127.9, 128.1, 128.3, 128.4, 128.5, 136.2, 137.9, 137.9 (CH_2Ph), 154.3 (C-3), 206.1 (C=O); HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{38}H_{44}O_6Na^+$, 511.2460; found, 511.2445.

***2,3,4,6-Tetrakis-benzyloxy-5-hydroxy-hexanoic acid* (71)**[41]

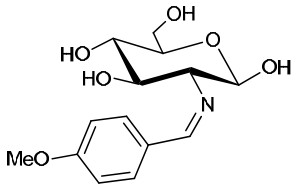


During the reaction between methyllithium and **92** TLC (petrol:EtOAc, 3:1) revealed a new spot (R_f 0) that was not the desired product (R_f 0.2). This product was isolated by column chromatography (petrol:EtOAc, 3:1) and was found to be the titled compound (13.5mg, 30.3%) as a clear oil.

TLC (petrol:EtOAc, 3:1, v/v): R_f = 0.2; $[\alpha]_D^{25}$ + 5.1 (c = 1, $CHCl_3$); ^1H NMR (500MHz MeOD) δ ppm 3.52 (2 H, d, *J* = 6.6 Hz, H-6a, H-6b), 3.83 (1 H, d, $J_{2,3}$ = 8.5 Hz, H-2), 4.13 (3 H, m, H-4, H-5), 4.23 - 4.35 (3 H, m, H-3, 2 x OCH_2Ph), 4.42 (1 H, d, *J* = 11.7 Hz, 1 x OCH_2Ph), 4.51 (2 H, d, *J* = 12.0 Hz, 2 x OCH_2Ph), 4.66 (2 H, d, *J* = 10.4 Hz, 2 x OCH_2Ph), 4.72 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 7.01 - 7.55 (20 H, m); MS *m/z* (ESI^+): 579.2 $[M+Na]^+$.

8.2.8 Synthesis of amino-trehalose 37

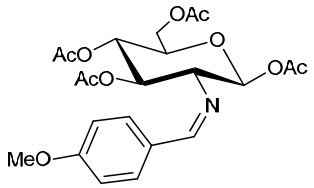
2-Deoxy-2-[p-methoxybenzilidine(amino)]-β-D-glucopyranoside (250)[42]



D-glucosamine **175** (51.5 g, 0.23 mol) was dissolved in a freshly prepared 1 M NaOH solution (300 mL) prior to the addition of *p*-anisaldehyde (34 mL, 0.28 mol) and the reaction mixture stirred vigorously for 10 mins until precipitation began. The reaction mixture was then cooled to 4 ^oC for 1 hour, following which, the precipitate was filtered and washed with cold water (150 mL) prior to a second wash (EtOH:Et_2O, 1:1, 100 mL) and the dried under reduced pressure. This afforded the desired product as a white crystalline solid (44.3 g, 65%).

TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.05; M.p.= 159-160 °C. (Lit. m.p. = 166 °C)[43]; $[\alpha]_D^{25}$ = -13.2 (c = 0.5, pyridine). [Lit. $[\alpha]_D^{21}$ = -33 (c = 1, pyridine)][42]; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 2.79 (1 H, at, $J_{1,2}$ = $J_{2,3}$ = 8.2 Hz, H-2), 3.14 (1 H, atd, J = 9.1 Hz, J = 5.2 Hz, H-4), 3.23 (1 H, ddd, $J_{4,5}$ = 9.6 Hz, $J_{5,6'}$ = 5.8 Hz, $J_{5,6}$ = 2.0 Hz, H-5), 3.42 (1 H, atd, J = 9.1 Hz, J = 5.6 Hz, H-3), 3.48 (1H, ddd, $J_{6a,6b}$ = 11.7 Hz, $J_{6a,OH-6}$ = 5.6 Hz, $J_{5,6a}$ = 5.2 Hz, H-6b). 3.72 (1 H, ddd, $J_{6a,6b}$ = 11.6 Hz, $J_{6b,OH-6}$ = 5.6 Hz, $J_{5,6b}$ = 1.9 Hz, H-6b), 3.80 (3H, s, OCH_3), 4.58 (1 H, at, J = 5.6 Hz, OH-6), 4.69 (1 H, at, J = 7.3 Hz, H-1), 4.85 (1 H, d, $J_{3,OH-3}$ = 5.6 Hz, OH-3), 4.97 (1 H, d, $J_{4,OH-4}$ = 5.8 Hz, OH-4), 6.54 (1 H, d, $J_{1,OH-1}$ = 6.8 Hz, OH-1), 6.98 (2 H, d, J = 10.4 Hz, Ar-H), 7.68 (2 H, d, J = 10.4 Hz, Ar-H), 8.11 (1 H, s, benzylic-H), 9.75 - 9.99 (1 H, s, NH); MS *m/z* (ESI^+): 298.2 $[M+H]^+$.

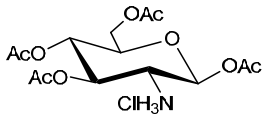
1,3,4,6-Tetra-O-acetyl-2-deoxy-2-[p-methoxybenzilidine(amino)]-β-D-glucopyranoside (176)[43]



250 (44.3 g, 0.149 mol) was dissolved in pyridine (170 mL) under an argon atmosphere The reaction mixture was cooled to 0°C prior to the addition of Ac_2O (150 mL) and stirred for 24 hours. After 24 hours TLC (petrol:EtOAc, 3:2) indicated the formation of a product (R_f 0.3) and consumption of starting material (R_f 0.0). The reaction solution was then added to ice water (1 L) and the precipitate formed was filtered and washed with cold water (200 mL) and dried under reduced pressure. This afforded the desired product (67.4 g, 97%) as a white crystalline solid.

TLC (petrol:EtOAc, 3:2, v/v): R_f = 0.3; M.p.= 175-176 °C. (Lit. m.p. = 188 °C)[43]; $[\alpha]_D^{25}$ = +71.7 (c = 1, $CHCl_3$). [Lit. $[\alpha]_D^{21}$ = +95 (c = 1, $CHCl_3$)][43]; ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.86 (3 H, s, 1 x $OCOCH_3$), 2.00 (3 H, s, 1 x $OCOCH_3$), 2.02 (3 H, s, 1 x $OCOCH_3$), 2.08 (3 H, s, 1 x $OCOCH_3$), 3.44 (1 H, at, $J_{1,2}$ = $J_{2,3}$ = 9.0 Hz, H-2), 3.85 (3 H, s, OCH_3), 3.97 (1 H, ddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ 4.5 Hz, $J_{5,6b}$ = 2.2 Hz, H-5), 4.11 (1 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6b}$ = 1.9 Hz, H-6b), 4.36 (1 H, dd, $J_{6a,6b}$ = 12.5 Hz, $J_{5,6a}$ 4.4 Hz, H-6a), 5.13 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.7 Hz, H-4), 5.42 (1 H, t, J 9.6 Hz, H-3), 5.93 (1 H, d, $J_{1,2}$ = 8.1 Hz, H-1), 6.90 (2 H, d, J = 8.8 Hz, Ar-H), 7.64 (2 H, d, J = 8.8 Hz, Ar-H), 8.14 (1 H, s, benzylic-H), 9.86 (1 H, s, NH); MS *m/z* (ESI^+): 466.2 $[M+H]^+$.

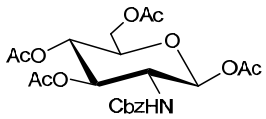
1,3,4,6-Tetra-O-acetyl-2-deoxy-β-D-glucosamine (177)[43]



176 (67.4 g, 0.145 mol) was dissolved in warm acetone (600 mL) before the addition of 5 M aqueous HCl (30 mL) causing the immediate formation of a precipitate. The reaction mixture was allowed to cool to RT prior to the addition of Et_2O (300 mL). The reaction mixture was then cooled to 4 °C before being filtered and washed with Et_2O and dried to afford the desired product (41.7 g, 75%).

TLC (petrol:EtOAc, 1:1, v/v): $R_f = 0$; M.p.= 220-222 °C. (Lit. m.p. = 235 °C)[43]; $[\alpha]_D^{25} = +36.5$ (c = 1.05, H_2O). [Lit $[\alpha]_D^{21} = +29.7$ (H_2O)][43]; ^{1}H NMR (400 MHz, DEUTERIUM OXIDE) δ ppm 1.94 (3 H, s, 1 x $OCOCH_3$), 2.00 (3 H, s, 1 x $OCOCH_3$), 2.07 (3 H, s, 1 x $OCOCH_3$), 2.08 (3 H, s, 1 x $OCOCH_3$), 3.64 (1 H, dd, $J_{2,3}$ 10.6 Hz, $J_{1,2}$ = 8.8 Hz, H-2), 4.04 - 4.11 (2 H, m, H-6b, H-5), 4.24 (1 H, d, $J_{6a,6b}$ = 13.5 Hz, $J_{5,6a}$ = 4.0 Hz, H-6a), 5.00 (1 H, at, J = 9.5 Hz, H-4), 5.35 (1 H, dd, $J_{2,3}$ = 10.7 Hz, $J_{3,4}$ = 9.0 Hz, H-3), 5.84 (1 H, d, $J_{1,2}$ = 8.8 Hz, H-1); MS *m/z* (ESI^+): 348.1 $[M+H]^+$.

1,3,4,6-Tetra-O-acetyl-2-deoxy-2- benzyloxycarbonylamino-β-D-glucopyranoside(178)[44]

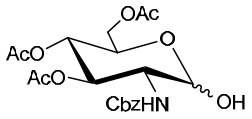


177 (16.5 g, 43.2 mmol, 1 equiv) was dissolved in H_2O (250 mL) before the addition of $NaHCO_3$ (6 g, 70.44 mmol, 1.6 equiv) following which the reaction was stirred for 10 mins prior to cooling to 0ºC. A solution of CBz-Cl (6.5 mL, 48 mmol, 1.14 equiv) in $CHCl_3$ (200 mL) and the reaction mixture stirred for 1.5 hours. The phases were then separated and the aqueous layer washed with $CHCl_3$ (70 mL). The combined organic phases were then washed with H_2O (100 mL) before drying with magnesium sulfate, filtering and concentrating. The crude product was then recrystallized from Et_2O to afford (10.5 g, 51%) of the desired product as a white crystalline solid.

TLC (petrol:EtOAc, 1:1, v/v): $R_f = 0.45$; M.p.= 116-118 °C. (Lit. m.p. = 130-132 °C)[44] $[\alpha]_D^{21} = +10.4$ (c = 1.0, $CHCl_3$). [Lit. $[\alpha]_D^{21} = +17$ (c = 2.8, $CHCl_3$)][44] ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.94 (3 H, s, CH_3), 2.03, (6 H, s, 2 x Me), 2.06, (3 H, s, Me), 3.80 (1 H, ddd, $J_{4,5}$ = 9.9 Hz, $J_{5,6}$ = 4.4 Hz, $J_{5,6'}$ = 1.7 Hz, H-5) 3.99-3.92 (1 H, m, H-2), 4.10 (1 H, ad, J = 12.5 Hz, H-6b), 4.27 (1 H, dd, $J_{6a,6b}$ = 12.6 Hz, $J_{5,6a}$ = 4.4 Hz, H-6a), 5.00-4.98 (1 H, m, H-3), 5.12 (2 H, m, 2 x OCH_2Ph Cbz), 5.17

(1 H, at, *J* = 9.9 Hz, H-4), 5.68 (1 H, d, $J_{1,2}$ = 8.5 Hz, H-1), 7.36-7.30 (5 H, m, Ar-H); MS *m/z* (ESI^-): 480 ($M-H^-$).

3,4,6-Tri-O-acetyl-2-benzyloxycarbonylamino-2-deoxy-β-D-glucopyranose (138)[45]

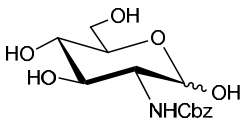


178 (5.56 g, 12.0 mmol, 1 equiv) was dissolved in THF (200 mL) before the addition of $BnNH_2$ (1.3mL, 12.1 mmol, 1.01 equiv) following which the reaction was stirred for 48 hours at which time TLC (petrol:EtOAc, 1:1) showed product formation (R_f 0.3) along with the presence of starting material (R_f 0.45). A further (1 mL, 9.2 mmol, 0.76 equiv) of $BnNH_2$ was added and stirring continued for a further 24 hours, upon which time TLC revealed complete disappearance of starting material starting material (R_f 0.6). The crude reaction mixture was concentrated under reduced pressure. Column chromatography (petrol:EtOAc, 2:1 to 1:1) afforded the desired product (4.9 g 75%) as an extremely viscous pale yellow oil.

TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.3; $[\alpha]_D^{21}$ = +114.6 (c = 2.0, $CHCl_3$). [Lit $[\alpha]_D^{21}$ = +31 (c = 1.0, $CHCl_3$)][45]; ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 2.06 (3 H, s, 1 x $OCOCH_3$), 2.02 (3 H, s, 1 x $OCOCH_3$), 1.94 (3 H, s, 1 x $OCOCH_3$), 4.02 (1 H, atd, *J* = 10.2, 3.5 Hz, H-2), 4.14-4.09 (1 H, m, H-6'), 4.23-4.16 (2 H, m, H-6 H-5), 5.02 (1 H, d, *J* = 12.4 Hz, 1 x OCH_2Ph Cbz), 5.10 (1 H, at, *J* = 9.6 Hz, H-4), 5.13 (1 H, d, *J* = 12.4 Hz, 1 x OCH_2Ph Cbz), 5.26 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1α), 5.25 (1 H, at, $J_{2,3} = J_{3,4}$ = 9.8 Hz, H-3), 7.36-7.26 (5 H, m, Ar-H); MS *m/z* (ESI^+): 462 $[M+Na]^+$.

8.2.8.1 Alternate synthesis of compound 138

2-Deoxy-2- benzyloxycarbonylamino-αβ-D-glucopyranoside(180)[46,47]

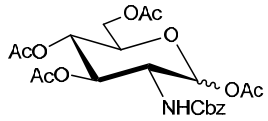


175 (52.4 g, 243.7 mmol) and Na_2CO_3 (30 g, 283 mmol) were dissolved in distilled water (400 mL) and to this was added benzyl chloroformate (40 mL, 272.7 mmol, 1.12 eq) and the reaction was stirred at RT for 10 minutes, following which it was cooled to O °C and stirred for 45 minutes upon which time the desired product had precipitaed from the solution as a white solid. TLC (MeOH:EtOAc, 1:1) revealed complete consumtion of starting material (R_f 0.0) and production of a single product (R_f

0.75). Precipitate was filtered off and washed with cold water (300 mL) and ether (300 mL) to yield the final product as a white solid (67.7 g, 87%)

TLC (MeOH:EtOAc, 1:1, v/v): R_f = 0.75; $[\alpha]_D^{25}$ + 69.8 (c = 1 in pyridine) [Lit. $[\alpha]_D^{20}$ +73 (c = 1 in pyridine)][46]; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 3.0 - 3.1 (2 H, C-5α, C-2β), 3.1 (1 H, ddd, $J_{4,5}$ = 10.2 Hz, $J_{5,6a}$ = 7.5 Hz, $J_{5,6b}$ = 2.0 Hz, C-5β), 3.3 (1 H, dd, $J_{2,3}$ = 9.0 Hz, $J_{1,2}$ = 3.9 Hz, H-2α), 3.4 (1 H, dd, $J_{6a,6b}$ = 12.4 Hz, $J_{5,6a}$ = 5.2 Hz, H-6αa), 3.5 - 3.5 (3 H, m, H-6β, H-3α, H-3β), 3.5 - 3.6 (3 H, H-6β, H-4α, H-4β), 3.7 (1 H, dd, $J_{6a,6b}$ = 12.1 Hz, $J_{5,6b}$ = 2.2 Hz, H-6αb), 4.4 (4 H, d, $J_{1,2}$ = 8.8 Hz, H-1β), 5.0 (1 H, d, $J_{1,2}$ = 3.4 Hz, H-1α), 5.0 (4 H, m, 4 x OCH_2Ph Cbz), 7.2 - 7.5 (10 H, m, Ar-H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 57.2 (C-2α), 59.5 (C-2β) 61.9 (C-6 α), 62.0 (C-6β), 65.9 (1 x OCH_2Ph Cbz αβ), 66.1 (1 x OCH_2Ph Cbz αβ), 71.1 (2 C, C-3α, C-3β), 71.7 (C-5α or C-5β) , 71.8 (C-5α or C-5β), 72.8 (C-4α, C-4β), 91.5 (C-1α), 96.3 (C-1 β), 127.3, 127.5, 128.6, 128.9, 129.2, 129.5, 129.7 (1 x OCH_2Ph Cbz αβ), 137.9 (1x 1C. 1 x OCH_2Ph Cbz αβ), 157.0 (1 x C=O Cbz αβ), 157.1 (1 x C=O Cbz αβ); MS *m/z* (ESI^+): 466.2 $[M+H]^+$.

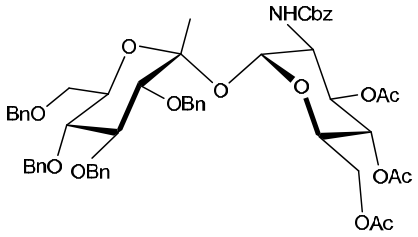
1,3,4,6-Tetra-O-acetyl-2-deoxy-2- benzyloxycarbonylamino-αβ-D-glucopyranoside (181)[47]



180 (51.4 g, 164.2 mmol, 1 eq) was dissolved in pyridine (400 mL) and to this was added acetic anhydride (100 mL). Reaction was stirred for 16 hours under argon atmostphere, upon which time TLC (petrol:EtOAc, 1:1) showed complete consumption of starting material (R_f 0) and production of a new product spot (0.45). Reaction was concentrated under reduced pressure and column chromatography (petrol:EtOAc, 1:1) yielded the desired product as a clear oil and 10:7 αβ mixture (29 g, 37%)

TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.45; ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.92 (3 H, s, 1 x $OCOCH_3$α), 1.94 (3 H, s, 1 x $OCOCH_3$β), 2.00 (3 H, s, 1 x $OCOCH_3$α.), 2.01 (3 H, s, 1 x $OCOCH_3$α) 2.02 (3 H, s, 1 x $OCOCH_3$β), 2.03 (3 H, s, 1 x $OCOCH_3$β), 2.06 (3 H, s, 1 x $OCOCH_3$β), 2.07 (3 H, s, 1 x $OCOCH_3$α), 3.80 (1 H, ddd, $J_{4,5}$ = 9.9 Hz, $J_{5,6}$ = 4.4 Hz, $J_{5,6'}$ = 1.7 Hz, H-5β) 3.92-4.04 (2 H, m, H-2β, H-5α), 4.04 (1 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6b}$ = 1.8 Hz, H-6βb), 4.09 (1 H, dd, $J_{6a,6b}$ = 10.2 Hz, $J_{5,6a}$ = 3.8 Hz, H-6αb), 4.12 (1 H, dd, $J_{6a,6b}$ = 10.2 Hz, $J_{5,6a}$ = 4.8 Hz, H-6αa), 4.19 (1 H, dd, $J_{2,3}$ = 10.1 Hz, $J_{1,2}$ = 3.4 Hz, H-2α), 4.26 (2 H, dd, $J_{6a,6b}$ = 12.6, $J_{5,6a}$ = 4.4 Hz, H-6βa), 4.98-5.0 (4 H, m, 2 x OCH_2Ph Cbz, H-3α, H-3β), 5.17 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.9 Hz, H-4β), 5.2 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 10.1 Hz, H-4α), 5.68 (1 H, d, $J_{1,2}$ = 8.5 Hz, H-1β), 6.2 (1 H, d, *J* = 3.5 Hz, H-1α), 7.36-7.30 (10 H, m, Ar-H); MS *m/z* (ESI^-): 480 $[M-H^-]$.

3,4,5,7-tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-3,4,6-tetra-O-acetyl-2-benzyloxycarbonylamino-2-deoxy-α-D-glucopyranoside (150)

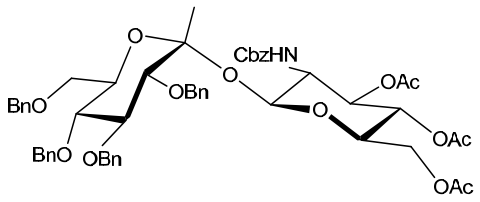


85 (173.8 g, 0.31 mmol, 1 eq) and **138** (168.0 mg, 0.38 mmol, 1.2 eq) were dried under reduced pressure for 1 hour and then dissolved in anhydrous DCM (10 mL) and added to a dry flask in the presence of molecular sieves (ca. 100 mg). Mixture was stirred with molecular sieves for 30 minutes at room temperature and then was cooled to –40 °C. To this was added TMSOTf (10 µl, 0.054 mmol. 0.15 eq) at –40 °C in the under an Ar atmosphere. The resulting mixture was stirred for 15 minutes until TLC (petrol:EtOAc, 3:2) revealed the production of two new products (R_f 0.6 and R_f 0.15) and consumption of starting material (R_f 0.35 and R_f 0.05). The reaction was then quenched by the addition triethylamine (0.1 mL), filtered through Celite® and concentrated. The crude product was purified by column chromatography (petrol:EtOAc, 5:2) to yield the desired compound α,α (214 mg) as well as the α,β (32 mg) for a net yield (81%, 6:1 α,α: α,β) as well as recovered **138** (14.5 mg, 8%).

TLC (petrol:EtOAc, 3:2, v/v): R_f = 0.60; $[\alpha]_D^{25}$ = 67.3 (c = 0.94, $CHCl_3$).^{1}H NMR (500 MHz, $CDCl_3$): δ 1.48 (3 H, s, C-1'), 1.96 (3 H, s, 1 x OCOC$\underline{H}_3$), 2.01 (3 H, s, 1 x OCOC$\underline{H}_3$), 2.04 (3 H, s, 1 x OCOC$\underline{H}_3$), 3.33 (1 H, d, $J_{3,4}$ = 9.4 Hz, C-3'), 3.43 (1 H, dd, $J_{7a,7b}$ = 10.9 Hz, $J_{6,7b}$ = 1.42 Hz, H-7b'), 3.53 (1 H, dd, $J_{7a,7b}$ = 11.03 Hz, $J_{6,7a}$ = 2.8 Hz, H-7a'), 3.68 - 3.79 (3 H, m, H-5', H-6b, H-6'), 3.97 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.1 Hz, H-4'), 4.04 (2 H, dd, $J_{5,6a}$ = 12.6 Hz, $J_{6a,6b}$ = 4.41 Hz, H-6a), 4.05 (1 H, dd, $J_{2,3}$ = 10.7 Hz, $J_{1,2}$ = 3.8 Hz, H-2), 4.26 (1 H, ddd, $J_{4,5}$ = 10.3 Hz, $J_{5,6a}$ 4.5 Hz, $J_{5,6b}$ = 2.21 Hz, H-5), 4.38 (1 H, d, J = 12.3 Hz, 1 x OCH_2Ph), 4.53 (1 H, d, J = 12.3 Hz, 1 x OCH_2Ph), 4.56 (1 H, d, J = 11.0 Hz, 1 x OCH_2Ph), 4.61 (1 H, d, J = 11.4 Hz, 1 x OCH_2Ph), 4.82 (1 H, d, J = 11.0 Hz, 1 x OCH_2Ph), 4.91 (1 H, d, J = 11.9 Hz, 1 x OCH_2Ph Cbz), 4.93 (1 H, d, J = 10.9 Hz, 1 x OCH_2Ph), 4.95 (1 H, d, J = 12.6 Hz, 1 x OCH_2Ph), 4.97 (1 H, d, J = 11.4 Hz, 1 x OCH_2Ph Cbz), 5.05 (1 H, d, J = 12.0 Hz, 1 x OCH_2Ph), 5.08 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.1 Hz, H-4), 5.30 (1 H, dd, $J_{2,3}$ = 10.7 Hz, $J_{3,4}$ = 9.4 Hz, H-3), 5.36 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1), 6.98 - 7.41 (25 H, m, Ar-H); ^{13}C NMR (126MHz, $CDCl_3$): δ 20.6, 20.7 (3 x $OCOCH_3$), 23.1 (C-1'), 54.6 (C-2), 61.7 (C-6), 67.0 (O$\underline{C}H_2Ph$ Cbz), 67.2 (C-5), 68.1 (C-7'), 68.3 (C-4), 70.9 (C-3), 72.3 (C-5'), 73.4 (O$\underline{C}H_2Ph$), 74.9

(O<u>C</u>H$_2$Ph), 75.5 (O<u>C</u>H$_2$Ph), 75.7(O<u>C</u>H$_2$Ph), 77.9 (C-6'), 82.8 (C-4'), 84.7 (C-3'), 90.7 (C-1), 101.6 (C-2'), 127.6, 127.6, 127.6, 127.8, 127.8, 128.1, 128.3, 128.4, 128.4 (5 x OCH$_2$Ph), 136.0, 138.1, 138.2, 138.3, 138.5 (5x 1C. 5 x OCH$_2$Ph), 155.6 (1 x C=O Cbz), 169.4, 170.6, 171.4 (3 x C=O) ; IR (thin film): 3346 (br. Amide), 3089, 2924, 2864 (C=CH), 1745 (C=O), 1720 (C=O), 1564, 1530, 1454, 1365, 1230, 1153, 1028, 737; HRMS *m/z* (ESI$^+$): $[M+NH_4]^+$ calcd. for $C_{55}H_{65}N_2O_{15}^+$, 993.4379; found, 993.4377; analysis (calcd., found for $C_{55}H_{61}NO_{15}$): C (67.68,67.61), H (6.30,6.21), N(1.44,1.44).

3,4,5,7-tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-3,4,6-tetra-O-acetyl-2-benzyloxycarbonylamino-2-deoxy-β-D-glucopyranoside (179)

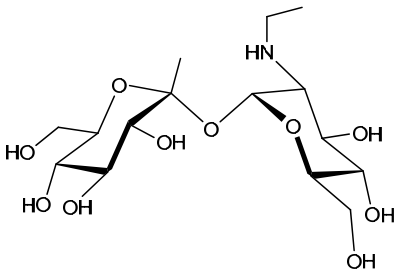


179 was isolated as the lower spot (R_f 0.15) of the reaction between **85** and **138** (212 mg, 7%).

TLC (petrol:EtOAc, 3:2, v/v): R_f = 0.15; $[\alpha]_D^{25}$ = 37.7 (c = 1.0, $CHCl_3$).^{1}H NMR (500 MHz, $CDCl_3$): δ 1.36 (3 H, s, C-1'), 1.94 (3 H, s, 1 x OCOC<u>H</u>$_3$), 2.00 (3 H, s, 1 x OCOC<u>H</u>$_3$), 2.01 (3 H, s, 1 x OCOC<u>H</u>$_3$), 3.35 (1 H, d, $J_{3,4}$ = 9.5 Hz, H-3'), 3.46 (1 H, br. s., H-2), 3.62 (2 H, adt, $J_{4,5}$ = 10.7 Hz , $J_{5,6a}$ = $J_{5,6b}$ = 1.1 Hz, H-5, H-7b'), 3.71 (2 H, at, $J_{4,5}$ = $J_{45,6}$ = 9.8 Hz, H-5'), 3.74 (1 H, dd, $J_{7a,7b}$ = 10.7 Hz, $J_{6,7a}$ = 3.2 Hz, H-7a'), 3.99 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.5 Hz, H-4'), 4.05 (1 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6b}$ = 2.2 Hz, H-6b), 4.12 (1 H, dd, $J_{6a,6b}$ = 12.0 Hz, $J_{5,6a}$= 4.7 Hz, H-6a), 4.17 (1 H, ddd, $J_{5,6}$ = 10.7 Hz, $J_{6,7a}$ = 3.3 Hz, $J_{6,7b}$ = 2.0 Hz, H-6'), 4.49 (1 H, d, *J* = 12.3 Hz, 1 x OCH$_2$Ph), 4.56 (1 H, d, *J* = 10.7 Hz, 1 x OCH$_2$Ph), 4.60 (1 H, d, *J* = 10.7 Hz, 1 x OCH$_2$Ph), 4.61 (1 H, d, *J* = 12.3 Hz, 1 x OCH$_2$Ph), 4.77 - 4.85 (3 H, m, 1 x OCH$_2$Ph Cbz, 2 x OCH$_2$Ph), 4.88 (2 H, d, *J* 11.0 Hz, 1 x OCH$_2$Ph), 4.95 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.8 Hz, H-4), 5.06 - 5.18 (2 H, m, H-1, 1 x OCH$_2$Ph Cbz), 5.46 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.6 Hz H-3), 7.14 - 7.37 (35 H, m, Ar-H); ^{13}C NMR (126MHz, $CDCl_3$): δ 20.6, 20.7, 21.0 (3 x OCOCH$_3$), 22.1 (CH$_3$), 56.5 (C-2), 62.4 (C-6), 66.9 (O<u>C</u>H$_2$Ph Cbz), 68.4 (C-4), 68.8 (C-7'), 71.85 (C-3), 72.0 (C-5), 72.3 (C-6'), 73.5 (O<u>C</u>H$_2$Ph), 74.9 (O<u>C</u>H$_2$Ph), 75.4 (O<u>C</u>H$_2$Ph), 75.8(O<u>C</u>H$_2$Ph), 78.2 (C-5'), 82.8 (C-4'), 84.7 (C-3'), 94.6 (C-1), 102.3 (C-2'), 127.5, 127.6, 127.6, 127.7, 127.8, 127.9, 128.0, 128.1, 128.3, 128.4, 128.5 (5 x OCH$_2$Ph), 136.2, 137.9, 138.2, 138.3, 138.7 (5x 1C. 5 x OCH$_2$Ph), 155.5 (1 x C=O Cbz), 169.6, 170.4, 170.6 (3 x C=O); IR (thin film): 3356 (br. amide), 3031, 2939 (C=CH), 1748 (C=O), 1587, 1454, 1367, 1231, 1043, 739; HRMS *m/z* (ESI$^+$): $[M+Na]^+$ calcd. for $C_{55}H_{61}NO_{15}Na^+$, 998.3933; found, 998.3905.

8.2.8.2 Deprotection of compound compound 150

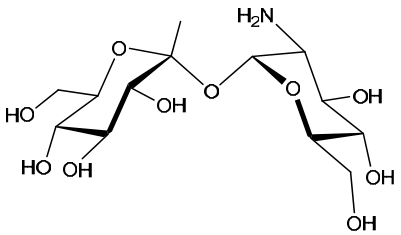
1- Deoxy-α-D-gluco-hept-2-ulopyranosyl -(2→1)-2 -N-ethyl-2-deoxy-α-D-glucopyranoside (181)



150 (45.0 mg, 0.046 mmol was dissolved in anhydrous methanol (10 mL). To this was added NaOMe (17.6 mg, 0.325 mmol). Reaction was stirred at room temperature for three hours until complete disappearance of starting material TLC (petrol:EtOAc, 2:1) (R_f 0.3) and formation of a new product spot (R_f 0.0) was detected. Product was also detected by ESI 872.3 ($M+Na^+$). Reaction was neutralized with Dowex 50WX8 (H^+ form) cation exchange resin and concentrated under reduced pressure. Product was redissolved in ethanol and to this was added $Pd(OH)_2/C$ (120 mg) and basic alumina (60 mg). Reaction was stirred for 18 hours under hydrogen atmosphere, upon which time the desired product was detected by ESI *m/z* 354.1(M-H) along with **181** *m/z* 382.1 as an inseparable mixture. TLC ($EtOH:NH_4:H_2O$, 5:3:1) revealed one spot as a smear (R_f 0.4-0.6). **181** was characterized by NMR and found to be an inseparable 2:1 mixture of the reductive amination product as well as another unidentified side product (14.9mg, 85%).

TLC ($EtOH:NH_4:H_2O$, 5:3:1, v/v): R_f = 0.4-0.6; HRMS *m/z* (ESI^+): $[M+H]^+$ calcd. for $C_{15}H_{30}NO_{10}^+$, 384.1870; found, 384.1860.

***1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2–amino-2-deoxy-α-D-glucopyranoside glucopyranoside* (37)**

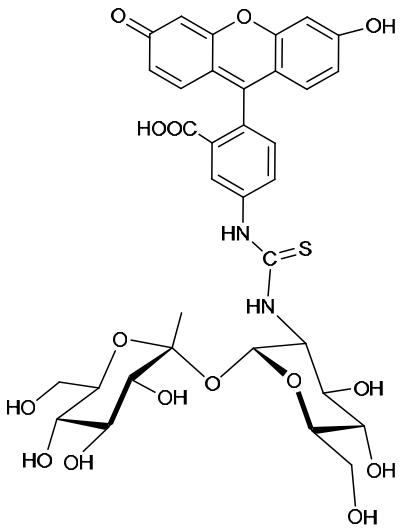


150 (578 mg, 0.59 mmol, 1 eq) was dissolved in anhydrous methanol (15 mL). To this was added sodium methoxide (25 mg, 0.46 mmol, 0.94 eq) and reaction was stirred for 1 hour, upon which time full conversion to product was detected by TLC (EtOAc) (R_f 0.75) and disappearance of **150** (R_f 1).

Reaction was neutralized with DOWEX 50WX8 (H^+ form) cation exchange resin (ca 50 mg). DOWEX was removed by filtration and reaction was concentrated under reduced pressure to yield the deacetylated product (507 mg, 100%). This product was split into two portions and each portion was dissolved in 20 mL (1:1 trifluoroethanol/water) with formic acid (50 µl) and cycled through the Thales Nano H-Cube® over a Pd/C cartridge at 70 bar for 10 hours. Upon completion product was detected by TLC (H_2O:*i*-PrOH:EtOAc, 1:2:2) (R_f 0.05) with complete disappearance of starting sugar (R_f 1). Reactions were concentrated under reduced pressure, redissolved in water and lyophilized to yield the desired product as a brownish, amorphous solid as the formate salt. Amine was further purified by a 10 mL column of DOWEX 50WX8 (H^+ form) cation exchange resin. Amine was loaded onto resin, washed with water (20 mL), 0.1% NH_4OH (20 mL),eluted with 5 % NH_4OH (20 mL) and evaporated to dryness to yield the desired product as a white amorphous solid. (221 mg, 93%).

TLC (water:isopropanol:ethyl acetate, 1:2:2, v/v): R_f=0.05; $[\alpha]_D^{25}$ = 83.8 (c = 0.21, MeOH); ^{1}H NMR (500 MHz, D_2O): δ 1.48 (3 H, s, C-1'), 2.74 (1 H, d, *J* = 10.4 Hz, H-2), 3.23 (1 H, d, $J_{3,4}$ = 9.7 Hz, H-3'), 3.35 (2 H, atd, *J* = 9.6, 5.7 Hz, H-5, H-6'), 3.60 - 3.83 (8 H, m, H-7a', H-7b', H-6a, H-6b, H-4', H-5', H-3, H-4), 5.24 (1 H, d, $J_{1,2}$ = 3.2 Hz, H-1); ^{13}C NMR (126 MHz, D_2O): δ 22.9 (C-1'), 55.5 (C-2), 60.5 (C-6), 60.6 (C-7'), 69.7 (C-5 or C-6'), 69.8 (C-5 or C-6'), 72.1 (C-4), 72.6 (C-4' or C-5'), 72.7 (C-4' or C-5'), 73.5 (C-3) 76.4 (C-3'), 91.95 (C-1), 101.3 (C-2'); HRMS *m/z* (ESI$^+$): $[M+Na]^+$ calcd. for $C_{13}H_{25}NO_{10}Na^+$, 378.1317; found: 378.1361.

***1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2-deoxy-2-(N'-(fluorescein-5-yl)-thioureido)-α-D-glucopyranoside* (9)**



37 (11 mg, 0.03 mmol, 1 eq) and fluorescein isothiocyanate (17 mg, 0.044 mmol, 1.4 eq) were dissolved in 75 mM $NaHCO_3$ buffer at pH 9 (1 mL) with acetonitrile (0.5 mL). Reaction was heated to

50 °C for 2 upon which time product (R_f 0.3) was detected by TLC (H_2O:*i*-PrOH:EtOAc, 1:2:2) with near complete disappearance of fluorescein starting material (R_f 0.8) and starting amine (R_f 0.05). Reaction mixture was purified by HPLC with a Phenomenex Synergi Hydro C_{18} column (150 mm x 21.2 mm, 4 μm) and a acetonitrile gradient with 1% aqueous TFA. Lyophilization yielded the desired product as a yellow solid (16 mg, 72 %).

TLC (H_2O:*i*-PrOH:EtOAc, 1:2:2, v/v/v): R_f = 0.30; $[\alpha]_D^{25}$ = 72.2 (c = 0.18, MeOH); ^{1}H NMR (500 MHz, D_2O): δ 1.48 (3 H, s, C-1'), 3.26 (1 H, d, $J_{3,4}$ = 9.8 Hz, H-3'), 3.49 (3 H, m, H4', H-5', H-4), 3.74 (6 H, m, H5, H-6', H-6a, H-6b, H-7a', H-7b'), 4.00 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.8 Hz, H-3), 4.41 (1 H, dd, $J_{2,3}$ = 10.9 Hz, $J_{1,2}$ = 2.9 Hz, H-2), 5.59 (1 H, d, $J_{1,2}$ = 3.6 Hz, H-1), 6.84 (4 H, t, *J* = 0.9 Hz), 7.12 - 7.18 (3 H, m), 7.68 (1 H, dt, *J* = 7.7, *J* = 0.9 Hz), 8.00 (1 H, d, *J* = 1.0 Hz); ^{13}C NMR (126 MHz, D_2O) δ 22.79 (C-1'), 58.8 (C-2), 60.0 (C-7'), 60.56 (C-6), 68.9 (C-4), 70.4 (C-3, C-4'), 71.9 (C-6'), 72.8 (C-5), 72.8 (C-5'), 76.3 (C-3'), 89.6 (C-1), 101.5 (C-2'), 102.5, 113.9, 115.1, 117.1, 117.4, 128.3, 130.2, 131.4, 140.4, 156.2, 162.8, 163.1, 166.0, 170.2, 181.6; HRMS *m/z* (ESI^+): $[M-H]^-$ calcd. for $C_{34}H_{35}N_2O_{15}S^-$, 743.1764; found, 743.1754; analysis (calcd., found for $C_{34}H_{36}N_2O_{15}S$): C (54.83,54.77), H (4.87,4.80), N (3.76,3.69), S(4.31, 4.26).

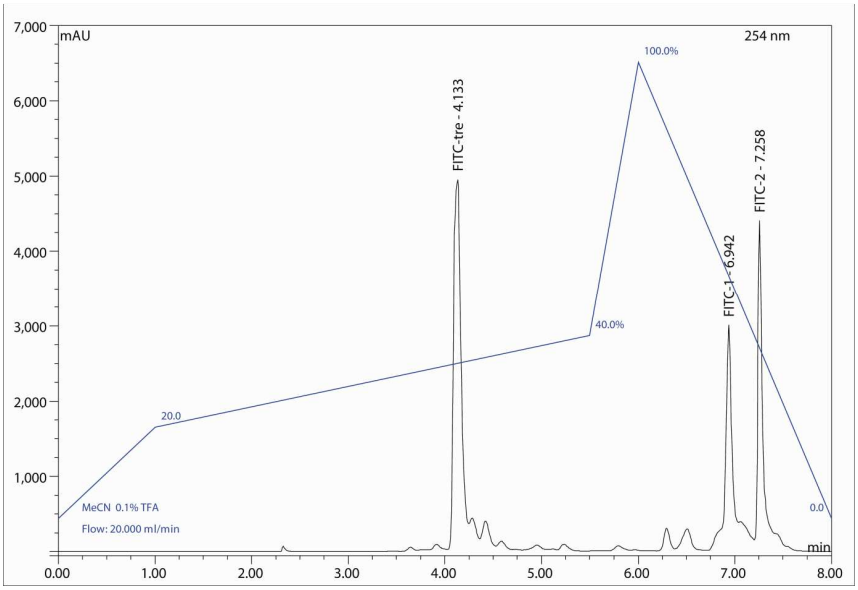
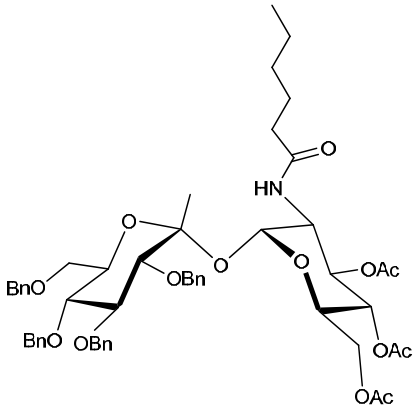


Figure 8.8. Preparative HPLC trace for the purification of FITC-Tre compound 40. Conditions: Phenomenex Synergi Hydro C18 column (150 mm x 21.2 mm, 4 μm), eluting with water with 0.1 % TFA (A) and acetonitrile (B) eluting at a flow rate of 20mL/min; Gradient: 0 min B=%0, 1 min B=20%, 5.5 min B=40%, 6 min B=100%, 8 min B=0%; Compound 40 elutes at 4.1 min (labeled FITC-Tre) while starting materials FITC elute at 6.9 and 7.3 min.

8.2.9 Other selective modifications of 37

3,4,6,7-Tetra-O-benzyl-1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-3,4,6-O-acetyl-2-N-hexanoyl-2-deoxy-α-D-glucopyranoside (185)

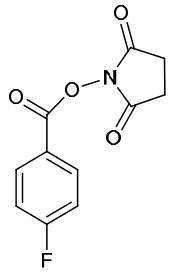


Crude **184** was dissolved in anhydrous DCM (5 mL). Reaction was stirred O/N at RT under argon atmosphere. TLC (petrol:EtOAc, 2:1) revealed complete dissappearance of starting material (R_f 0.05) and appearance of product (R_f 0.35). Reaction was concentrated under reduced pressure and purified by column chromatography (petrol:EtOAc, 5:2) with 1% NEt_3 to yield the desired product (12.1 mg, 42.1% over two steps) as a clear oil.

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.35; $[\alpha]_D^{21}$ = 42.6 (c = 0.5, $CHCl_3$).^{1}H NMR (500 MHz, CHLOROFORM-*d*) δ ppm 0.83 (3 H, t, $J_{CH3,CH2}$ = 7.3 Hz, $OCO(CH_2)_4C\underline{H}_3$), 1.03 - 1.14 (2 H, m, $OCO(CH_2)_3\ C\underline{H}_2CH_3$), 1.19 (2 H, quin, *J* = 6.9 Hz, $OCO(CH_2)_2\ \underline{C}H_2\ CH_2CH_3$), 1.34 - 1.52 (4 H, m, $OCO(CH_2)_2\ C\underline{H}_2\ CH_2CH_3$, $OCOCH_2\ C\underline{H}_2\ (CH_2)_2CH_3$), 1.48 (3 H, s, CH_3), 1.97 - 2.01 (2 H, m, $OCOC\underline{H}_2(CH_2)_3\ CH_3$), 2.02 (3 H, s, 1 x $OCOC\underline{H}_3$), 2.03 (3H, s, 1 x $OCOC\underline{H}_3$), 2.05 (3 H, s, 1 x $OCOC\underline{H}_3$), 3.34 (1 H, d, $J_{3,4}$ = 9.5 Hz, H-3'), 3.55 (1 H, dd, $J_{6,7}$ = 10.4 Hz, $J_{7a,7b}$ = 2.2 Hz, H-7b'), 3.64 (2 H, m, H-7a', H-6'), 3.72 (1 H, at, $J_{4,5} = J_{5,6}$= 10.1 Hz, H-5'), 3.84 (1 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6a}$ = 2.5 Hz, H-6b'), 3.98 (1 H, at, $J_{3,4} = J_{4,5}$= 9.1 Hz, H-4'), 4.07 (1 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6a}$ = 4.7 Hz, H-6a) 4.22 (1 H, ddd, $J_{4,5}$ = 10.2 Hz, $J_{5,6a}$ = 4.6 Hz, $J_{5,6b}$ = 2.2 Hz, H-5), 4.32 (1 H, ddd, $J_{2,3}$ = 11.0 Hz, $J_{2,NH}$ = 8.5 Hz, $J_{1,2}$ = 3.8 Hz, H-2), 4.45 (1 H, d, *J* = 12.3 Hz, 1 x OCH_2Ph), 4.54 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.60 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.64 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 4.83 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.92 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.96 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.98 (1 H, d, *J* = 11.3 Hz, 1 x OCH_2Ph), 5.11 (1 H, at, $J_{3,4} = J_{4,5}$= 9.8 Hz, H-4), 5.31 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1), 5.33 (1 H, dd, $J_{2,3}$ = 11.2 Hz, $J_{3,4}$ = 9.3 Hz, H-3), 5.72 (1 H, d, *J* = 8.8 Hz, NH), 6.80 - 7.59 (20 H, m, Ar-H);^{13}C NMR (126 MHz, CHLOROFORM-*d*) δ ppm 13.9 ($OCO(CH_2)_4\underline{C}H_3$), 20.6 (1 x $OCOCH_3$), 20.7 (1 x $OCOCH_3$), 20.8 (1 x $OCOCH_3$), 22.3 (CH_3) 23.1($OCO(CH_2)_3\ \underline{C}H_2CH_3$), 24.9 ($OCO(CH_2)_2\ \underline{C}H_2\ CH_2CH_3$), 31.2 ($OCOCH_2\ \underline{C}H_2\ (CH_2)_2CH_3$), 36.3

($OCO\underline{C}H_2(CH_2)_3\ CH_3$), 52.7 (C-2), 61.8 (C-6), 67.4 (C-5), 68.1 (C-4), 68.1 (C-7'), 71.2 (C-3), 72.3 (C-6'), 73.5 (1 x OCH_2Ph), 75.1 (1 x OCH_2Ph), 75.5 (1 x OCH_2Ph), 75.6 (1 x OCH_2Ph), 77.9 (C-5') 82.7 (C-4'), 84.6 (C-3'), 90.5 (C-1), 101.5 (C-2'), 127.6, 127.6, 127.7, 127.8, 127.9, 128.3, 128.3, 128.3, 128.4, 128.4 (4 x OCH_2Ph), 137.9, 138.0, 138.3, 138.4 (4 x 1C, 4 x OCH_2Ph), 169.3, 170.6, 171.8, 172.9 (4 x C=O); HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{55}H_{65}NO_{14}Na^+$, 962.429; found, 962.4314.

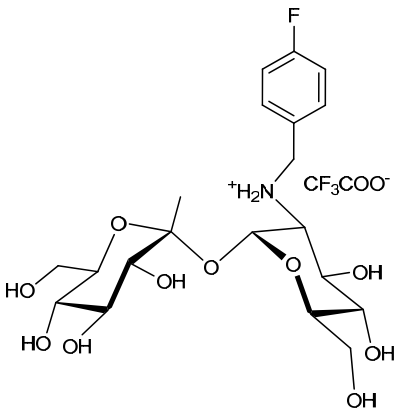
N-succinimidyl 4-fluorobenzoate (188)



4-Fluorobenzoic acid **187** (1.99 g, 14.2 mmol) DCC (3.5 g, 17.4 mmol) and NHS (1.9 g, 17.04 mmol) were dissolved in anhydrous THF (60mL). Reaction mixture was stirred O/N at RT. Product (R_f 0.35) was detected by TLC (petrol:EtOAc, 2:1) with nearly complete dissappearance of starting materials (R_f 0.0, R_f 0.98, (R_f 0.1-0.4)). Reaction was concentrated under reduced pressure and purified via column chromatography (petrol:EtOAc, 2:1 followed by petrol:EtOAc, 1:1). The desired compound was obtained as a white, crystaline solid (2 g, 60%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.35; ^{1}H NMR (400 MHz, CHLOROFORM-*d)* δ ppm 2.90 (4 H, s, CH_2), 7.05–7.30 (2 H, m, Ar-H meta), 8.05–8.30 (2 H, m, Ar-H ortho).

1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2-(4-fluorobenzylamino)-2-deoxy-α-D-glucopyranoside trifluoroacetic acid salt (38)



37 (11 mg, 0.03 mmol, 1 eq) was dissolved in methanol (1 mL). 4-fluoro-benzaldehyde (9.5 µl, 0.06 mmol, 2 eq) was added to the reaction. After 10 minutes, complete conversion to imine was detected by TLC (EtOAc:MeOH, 2:1) product (R_f 0.4). To this was added $NaBH_4$ (15.2 mg, 0.4 mmol, 13.1 eq) and reaction was allowed to stir for a further 10 minutes, upon which time reductive amination was detected by ESI mass spec as well as TLC (H_2O:*i*-PrOH:EtOAc, 1:2:2) product (R_f 0.75). Upon conversion, reaction was purified by HPLC with a HPLC with a Phenomenex Synergi Hydro C18 column (150 mm x 21.2 mm, 4 µm) and an MeCN/H_2O gradient (5%/min) with 0.1% NH_4OH. Lyophilization yielded the desired compound as a off-white amorphous solid as the TFA salt. (8.4 mg, 59%).

TLC (H_2O:*i*-PrOH:EtOAc, 1:2:2, v/v/v): R_f = 0.75; $[\alpha]_D^{25}$ = 32.4 (c = 0.29, MeOH); ^{1}H NMR (500 MHz, D_2O): δ 1.47 (3 H, s, C-1'), 2.63 (1 H, dt, $J_{2,3}$ = 10.4 Hz, $J_{1,2}$ = 2.8 Hz, $J_{2,NH}$ = 2.8 Hz, H-2), 3.22 (1 H, d, $J_{3,4}$ = 9.8 Hz, H-3'), 3.29 (1 H, at, $J_{4,5}$ = $J_{5,6}$ = 9.5 Hz, H-5'), 3.36 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4), 3.59 (1 H, dd, $J_{7a,7b}$ = 12.3 Hz, $J_{6,7a}$ = 5.0 Hz, H-7a'), 3.65 (1 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6a}$ = 5.0 Hz, H-6a), 3.67 (1 H, dd, $J_{6a,6b}$ = 12.6 Hz, $J_{5,6b}$ = 2.5 Hz, H-6b), 3.71 - 3.77 (3 H, m, H-7b', H-6', H-4'), 3.79 - 3.86 (3 H, m, H-3, 2 x NC<u>H</u>$_2$Ph), 3.86 (1 H, ddd, $J_{4,5}$ = 9.8 Hz, $J_{5,6a}$ = 4.4 Hz, $J_{5,6b}$ = 2.2 Hz, H-5), 5.36 (1 H, d, $J_{1,2}$ = 2.8 Hz, H-1), 7.05 (2 H, at, $J_{ortho,meta}$ = $J_{H,F}$ = 9.0 Hz, Ar-H_{ortho}), 7.30 (2 H, dd, $J_{ortho,meta}$= 8.4 Hz, $J_{H,F}$ = 5.8 Hz, Ar-H_{meta}); ^{13}C NMR (126 MHz, D_2O): δ 22.5 (C-1'), 50.0 (N<u>C</u>H$_2$Ph-F), 60.3 (C-2), 60.6 (C-6), 60.7 (C-7'), 69.6 (C-4), 70.4 (C-5'), 71.8 (C-4'), 72.4 (C-3), 72.6 (C-6'), 72.7 (C-5), 76.5 (C-3'), 90.5 (C-1), 101.4 (C-2'), 115.1, 115.2, 115.3, 117.3, 117.5, 119.8, 129.7, 129.8 (Ar-C), 162.8, 162.9 (C=O, TFA);^{19}F NMR (1H) (377 MHz, D_2O): δ –75.6 (TFA), -116.3 (4-F-benzyl); HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{20}H_{30}FNO_{10}Na^+$, 486.1746; found, 486.1770.

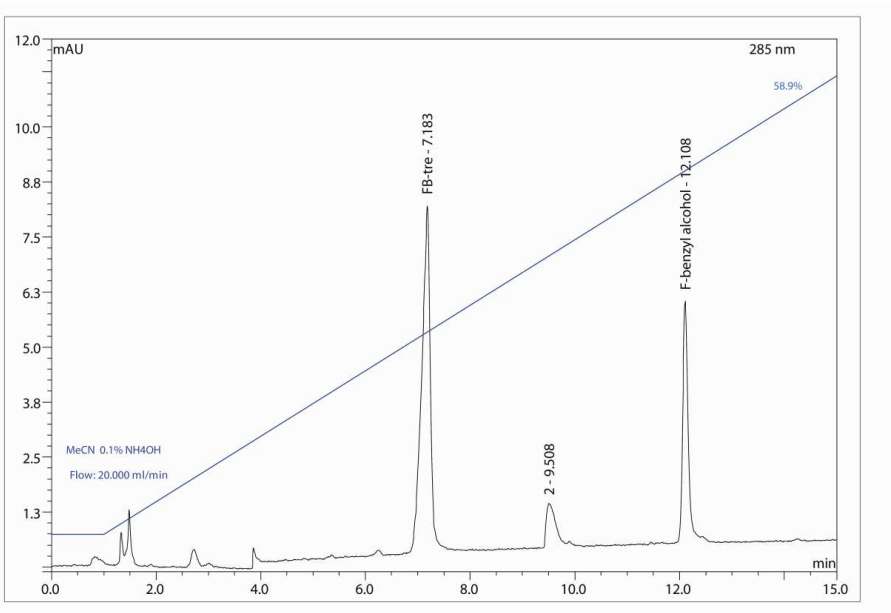
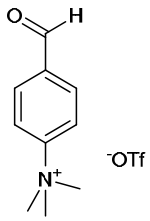


Figure 8.9. Preparative HPLC trace for the purification of 38. Conditions: Phenomenex Synergi Hydro C18 column (150 mm x 21.2 mm, 4 µm), eluting with water with 0.1 % NH_4OH (A) and acetonitrile (B) at a flow rate of 20 mL/min; Gradient: 0 min B=%0, 1 min B=0%, 15 min B=60%; Compound 38 elutes at 7.2 min (labeled FB-tre) while side product 4-fluorobenzyl alcohol elutes at 12.1 min.

8.2.9.1 Synthesis of ^{18}F-Trehalose 39

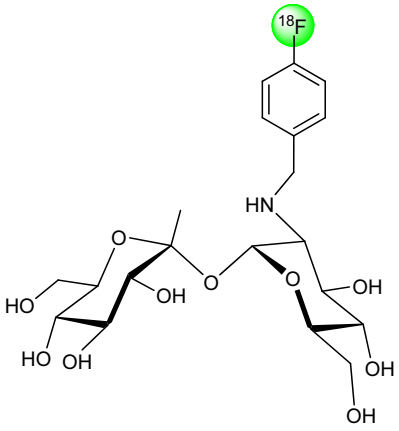
(4-Formyl-phenyl)-trimethyl-ammonium trifluoro-methanesulfonate (191)



Methyl trifluoro-methanesulfonate (0.61 mL, 4.5 mmol) was added to a solution of N,N-dimethylamino benzaldehyde **190** (0.75 g, 5 mmol) in dichloromethane (70 mL) at RT under argon. The dark green reaction mixture was stirred at RT for 24 h. Reaction was then concentrated to 20 mL by rotovap under reduced pressure and concentrated solution was added to ether (200 mL) and then chilled to 0 ˚C on ice. The resulting white precipitate was filtrated and washed extensively with diethyl ether (5 x 40 mL) to yield **191** (368 mg, 24%) as a white powdery solid.

^{1}H NMR (400 MHz, DMSO-*d*) δ ppm 3.38 (3 H, s, CH_3), 8.12,8.17, 8.19, 8.24 (dd, 4H, CHarom, J ¼ 8:79 Hz), 10.11 (s, 1H, CHO). ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 56.43 (CH_3), 121.73 (CHarom, ortho to –$N(CH3)^{3+}$), 130.92 (CHarom), 136.81 (Cquart, CHO), 151.07 (Cquart,C-$N(CH3)^{3+}$), 192.17 (CHO); MS *m/z* (ESI^+): 164.13 $[M+H]^+$.

Deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2-N-para-18-fluorobenzoyl-2-deoxy-α-D-glucopyranoside (39)



To one 13 x 100 test tube was added 100 µL Kryptofix (K222) solution (4.5 mg/100 µL), 60 µL K_2CO_3 solution (0.1 M in water), and $^{18}F^-$ (57.9 mCi) at 8:21 AM. Reaction was evaporated under stream of argon in 105 °C heat block and was azeotroped 3 times with 200 µL CH_3CN. To this was added the substrate 4-trimethylammonium benzaldehyde triflate in 300 µL CH_3CN and reaction was heated 5 min at 30 °C. Reaction was cooled to RT in a water bath and to it was added 8 µL HOAc, 100 µL solution of aminotrehalose (4 mg in 100 µL) and returned to 130 °C block for about 1 min at which time was added 100 µL of $NaCNBH_3$ solution (1 M in THF). Reaction was heated for a further 9 minutes upon which time it was cooled to RT in a water bath and diluted to ~12 mL with 0.1% NH_4OH and loaded onto 500 mg BONDELUT C18. The column was washed with 1 mL water and reaction mixture was eluted with 1 mL MeOH. Eluant was assayed 26.1 mCi at 8:56 AM and reaction mixture was concentrated under stream of argon. Purification was conducted on a semiprep scale, utilizing a Synergi HydroRP 10 x 100 mm and Elution with 10% ethanol 90% 0.1% NH_4OH. Product eluted a roughly 13 minutes. Radio TLC (water:isopropanol:ethyl acetate, 1:2:2) indicated the desired product (R_f 0.75). Reaction was neutralized with 1M HCl (30 µL) and filter sterilized through a 13 mm sterile filter to yield 3 mL ethanol solution of the desired compound (6.37 mCi, 22%). TLC (H_2O:*i*-PrOH:EtOAc, 1:2:2, v/v/v): R_f=0.75.

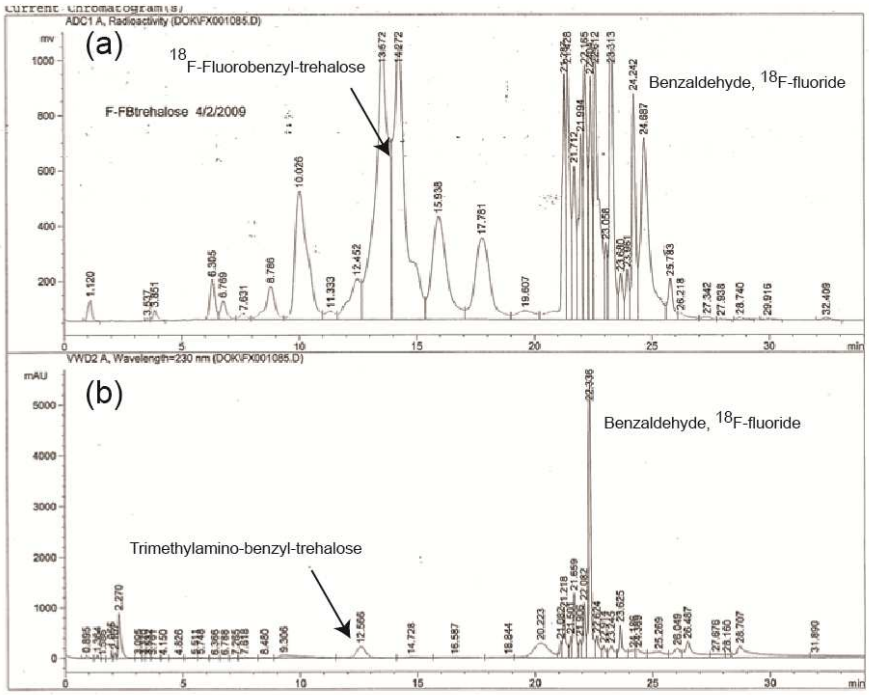


Figure 8.10. Prep HPLC traces of ^{18}FB-trehalose. Elution with 10% ethanol 90% with 0.1% NH4OH. (a) Radiochemical trace, peak for ^{18}F-Fluorobenzyl-trehalose 39 shows as a doublet because the detector saturates and then recovers detection of the signal. (b) UV trace shows elution of non-radioactive side product Trimethylamino-benzyl-trehalose prior to the target compound.

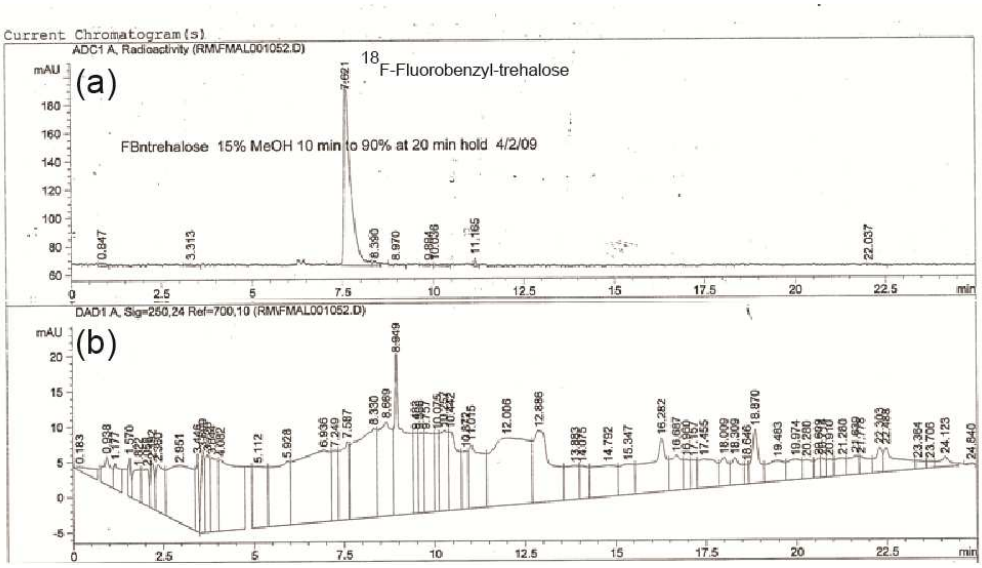


Figure 8.11. Analytical HPLC traces of purified FB-trehehalose 39 shows no significant, UV active or radioactive contaminants of 39. Elution with 15% MeOH at 10 min to 90% MeOH at 20 min with 0.1% NH_4OH (a) radiochemical trace. (b) UV trace.

Radiographic TLC was employed to identify the product as the desired compound as well as to determine how much of the desired compound was in the breakthrough of the C18 cartridge. None of the product is in the breakthrough, which had been a problem in earlier syntheses.

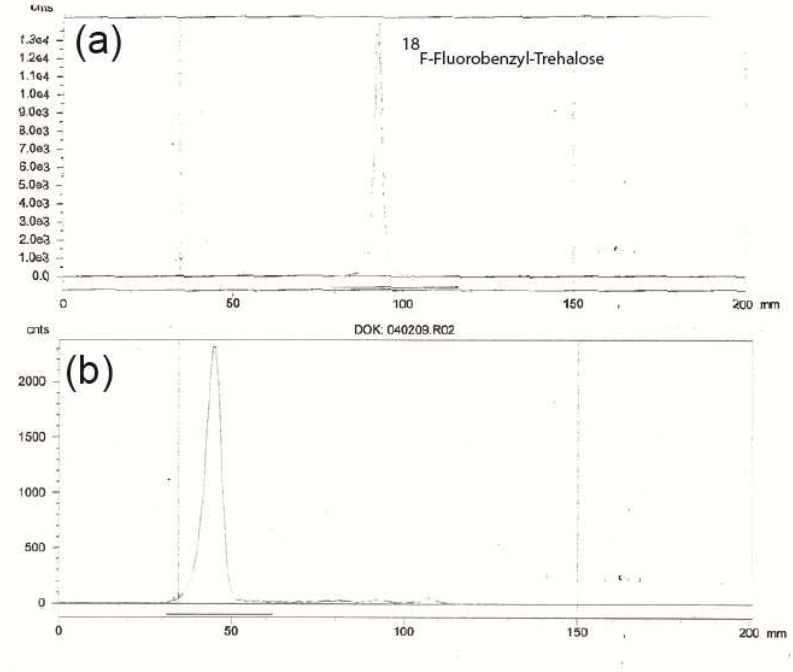
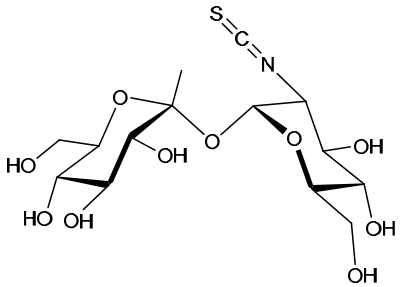


Figure 8.12. Radiographic TLC (H2O:i-PrOH:EtOAc, 1:2:2, v/v/v) (a) product 39 (b) breakthrough of Sep-Pak purification. Notably, no product is present in the breakthrough.

8.2.10 Synthesis of Quantum dot-trehalose 55

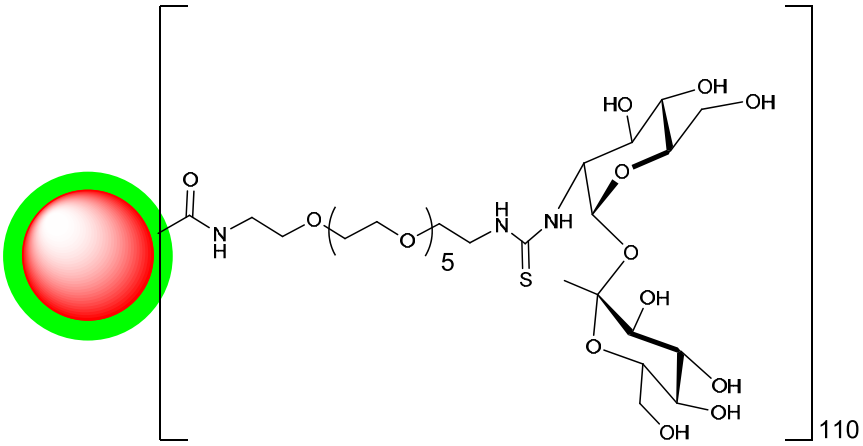
1-deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2-deoxy-2–isothiocyanato-α-D-glucopyranoside (194)



37 (2.5 mg, 0.007 mmol, 1 eq) was dissolved in 75 mM $NaHCO_3$ buffer pH 9 (200 µl). To this was added thiophosgene as a solution (20 µl of thiophosgene into 1 mL of chloroform). 100 µl of thiophosgene solution was added to reaction (3 mg, 0.26, 4.2 eq). The resulting biphasic mixture was stirred at room temperature for 3 hours upon which TLC (EtOH:NH_4OH: H_2O, 5:3:1) showed

complete consumption of starting material (R_f 0.2) and conversion to a single product (R_f 0.65). Excess thiophosgene and chloroform were removed *in vacuo* and crude product was used without further purification. TLC (EtOH:NH_4OH: H_2O, 5:3:1, v/v): R_f=0.65;

Trehalose Quantum Dots **(55)**



An 8 µM solution of CdSe-ZnS (50 µl) core-shell quantum dots (emission λ_{max} 655nm) in borate buffer at pH 8 (Invitrogen) was buffer exchanged into water by repeated (x5) centrifugal filtration through a 10 kDa cutoff spin filter. Dots were then resuspended in water. 50 µl of this 8 µM quantum dot solution was then added to the solution of **194**, and the total volume made up to 1.0 mL with 75 mM $NaHCO_3$ buffer at pH 9.0 (pH electrode). Reaction mixture was shaken at 4 °C for 14 hours. Excess sugar and salt was removed from the reaction mixture by size exclusion chromatography (PD 10 column, Amersham) with water as the mobile phase. The quantum dot solution was concentrated to 1 mL, using 10 Kda spin filter, and the concentration determined using previously reported procedures[48] to be 0.44 µM (ε_{350} = 3880000 $M^{-1}cm^{-1}$). The modification of the quantum dots was confirmed by agarose gel electrophoresis.

The carbohydrate loading on the quantum dots was determined using the phenol sulphuric acid method. An aliquot of the quantum dot solution (50 µl) was treated with concentrated sulphuric acid (75 µl) and aqueous phenol (5% w/w, 10 µl) and heated to 90 °C. After 5 minutes the sample was cooled to room temperature and A_{490} measured, referenced to a solution of carbohydrate modified quantum dots and acid. The concentration of trehalose was determined by comparison to a standardized curve (**Figure 8.13**). The carbohydrate content per dot was calculated from the ratio of trehalose concentration to the concentration of ZnS-CdSe quantum dots and was found to be ~110 sugars/dot.

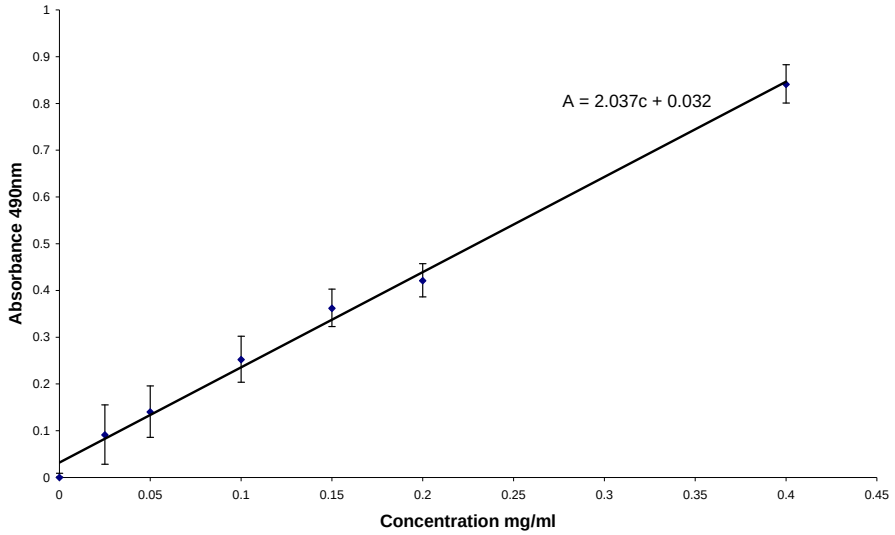
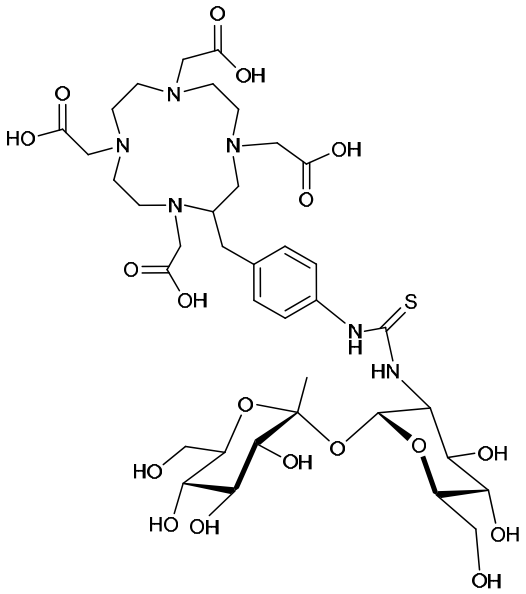


Figure 8.13. Standardized curve for determining trehalose concentration on modified quantum dots using the phenol-sulphuric acid method. Shown is the absorbance at 490 nm after reaction with phenol and sulphuric acid against standardized concentrations of trehalose.

8.2.11 Synthesis of DOTA-trehalose 57

1-Deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2-N-isothiourea-benzyl-(1,4,7,10-tetraazacyclododecane-N,N',N'',N'''-tetraaceticacid)-2-deoxy-α-D-glucopyranoside trifluoroacetic acid salt (57)

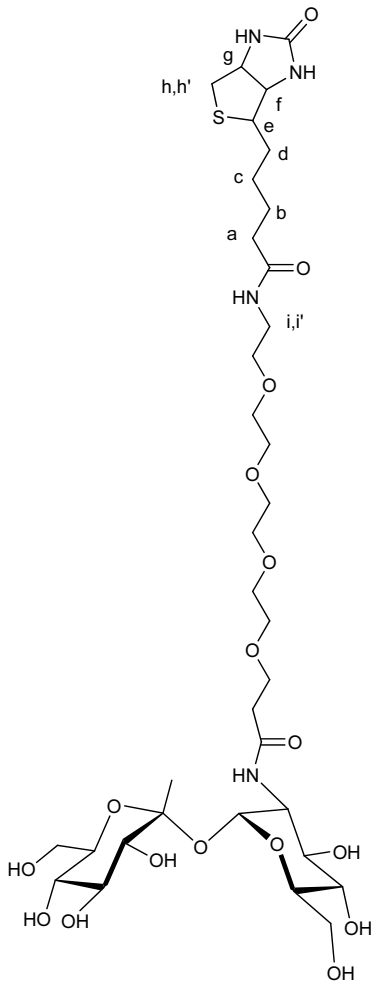


37 (8.3 mg, 0.023 mmol, 1 eq) and **195** (13.4 mg, 003 mmol, 1.1 eq) (Macrocyclics, B-205) were dissolved in sodium carbonate buffer (1 mL, 75 mM, pH 9). To this was added acetonitrile (0.5 mL) and the reaction mixture was stirred at 50 °C for 5 hours followed by 72 hours at RT. Upon completion TLC (5 EtOH:NH_4OH:H_2O, 5:3:1, v/v/v) revealed complete disappearance of the starting sugar (R_f 0.5) and Dota (R_f 0.9) and appearance of a single product (R_f 0.6). Reaction was neutralized with NH_4Cl solution until neutral pH was obtained and the resulting mixture was lyophilized to give a brownish solid. Reaction mixture was purified by cation and anion exchange chromatography. First, mixture was loaded onto Dowex 550A (OH- form) anion exchange resin (ca 100 mg) and resin was washed with 10 mL water and then eluted with 10 mL saturated NaCl solution. Eluant was then loaded onto DOWEX 50WX8 (H^+ form) cation exchange resin (ca 100 mg) and resin was washed with water (5 mL), NH_4OH (0.1%) followed by elution in NH_4OH (5%). Lyophilization yielded the desired product as a brownish amorphous solid (3.7 mg, 20%).

TLC (EtOH:NH_4OH:H_2O, 5:3:1, v/v/v): R_f = 0.6; $[\alpha]_D^{25}$ = 21.6 (c = 0.43, water); ^{1}H NMR (500 MHz, DEUTERIUM OXIDE) δ ppm 1.44 (3 H, s), 1.86 (1 H, d, *J* = 12.3 Hz, 1 x C<u>H</u>$_2$ Dota), 2.00 (1 H, d, *J* = 12.6 Hz, 1 x C<u>H</u>$_2$ Dota), 2.05 - 2.36 (7 H, m, 6 x C<u>H</u>$_2$ DOTA, 1x C-H Benzyl), 2.45 (1 H, d, *J* = 16.1 Hz, 1 x C<u>H</u>$_2$ Dota), 2.49 - 2.89 (7 H, m, 7 x C<u>H</u>$_2$ DOTA), 2.87 (1 H, d, *J* = 16.1 Hz, 1 x C<u>H</u>$_2$ Dota), 2.96 - 3.29 (8 H, m, C-3', 6 x C<u>H</u>$_2$ DOTA, 1x C-H Benzyl), 3.38 - 3.54 (7 H, m, H-4; H-4, H-5', 4 x C<u>H</u>$_2$ DOTA), 3.63 (1 H, ddd, $J_{5,6}$ = 10.5 Hz, $J_{6,7a}$ = 5.8 Hz, $J_{6,7b}$ = 3.5 Hz, H-6'), 3.67 - 3.81 (5 H, m, H-6a, H-6b, H-7a, H7-b, H-5), 3.89 (1 H, at, *J* = 9.8 Hz, H-3), 4.31 (1 H, dd, $J_{1,2}$ = 3.2 Hz, $J_{2,3}$ = 9.1 Hz, H-2), 5.59 (1 H, d, $J_{1,2}$ = 3.2 Hz, H-1), 7.16 (2 H, d, *J* = 7.9 Hz, H-Meta), 7.23 (2 H, m, *J* = 8.8 Hz, H-Ortho); ^{13}C NMR (126 MHz, DEUTERIUM OXIDE) δ ppm 22.8 (C-1'), 30.3 (CH_2 DOTA benzyl), 42.6 (CH_2 DOTA), 48.2 (CH_2 DOTA), 48.7 (CH_2 DOTA), 51.0 (CH_2 DOTA), 52.5 (CH_2 DOTA), 52.9 (CH_2 DOTA), 53.4 (CH_2 DOTA), 55.1 (CH_2 DOTA), 55.7 (CH_2 DOTA), 58.7 (CH_2 DOTA), 58.8 (C-2), 59.3 (CH_2 DOTA), 60.0 (C-6), 60.5 (C-7'), 61.3 (CH-Dota) 68.9 (C-4'), 70.4 (C-4), 70.4 (C-3), 71.9 (C-5), 72.7 (C-6'), 72.7 (C-5'), 76.4 (C-3'), 89.8 (C-1), 101.4 (C-2'), 101.6 (C isothiourea), 126.3 (C-H-meta), 129.4 (C-H ortho), 179.7 (C=O), 179.8 (C=O), 180.1 (C=O), 180.1 (C=O); HRMS *m/z* (ESI^+): [M-H^-] calcd. for $C_{37}H_{57}N_6O_{18}S^-$, 905.3450; found, 905.3449.

8.2.12 Synthesis of biotin-trehalose, 54

1-Deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2-N-carboxy-tetra-ethyleneglycol-N-biotin-2-deoxy-α-D-glucopyranoside (54)



37 (26.2 mg, 0.07 mmol, 1.5 eq) and **192** (29.1 mg, 0.049 mmol, 1 eq) were dissolved in anhydrous DMF (2 mL) and to this was added Na_2CO_3 (1 mg, 0.0095 mmol, 0.2 eq). Reaction was conducted under Argon and heated to 40 °C for 24 hours and stirred for a further 4 days at RT upon which time product (R_f 0.3) was detected by TLC TLC (EtOH:NH_4OH:H_2O, 5:3:1, v/v/v) with near complete disappearance of biotin NHS starting material (R_f 0.8) and starting amine (R_f 0.4) and appearance of a new product spot (R_f 0.85). Reaction mixture was purified by HPLC with a Phenomenex Synergi Hydro C18 column (150 mm x 21.2 mm, 4 μm) and a acetonitrile gradient with 0.1% aqueous NH_4OH. Lyophilization yielded the desired product as a white solid (15.4 mg, 38 %).

TLC (EtOH:NH_4OH:H_2O, 5:3:1, v/v/v): R_f = 0.85; $[\alpha]_D^{25}$ = 44.3 (c = 0.23, water); ^{1}H NMR (500 MHz, DEUTERIUM OXIDE) δ ppm 1.33 (1 H, quin, $J_{c,d} = J_{b,c}$ = 7.5 Hz, H-c), 1.44 (3 H, s, H-1'),

1.46 - 1.70 (2 H, m, H-b, H-d), 2.20 (1 H, t, $J_{a,b}$ = 7.3 Hz, H-a), 2.41 (1 H, t, J = 6.0 Hz, H-i), 2.44 (1 H, at, J = 5.4 Hz, H-j), 2.57 (1 H, dd, J = 7.7, J = 5.8 Hz, H-i'), 2.60 (1 H, td, J = 7.8, J = 6.5 Hz, H-j'), 2.43 (1 H, dt, $J_{i,i'}$ = 15.1 Hz, $J_{i,CH2}$ = $J_{i,CH2}$ =5.7 Hz, H-i), 2.59 (1 H, ddd, J = 15.0 Hz, $J_{i,CH2}$ = 7.9 Hz, $J_{i,CH2}$ = 6.5 Hz, H-i'),2.70 (1 H, d, J = 12.9 Hz, H-h), 2.91 (1 H, dd, $J_{h,h'}$ = 13.2 Hz, $J_{g,h'}$ = 5.0 Hz, H-h'), 3.23 (1 H, d, $J_{3,4}$ = 9.8 Hz, H-3'), 3.26 (1 H, dt, $J_{d',e}$ = 9.1 Hz, $J_{e,f}$ = 4.7 Hz, $J_{d,e}$ = 4.7 Hz, H-e), 3.31 (2 H, at, $J_{CH,CH2}$ = 5.2 Hz, PEG-C<u>H</u>$_2$), 3.38 (1 H, at, $J_{4,5}$ = $J_{5,6}$ = 9.8 Hz, H-5'), 3.41 (1 H, dd, $J_{4,5}$ = 9.8 Hz, $J_{3,4}$ = 8.4 Hz, H-4), 3.45 (1 H, ddd, $J_{5,6}$ = 10.1 Hz, $J_{6,7a}$ = 4.0 Hz, $J_{6,7b}$ = 3.2 Hz, H-6'), 3.55 (2 H, t, $J_{CH,CH2}$ = 5.4 Hz, PEG-CH$_2$), 3.59 - 3.76 (18 H, m, PEG-CH$_2$, H-6^a, H-7^a,H-7b, H-4'), 3.65 (1 H, dd, $J_{6a,6b}$ = 12.6 Hz, $J_{5,6b}$ = 2.8 Hz, H-6b), 3.79 (1 H, ddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 4.7 Hz, $J_{5,6b}$ = 2.2 Hz, H-5), 3.84 (1 H, dd, $J_{2,3}$ = 10.7 Hz, $J_{1,2}$ = 3.2 Hz, H-2), 3.85 (1 H, at, $J_{2,3}$ = $J_{4,3}$ = 9.7 Hz, H-3), 4.34 (1 H, dd, $J_{f,g}$ = 7.9 Hz, $J_{e,f}$ = 4.4 Hz, H-f), 4.53 (1 H, dd, $J_{f,g}$ = 7.9, $J_{g,h}$ = 5.0 Hz, H-g), 5.25 (1 H, d, $J_{1,2}$ = 2.8 Hz, H-1); ^{13}C NMR (126 MHz, DEUTERIUM OXIDE) δ ppm 22.7 (C-1), 25.1 (C-b), 27.65 (C-d), 27.83 (C-c), 35.4 (C-a), 35.6 (C-i), 38.8 (Peg-<u>C</u>H$_2$), 39.6 (C-h), 54.3 (C-2), 55.3 (C-e), 60.1 (C-7' or C-g), 60.2 (C-7' or C-g), 60.4 (C-6), 62.1 (C-f), 66.3 (Peg-<u>C</u>H$_2$), 68.8 (Peg-<u>C</u>H$_2$), 69.3 (Peg-<u>C</u>H$_2$), 69.4 (C-5'), 69.5 (Peg-<u>C</u>H$_2$), 69.5 (2 C, Peg-<u>C</u>H$_2$), 69.5 (Peg-<u>C</u>H$_2$), 69.61 (Peg-<u>C</u>H$_2$), 70.0 (C-3), 70.27 (C-4), 71.8 (C-5), 72.71 (C-4'), 72.9 (C-6'), 76.37 (C-3'), 90.1 (C-1), 101.3 (C-1'), 165.3 (C=O), 174.0 (C=O), 176.9 (C=O); MS *m/z* (ESI$^-$) 829.4 (M+H); HRMS *m/z* (ESI$^+$): $[M+Na]^+$ calcd. for $C_{34}H_{60}N_4O_{17}SNa^+$, 851.3572; found, 851.3584.

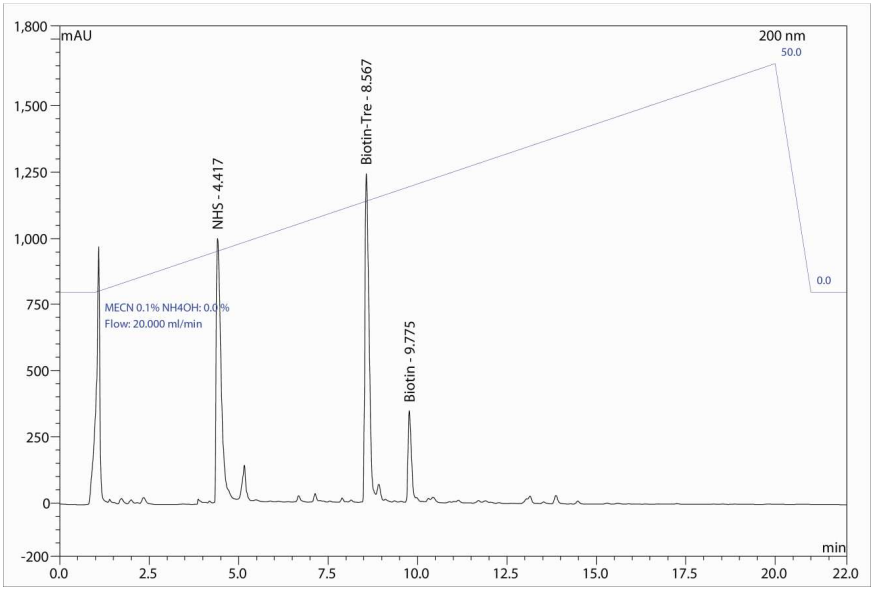
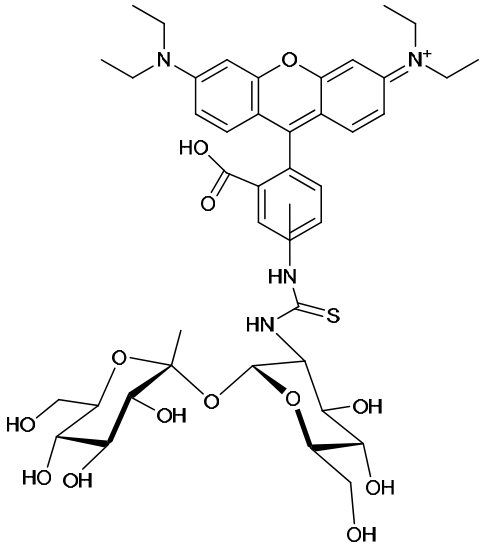


Figure 8.14. Preparative HPLC trace for the purification of 54. Conditions: Phenomenex Synergi Hydro C18 column (150 mm x 21.2 mm, 4 µm), eluting with water with 0.1 % NH_4OH (A) and acetonitrile (B) at a flow rate of 20 mL/min; Gradient: 0 min B=%0, 1 min B=0%, 20 min B=50%; Compound 54 elutes at 8.7 min (labeled Biotin-Tre) while side product hydrolyzed Biotin elutes at 9.8 min.

8.2.13 Synthesis of rhodamine-trehalose, 56 and 201

1-Deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2-N-isothiourea-5-(and-6)-rhodamine-B-2-deoxy-α-D-glucopyranoside (201)



37 (50 mg, 0.14 mmol, 1.5 eq) and rhodamine B isothiocyante (RBITC, **200**) (88 mg, 0.016 mmol, 1.16 eq) were dissolved in sodium carbonate buffer (1 mL, 75 mM, pH 9). To this was added acetonitrile (0.5 mL). Reaction was stirred at RT for 3 days upon which time product (R_f 0.1) was detected by TLC (EtOAc:MeOH, 1:1, v/v) with near complete disappearance of rhodamine starting material (R_f 0.4).Reaction mixture was purified by HPLC with a Phenomenex Synergi Hydro C18 column (150 mm x 21.2 mm, 4 μm) and a acetonitrile gradient with 0.1% aqueous NH_4OH or 0.1% aqueous TFA. Lyophilization yielded the desired product as a pink solid mixed with impurities. Further purifications are under investigation.

TLC (EtOAc:MeOH, 1:1, v/v): R_f = 0.1; HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{42}H_{55}N_4O_{13}SNa^+$, 877.3300; found, 877.3305.

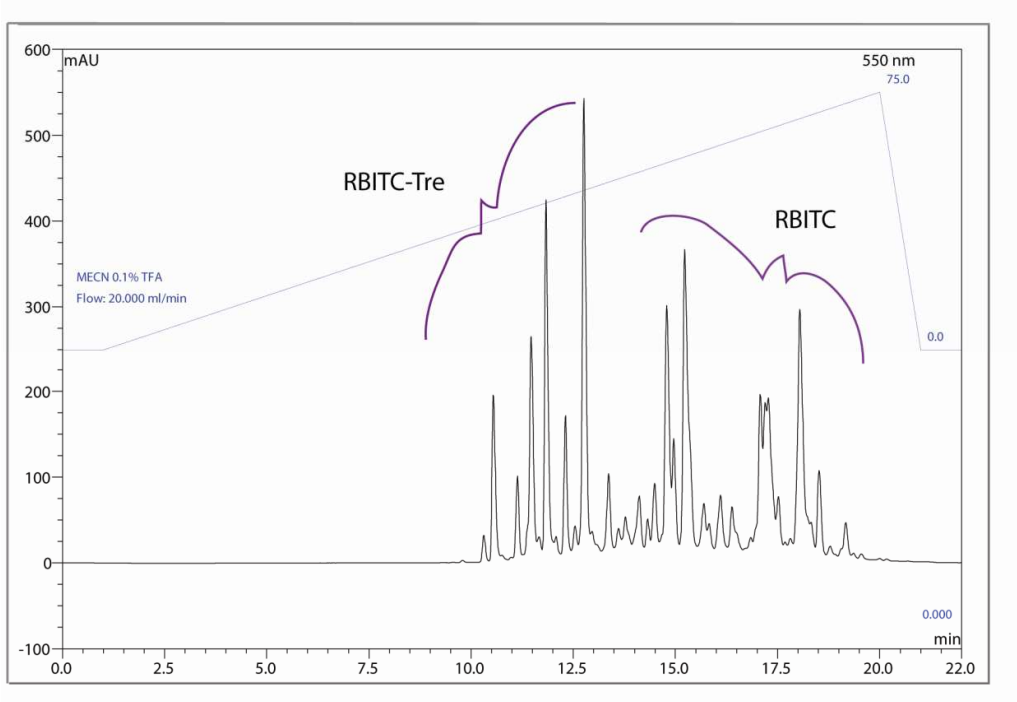
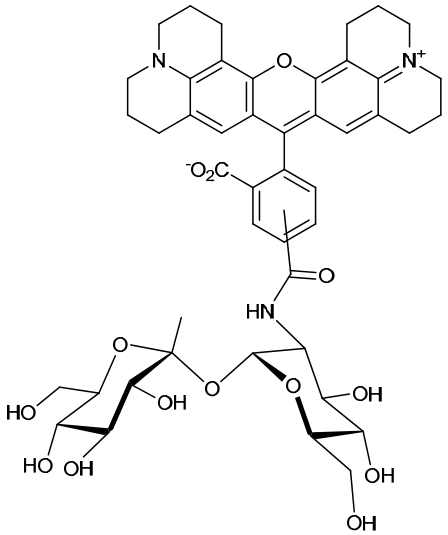


Figure 8.15. Preparative HPLC trace for the purification of 201. Conditions: Phenomenex Synergi Hydro C18 column (150 mm x 21.2 mm, 4 µm), eluting with water with 0.1 % NH4OH (A) and acetonitrile (B) at a flow rate of 20 mL/min; Gradient: 0 min B=%0, 1 min B=0%, 20 min B=75%; Compound 201 elutes at 10-12.5 min (labeled RBITC-Tre) while parent compound (RBITC) elutes at 15-20 min.

1-Deoxy-α-D-gluco-hept-2-ulopyranosyl-(2→1)-2-N-5-(and-6)-Carboxyrhodamine-2-deoxy-α-D-glucopyranoside (56)



37 (44.5 mg, 0.123 mmol, 1.1 eq) and **202** (68 mg, 0.107 mmol, 1 eq) were dissolved in anhydrous DMF (1 mL) and to this was added N,N-diisopropylethylamine (10 µL, 0.06 mmol, 0.6 eq). Reaction

was stirred 24 h at 50 °C upon which time product (R_f 0.25) was detected by TLC (MeOH:EtOAc, 2:1) with near complete disappearance of rhodamine material (R_f 0.6) and starting amine (R_f 0.05). Reaction mixture was purified by HPLC with a Phenomenex Synergi Hydro C18 column (150 mm x 21.2 mm, 4 μm) and an acetonitrile gradient with 0.1% aqueous NH_4OH. Lyophilization yielded the desired product as a pink solid (30 mg, 32 %). Product was obtained as a mixture of the two isomers present in **202**. The major isomer was assigned by NMR.

TLC (EtOH:NH_4OH:H_2O, 5:3:1, v/v/v): R_f = 0.9; $[\alpha]_D^{25}$ = 53.3 (c = 0.18, water); ^{1}H NMR (500 MHz, METHANOL-*d4*) δ ppm 1.54 (3 H, s, C$\underline{H}_3$), 1.95 (4 H, quin, $J_{CH2,CH2}$ = 4.4 Hz, 2 x Rho-C$\underline{H}_2$), 2.11 (4 H, quin, $J_{CH2,CH2}$ = 5.7 Hz, 2 x Rho-C$\underline{H}_2$), 2.72 (4 H, sxt, $J_{CH2,CH2}$ = 6.0 Hz, 2 x Rho-C$\underline{H}_2$), 3.06 - 3.11 (4 H, m, 2 x Rho-C$\underline{H}_2$), 3.15 (1 H, d, $J_{3,4}$ = 9.8 Hz, H-3'), 3.26 (1 H, m, H-5'), 3.38 (1 H, dd, $J_{6a,6b}$ = 10.8 Hz, $J_{5,6a}$ = 4.9 Hz, H-6a), 3.39 (1 H, dd, $J_{6a,6b}$ = 10.8, $J_{5,6b}$ = 2.8 Hz, H-6b), 3.45 - 3.53 (6 H, m, H=6', H-4, 2 x Rho-C$\underline{H}_2$), (4 H, at, $J_{CH2,CH2}$ = 4.9 Hz, 2 x Rho-C$\underline{H}_2$), 3.75 (1 H, dd, $J_{7a,7b}$ =12.0 Hz, $J_{6,7a}$ = 6.3 Hz, H-7a'), 3.76 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.1 Hz, H-4'), 3.82 (1 H, dd, $J_{7a,7b}$ =11.8, $J_{6,7b}$ = 2.8 Hz, H-7b'), 3.91 (1 H, ddd, $J_{4,5}$ = 8.2 Hz, $J_{5,6a}$ = 5.0 Hz, $J_{5,6b}$ = 2.7 Hz, H-5), (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.5 Hz, H-3), 4.09 (1 H, dd, $J_{2,3}$ = 8.7 Hz, $J_{1,2}$ = 3.2 Hz, H-2), 5.52 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1), 6.28 (1 H, s), 6.78 (1 H, s, Ar-H), 6.82 (1 H, s, Ar-H), 7.71 (1 H, d, $J_{CH,\,CH}$ = 1.3 Hz, Ar-H), 8.08 (1 H, dd, $J_{CH,\,CH}$ = 8.2, 2.5 Hz, Ar-H), 8.18 (1 H, dd, $J_{CH,\,CH}$ = 8.2, 1.6 Hz, Ar-H); ^{13}C NMR (126 MHz, METHANOL-*d4*) δ ppm 21.0 (4 C, 4 x Rho-CH_2), 21.9 (Rho-CH_2), 21.9 (Rho-CH_2), 22.0 (Rho-CH_2), 23.9 (CH_3), 28.4 (Rho-CH_2), 28.5 (Rho-CH_2), 48.5 (Rho-CH_2), 51.3 (2 C, 2 x Rho-CH_2), 51.8 (2 C, 2 x Rho-CH_2), 57.1 (C-2), 61.9 (C-6), 62.5 (C-7'), 71.2 (C-3), 71.7 (C-5'), 72.6 (C-4), 73.7 (C-5), 74.1 (C-4'), 74.9 (C-6'), 78.4 (C-3'), 91.6 (C-1), 102.6 (C-2'), 106.2 (Rho C-N), 114.5 (Rho C-N), 124.8 (Rho Ar-C), 124.8 (Rho Ar-C), 128.0 (Rho Ar-C), 128.0 (Rho Ar-C),128.1 (Rho Ar-C), 129.6 (Rho Ar-C), 130.2 (Rho Ar-C), 130.7 (Rho Ar-C), 130.7 (Rho Ar-C), 134.1 (Rho Ar-C), 135.6 (Rho Ar-C), 152.3 (Rho Ar-C), 152.3 (Rho Ar-C), 152.3 (Rho Ar-C), 153.6 (Rho Ar-C), 153.7 (Rho Ar-C), 169.5 (Rho Ar-C), 169 (C=O), 172 (C=O); HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{46}H_{55}N_3O_{14}\,Na^+$, 894.3420; found, 894.3416.

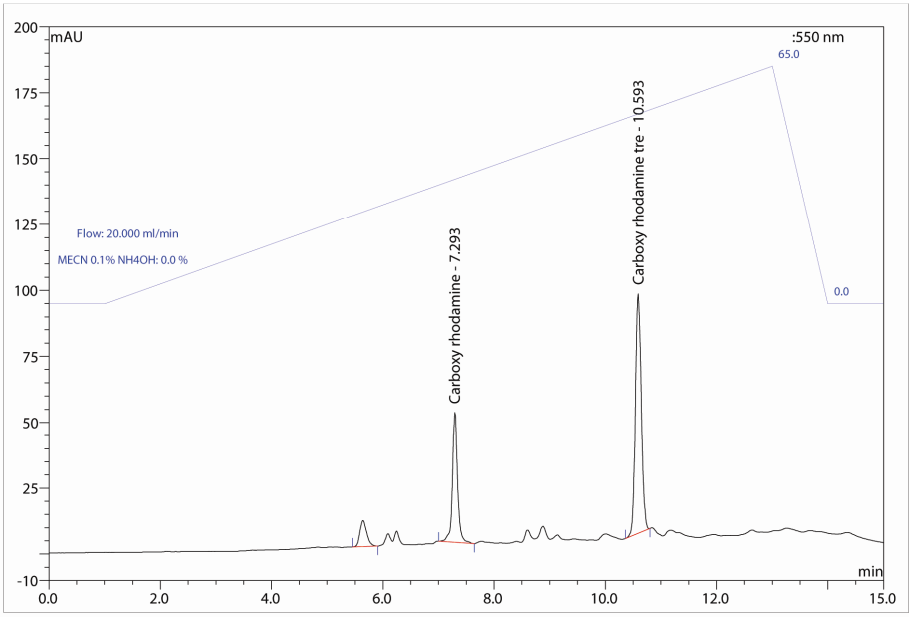
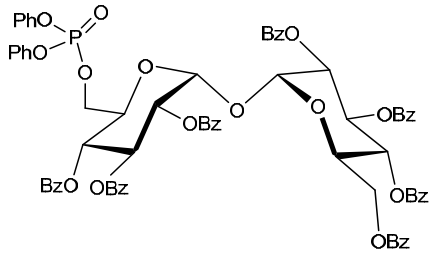


Figure 8.16. Preparative HPLC trace for the purification of 56. Conditions: Phenomenex Synergi Hydro C18 column (150 mm x 21.2 mm, 4 µm), eluting with water with 0.1 % NH4OH (A) and acetonitrile (B) at a flow rate of 20 mL/min; Gradient: 0 min B=%0, 1 min B=0%, 15 min B=65%; Compound 56 elutes at 10.59 min (labeled RBITC-Tre) while Carboxyrhodamine elutes at 7.29 min.

8.2.14 Synthesis of trehalose-6-Azide 48[ii]

***6-O-diphenoxyphosphoryl-2,2',3,3',4,4',6'-hepta-O-benzoyl-α,α-D-trehalose* (252)**



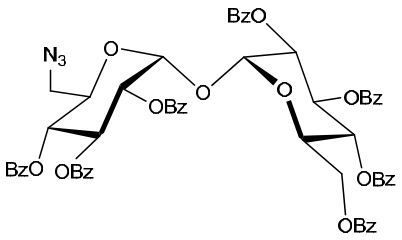
To a solution of **251** (100 mg, 0.18 mmol, 1 eq) in dry pyridine (5 mL) at RT was added dropwise benzoyl chloride (0.22 mL, 1.93 mmol, 11 eq). The reaction mixture was stirred for 15 hours after which TLC (petrol:EtOAc, 1:1) showed full consumption of starting material (R_f 0) and the formation of a single product (R_f 0.5). The reaction was quenched with methanol (1 mL) and the mixture was concentrated under reduced pressure. The resultant residue was partitioned between ethyl acetate (25 mL) and water (25 mL) and the aqueous layer was extracted with ethyl acetate (2 x 25 mL). The combined organics were washed with 1M HCl (3 x 25 mL), saturated $NaHCO_3$ (25 mL), brine (25

[ii] Synthesis conducted with Mitul K. Patel

mL), dried over $MgSO_4$ and concentrated *in vacuo*. Silica gel chromatography (petrol:EtOAc, 1:1) yielded the target compound as a colorless oil (0.091g, 40%).

TLC (petrol:EtOAc, 1:1, v/v): R_f=0.50; $[\alpha]_D^{20}$ +148.2 (c = 0.46, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 3.78 (1 H, ddd, $J_{6a,6b}$ = 11.4 Hz, $J_{P,6a}$ = 6.9 Hz, $J_{5,6a}$ = 4.0 Hz, H-6a), 3.87 (1 H, m, H-6b), 3.91 (1 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6a}$ = 4.6 Hz, H-6'a), 4.03 (1 H, dd, $J_{6a,6b}$ = 12.6 Hz, $J_{5,6b}$ = 2.7 Hz, H-6'b), 4.14 (1 H, m, H-5), 4.28 (1 H, ddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 4.3 Hz, $J_{5,6b}$ = 2.9Hz, H-5'), 5.31 (1 H, dd, $J_{2,3}$ = 10.2 Hz, $J_{1,2}$ = 3.8 Hz, H-2), 5.48 (1 H, dd, $J_{2,3}$ = 10.2 Hz, $J_{1,2}$ = 3.8 Hz, H-2'), 5.58 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.9 Hz, H-4), 5.62 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1), 5.65 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1'), 5.69 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.9 Hz, H-4'), 6.24 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.5 Hz, H-3), 6.29 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.9 Hz, H-3'), 7.12-7.60 (31 H, m, Ar-H), 7.80, 7.81, 8.08 (3 x 2 H, 3 x d, $J_{ortho\text{-}meta}$ = 7.2 Hz, Bz Ar-H_{ortho}), 7.94 (4 H, d, $J_{ortho\text{-}meta}$ = 8.2 Hz, Bz Ar-H_{ortho}), 8.00, 8.12 (2 x 2 H, 2 x d, $J_{ortho\text{-}meta}$ = 7.3 Hz, Bz Ar-H_{ortho}); ^{13}C NMR (101 MHz, $CDCl_3$): δ 61.9 (C-6'), 65.6 (d, $J_{P,6}$ = 6.4 Hz, C-6), 67.7 (C-4'), 68.5 (C-4), 68.6 (C-5'), 68.7 (d, $J_{P,5}$ = 6.4 Hz, C-5), 70.3 (2C, C-3, C-3'), 71.0 (C-2'), 71.2 (C-2), 92.4 (C-1'), 92.4 (C-1), 120.1 (d, $J_{P,C}$ = 4.8 Hz, P-OPh Ar-C_{ortho}), 125.4 (d, $J_{P,C}$ = 7.2 Hz, P-OPh Ar-C_{para}), 128.4-129.9 (m, ArC), 133.2, 133.3, 133.3, 133.5, 133.6, 133.7, 133.8 (7 x Bz Ar-C_{para}), 150.3 (d, $J_{P,C}$ = 7.2 Hz, P-OPh Ar-C_{ipso}), 150.4 (d, $J_{P,C}$ = 6.4 Hz, P-OPh Ar-C_{ipso}), 164.6, 164.9, 165.2, 165.2, 165.6, 165.6, 165.9 (7 x C=O); IR (thin film) 1640 (C=O) cm^{-1}; MS *m/z* (ESI^+): 1325.3$[M+Na]^+$; Isotopic distribution: species observed $[M+Na]^+$, peaks observed 1325.27 (100%), 1326.28 (74%), 1327.28 (29%), 1328.29 (8%), 1329.29 (2%) peaks calculated 1325.32 (100%), 1326.32 (80%), 1327.32 (36%), 1328.32 (12%), 1329.33 (3%).

***6-azido-2,2',3,3',4,4',6'-hepta-O-benzoyl-6-deoxy-α,α-D-trehalose* (253)**

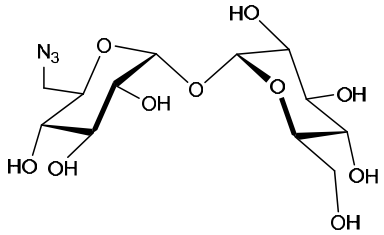


A solution of **252** (90 mg, 0.069 mmol, 1 equi) in DMF (3 mL) was heated to 90 °C with sodium azide (9 mg, 0.14 mmol, 2 equi). The reaction mixture was maintained at this temperature with aggressive stirring for 18 hours after which TLC (petrol:EtOAc, 1:1, v/v) showed complete consumption of starting material (R_f 0.5) and the formation of a product (R_f 0.4). The reaction mixture was concentrated *in vacuo* and the residue partitioned between ethyl acetate (30 mL) and water (20 mL). The aqueous layer was extracted with ethyl acetate (2 x 20 mL) and the combined organics washed with brine (3 x 25 mL), dried over $MgSO_4$ and the solvent removed *in vacuo*. The compound

was purified by silica gel chromatography (petrol:EtOAc, 3:2, v/v) to afford the desired compound as a colorless oil (52 mg, 69%).

TLC (petrol:EtOAc, 1:1, v/v): R_f = 0.40; $[\alpha]_D^{17}$ +210.4 (c = 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 2.89-2.91 (2 H, m, H-6a, H-6b), 3.88 (1 H, dd, $J_{6a,6b}$ = 12.5 Hz, $J_{5,6a}$ = 4.7 Hz, H-6'a), 4.03 (1 H, dd, $J_{6a,6b}$ = 12.6 Hz, $J_{5,6b}$ = 2.5 Hz, H-6'b), 4.14 (1 H, ddd, $J_{4,5}$ = 9.8 Hz, $J_{5,6a}$ = 7.6 Hz, $J_{5,6b}$ = 3.0 Hz, H-5), 4.32 (1 H, ddd, $J_{4,5}$ = 10.4 Hz, $J_{5,6a}$ = 4.3 Hz, $J_{5,6b}$ = 2.8 Hz, H-5'), 5.50 (2 H, dd, $J_{2,3}$ = $J_{2,3}$ = 10.1 Hz, $J_{1,2}$ = $J_{1,2}$ = 3.8 Hz, H-2, H-2'), 5.57 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.9 Hz, H-4), 5.70 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 10.0 Hz, H-4'), 5.76 (1 H, d, $J_{1,2}$ = 4.0 Hz, H-1), 5.78 (1 H, d, $J_{1,2}$ = 4.0 Hz, H-1'), 6.26 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.9 Hz, H-4), 6.30 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 10.0 Hz, H-3'), 7.32-7.60 (21 H, m, Ar-H), 7.83, 7.85 (2 x 2 H, 2 x dd, $J_{ortho\text{-}meta}$ = 8.2 Hz, $J_{ortho\text{-}ortho}$ = 1.3 Hz, Ar-H_{ortho}), 7.93, 7.94 (2 x 2 H, 2 x dd, $J_{ortho\text{-}meta}$ = 8.7 Hz, $J_{ortho\text{-}ortho}$ = 1.5 Hz, Ar-H_{ortho}), 7.96 (2 H, dd, $J_{ortho\text{-}meta}$ = 8.2 Hz, $J_{ortho\text{-}ortho}$ = 1.0 Hz, Ar-H_{ortho}), 8.08, 8.14 (2 x 2 H, 2 x dd, $J_{ortho\text{-}meta}$ = 7.4 Hz, $J_{ortho\text{-}ortho}$ = 1.3 Hz, Ar-H_{ortho}); ^{13}C NMR (101 MHz, $CDCl_3$): δ 50.0 (C-6), 61.9 (C-6'), 68.5 (C-5'), 68.6 (C-4'), 69.1 (C-4), 69.7 (C-5), 70.1 (C-3'), 70.2 (C-3), 71.1 (C-2'), 71.3 (C-2), 92.6 (C-1'), 92.7 (C-1), 128.4, 128.4, 128.7, 128.7 (4 x as, Ar-C_{meta}), 128.5, 128.7, 128.7, 128.8 129.0, 129.0, 129.4 (Ar-C_{ipso}), 129.8, 129.9, 129.9, 129.9 (4 x as, Ar-C_{ortho}), 133.1, 133.3, 133.3, 133.5, 133.6, 133.7, 133.8 (Ar-C_{para}), 164.9, 164.9, 165.2, 165.3, 165.5, 165.6, 165.8 (C=O); IR (thin film) 1732 (C=O), 2107 (N_3) cm^{-1}; MS (m/z) (ESI$^+$): 1118.4$[M+Na]^+$; Isotopic distribution: peaks observed: $[M+Na]^+$ peaks 1118.30 (100%), 11190.31 (66%), 1120.31 (24%), 1121.31 (7%), 1122.32 (2%) peaks calculated 1118.30 (100%), 1119.30 (68%), 1120.30 (26%), 1121.30 (7%), 1122.31 (2%).

***6-azido-6-deoxy-α,α-D-trehalose* (253)**[49]

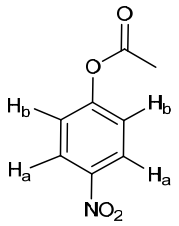


To a stirring solution of **252** (52 mg, 0.047 mmol, 1 equi) in methanol (3 mL) was added sodium methoxide (2.5 mg, 0.047 mmol, 1 equi). The reaction mixture was stirred for 13 hours at which point TLC (1:4:4 water/isopropanol/ethyl acetate) indicated complete consumption of the starting material (R_f 0) and the formation of a single product (R_f 0.3). The reaction was quenched with Dowex 50WX8 100-200 mesh (H^+ form), filtered and the solvent removed *in vacuo*. The compound was purified by silica gel chromatography (H_2O:*i*-PrOH:EtOAc, 1:4:4). Lyophilization gave the title compound as a white amorphous solid (16 mg, 92 %).

TLC (H_2O:*i*-PrOH:EtOAc,1:4:4, v/v/v): R_f=0.30; $[\alpha]_D^{25}$ + 167.3 (c = 0.26 in H_2O) [Lit. $[\alpha]_D$ + 149 (c = 0.81 in MeOH)][49] ;^{1}H NMR (500 MHz, D_2O): δ 3.35 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.3 Hz, H-4'), 3.37 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.5, H-4), 3.47 (1 H, dd, $J_{6a,6b}$ = 13.6 Hz, $J_{5,6a}$ = 5.7 Hz, H-6a'), 3.56 (1 H, dd, $J_{2,3}$ = 10.4 Hz , $J_{1,2}$ = 3.8 Hz, H-2'), 3.58 (1 H, dd, $J_{2,3}$ = 10.1 Hz, $J_{1,2}$ = 4.0 Hz, H-2), 3.59 (1 H, dd, $J_{6a,6b}$ = 13.7 Hz, $J_{5,6b}$ = 2.5 Hz, H-6b'), 3.67 (1 H, dd, $J_{6a,6b}$ = 11.7 Hz, $J_{5,6a}$ = 5.0 Hz, H-6a), 3.72 - 3.79 (4 H, m, H-5, H-3,H-3', H-6b), 3.88(1 H, ddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 3.8 Hz, $J_{5,6b}$ = 2.2 Hz, H-5'), 5.10 (1 H, d, $J_{1,2}$ = 4.2 Hz, H-1), 5.11 (1 H, d, $J_{1,2}$ = 4.7 Hz, H-1'); ^{13}C NMR (126MHz, D_2O): δ 50.8 (C-6'-azide), 60.4 (C-6), 69.6 (C-4'), 70.4 (C-4), 70.9 (C-2'), 70.9 (C-2), 70.9 (C-5), 72.2 (C-5'), 72.3 (C-3'), 72.5 (C-3), 93.4 (C-1'), 93.6 (C-1); IR (KBr disc) ν = 2110 (N_3), 3456 cm^{-1}(br OH); MS *m/z* (ESI$^+$): 390.1 $[M+Na]^+$.

8.2.15 Synthesis of nitrophenol esters, 203-206

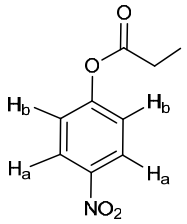
para-Nitrophenyl acetate (203)



p-Nitrophenol (82mg, 0.6mmol) was dissolved in anhydrous pyridine and to this was added propionic anhydride (2 mL). Reaction was stirred at RT for 30 minutes upon which time TLC (petrol:EtOAc, 5:1) indicated complete conversion to product (R_f 0.95) and disappearance of starting material (R_f 0.3). Reaction was extracted into DCM, washed with satd bicarbonate solution and concentrated *in vacuo* and azeotroped with toluene 3 x 20 mL to yield the desired product as a yellow crystalline solid (1.96mg, 94%).

TLC (petrol:EtOAc, 5:1, v/v): R_f = 0.95; M.p. = 74-76^{o}C; ^{1}H NMR (500 MHz CHLOROFORM-*d*) δ ppm 2.36 (3 H, s, CH_3), 7.29 (2 H, d, J_{Ha-Hb} 9.1 Hz, H-b), 8.27 (2 H, d, J_{Ha-Hb} 9.1 Hz, Ha) MS *m/z* (FI+) 181.0 (M).

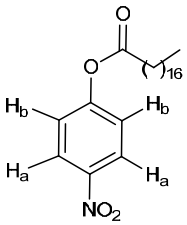
para-Nitrophenyl propionate (204)



p-Nitrophenol (945 mg, 6.79 mmol) was dissolved in anhydrous pyridine and to this was added acetic anhydride (2 mL). Reaction was stirred at RT for 30 minutes upon which time TLC (petrol:EtOAc, 5:1) indicated complete conversion to product (R_f 0.55) and disappearance of starting material (R_f 0.3). Reaction was extracted into DCM, washed with satd bicarbonate solution and concentrated *in vacuo* and azeotroped with toluene 3 x 20 mL to yield the desired product as a yellow crystalline solid (94mg, 74%).

TLC (petrol:EtOAc, 5:1, v/v): R_f = 0.55; M.p. = 78-82°C ; ^{1}H NMR (500 MHz CHLOROFORM-*d*) δ ppm 1.28 (3 H, t, $J_{CH2\text{-}CH3}$ 7.6 Hz CH_3), 2.64 (2 H, q, $J_{CH2\text{-}CH3}$ 7.6 Hz, CH_2), 7.27 (2 H, d, *J* 9.1 Hz, H-a), 8.27 (2 H, d, *J* 9.1 Hz, H-b); MS *m/z* (FI+) 195.1 (M).

para-Nitrophenyl stearate (205)

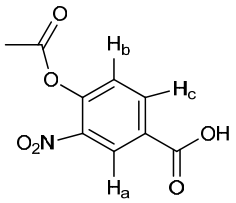


p-Nitrophenol (82mg, 0.6mmol) was dissolved in anhydrous pyridine and to this was added stearoyl chloride (175.16mg, 0.58mmol). Reaction was stirred at RT for 30 minutes upon which time TLC (petrol:EtOAc, 5:1) indicated complete conversion to product (R_f 0.95) and disappearance of starting material (R_f 0.3). Reaction was concentrated *in vacuo* and purified by column chromatography (6:1 petrol/ethyl acetate) to yield the desired product as a yellow crystalline solid (200mg, 85%).

TLC (petrol:EtOAc, 5:1, v/v): R_f = 0.95; M.p. = 72-74°C ; ^{1}H NMR (400 MHz CHLOROFORM-*d*) δ ppm 0.84 - 0.90 (3 H, t, *J* 7.1 Hz, CH_3), 1.18 - 1.30 (28 H, m, 14 x CH_2), 1.76 (3 H, quin, *J* 7.5 Hz, CH_2), 2.60 (3 H, t, *J* 7.6 Hz, CH_2), 7.28 (3 H, $J_{Ha\text{-}Hb}$ 9.1 Hz, Hb), 8.27 (2 H, d, $J_{Ha\text{-}Hb}$ 9.1 Hz, Ha); ^{13}C NMR (101 MHz, CHLOROFORM-*d*) δ ppm 14.12 (CH_3), 22.69 (CH_2), 24.73(CH_2), 25.30(CH_2), 29.04(CH_2), 29.08(CH_2), 29.21(CH_2), 29.26(CH_2), 29.36(CH_2), 29.43, 29.58(CH_2), 29.69 (m, 4C, CH_2), 31.92(CH_2), 34.32(CH_2), 122.44(C-b), 125.18(C-a), 145.21 (C-O), 155.50 (C-NO_2) 171.36

(C=O); IR (thin film): ν = 2917.44, 2850.30 (Ar-H), 1753.93 (C=O), 1536.57 (N-O), 1348.37 (N-O), 1219.33, 1139.51, 867.74 cm^{-1}**;** HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd for $C_{24}H_{39}NO_4Na^+$, 428.2777; found, 428.2775 $[M + Na]^+$.

3-Nitro-4-O-acetyl benzoic acid (206)

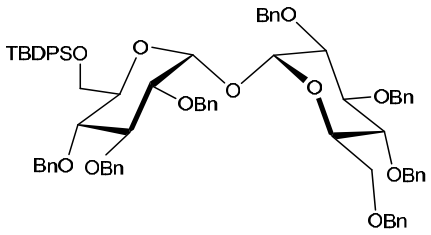


3-nitro-4-hydroxyl benzoic acid (350mg, 1.92mmol) was dissolved in anhydrous pyridine (5mL) and to this was added acetic anhydride (.5mL). After 1 hr at RT, TLC (petrol:EtOAc, 2:1) indictated formation of a uv active product (R_f 0.2) and disappearance of starting material (R_f 0.1). Reaction was extracted into DCM, acidified to pH 3 and azeotroped with toluene (3 x 10 mL). Product crystallized as the pyridinium salt (indicated by nmr). Additional washes with 1M $CuSO_4$ solution yielded the product as a yellow solid (120 mg, 27%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.2; M.p. = 78-82^oC ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 2.43 (3 H, s, C*H*$_3$CO), 7.40 (1 H, s, H-a), 7.41 (4 H, s), 8.83 (1 H, d, *J* 2.0 Hz, H-b), (1 H, d, *J* 2.0 Hz, H-c); m/z (ESI^+) 242.3 ($M+NH_4$); ^{13}C NMR (101 MHz, CHLOROFORM-*d*) δ ppm 20.83 (C*H*$_3$CO), 125.83 126.81. 129.59, 136.00, 138.33, 148.05 (6 x Ar-C) 158.80 (COO), 168.07 (C*H*$_3$CO); MS *m/z* (FI+) 225.0273 (M).

8.2.16 Synthesis of compounds TMH and TDH 212 and 215

6-O-tert-butyldiphenylsilyl-2,3,4,2',3',4',6'-hepta-O-benzyl-α,α-D-trehalose (209)

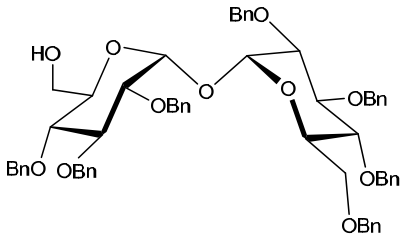


D-Trehalose (7.4 g, 21.75 mmol, 1 eq) was dissolved in anhydrous DMF (40 mL). To this was added tert-butyldiphenylchlorosilane (TBDPS-Cl) (5 mL, 18 mmol, 0.9 eq) and imidazole (1.4 g, 21 mmol, 0.95 eq). Solution was stirred at RT under Ar atmosphere for 18 h. TLC (petrol:EtOAc, 2:1) indicated

primarily starting material (R_f 0.1) and a small amount of the mono-TBDPS-trehalose. Additional TBDPS-Cl was added (2.5 mL, 9 mmol, 0.45 eq) and reaction was left for a further 18 hours. NaH (8 g, 339 mmol, 15 eq) and benzyl bromide (25 mL, 145 mmol, 7 eq) were added *in situ*. Reaction was stirred for a further 24 hours under argon, until the desired product could be detected by TLC (petrol:EtOAc, 5:1) (R_f 0.85). Column chromatography yielded the desired product as a slightly yellow oil. (4.83 g, 19% over two steps).

TLC (petrol:EtOAc, 5:1, v/v): R_f = 0.85; $[\alpha]_D^{25}$ + 16.1 (c = 0.88, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.14 (9 H, s, 3 x C($C\underline{H}_3$)$_3$), 3.47 (1 H, dd, $J_{6a,6b}$ = 10.2 Hz, $J_{5,6b}$ = 1.8 Hz, H-6b'), 3.60 (1 H, dd, *J* $J_{6a,6b}$ = 10.2 Hz, $J_{5,6a}$ = 2.7 Hz, H-6a'), 3.64 (1 H, dd, $J_{2,3}$ = 7.7 Hz, $J_{1,2}$ = 4.2 Hz, H-2'), 3.67 (1 H, dd, $J_{2,3}$ = 7.8 Hz, $J_{1,2}$ = 3.8 Hz, H-2), 3.72 (1 H, dd, $J_{6a,6b}$ = 9.7 Hz, $J_{5,6a}$ = 3.7 Hz, H-6a), 3.78 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.4 Hz, H-4'), 3.87 (1 H, dd, $J_{6a,6b}$ = 10.4 Hz, $J_{5,6b}$ = 2.3 Hz, H-6b), 3.97 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.5 Hz, H-4), 4.13-4.17 (3 H, m, H-5, H-3, H-3') 4.27 (1 H, ddd, $J_{4,5}$ = 10.0 Hz, $J_{5,6a}$ = 2.8, $J_{5,6b}$ = 2.7 Hz, H-5'), 4.47 (1 H, d, *J* = 11.9 Hz, 1 x OCH_2Ph), 4.52 (1 H, d, *J* = 10.6 Hz, 1 x OCH_2Ph), 4.57 (1 H, d, *J* = 12.6 Hz, 1 x OCH_2Ph), 4.58 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph), 4.87 (1 H, d, *J* = 12.1 Hz, 1 x OCH_2Ph), 4.90 (1 H, d, *J* = 11.1 Hz, 1 x OCH_2Ph), 4.94 (1 H, d, *J* = 9.3 Hz, 1 x OCH_2Ph), 4.96 (1 H, d, *J* = 11.6 Hz, 1 x OCH_2Ph), 4.99 (1 H, d, *J* = 9.6 Hz, 1 x OCH_2Ph), 5.01 (1 H, d, *J* = 9.6 Hz, 1 x OCH_2Ph), 5.04 (1 H, d, *J* = 12.2 Hz, 1 x OCH_2Ph), 5.07 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 5.09 (1 H, d, *J* =10.6 Hz, 1 x OCH_2Ph), 5.28 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1'), 5.37 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1), 7.00 - 7.61 (45 H, m, Ar-H); ^{13}C NMR (101 MHz, $CDCl_3$): δ 20.7 (3 x C($C\underline{H}_3$)$_3$), 28.1 (1 x $\underline{C}(CH_3)_3$), 62.2 (C-6), 68.2 (C-6'), 70.6 (C-5'), 71.69 (C-5), 72.8 (OCH_2Ph), 72.9 (OCH_2Ph), 73.6 (OCH_2Ph), 75.1 (OCH_2Ph), 75.3 (OCH_2Ph), 75.7 (OCH_2Ph), 75.9 (OCH_2Ph), 77.70 (C-4 or C-4'), 77.7 (C-4 or C-4'), 79.5 (C-2'), 80.1 (C-2), 81.9 (C-3 or C-3'), 81.9 (C-3 or C-3'), 94.2 (2 C, C-1, C-1'), 126.0, 126.3, 126.4, 127,1 127.2, 127.3, 127.4, 127.5, 127.6, 127.7, 127.8, 127.9, 128.0, 128.0, 128.1, 128.2, 128.4, 128.4, 128.5, 128.6, 128.9, 129.1, 129.5, 129.7 (7 x OCH_2Ph), 138.0, 138.4, 138.5, 138.6, 138.9, 139.0, 139.6 (7 x 1 C, 7 x OCH_2Ph); IR (thin film): 2925, 2875 (C=CH), 1454, 1444 (C=C), 698 cm^{-1}; MS *m/z* (ESI^+): 1228.6 $[M+NH_4]^+$; isotopic distribution: Species observed $[M+Na]^+$, peaks observed 1233.53 (100%), 1234.53 (86.6%), 1235.53 (43.1%), 1236.53 (14.5%), 1237.53 (4.0%), 1238.54 (1.1%) peaks calculated 1233.55 (100.0%), 1234.55 (89.7%), 1235.55 (45.2%), 1236.56 (16.0%), 1237.56 (4.2%), 1238.56 (0.6%).

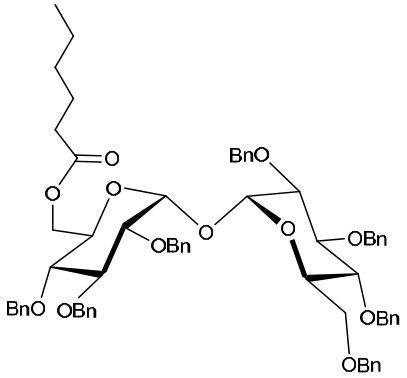
2,3,4,2',3',4',6'-hepta-O-benzyl-α,α-D-trehalose (210)[50]



209 (4 g, 3.3 mmol, 1 eq) was dried under reduced pressure for 1 hour and dissolved in anhydrous methanol (20 mL). To this was added acetyl chloride (5 mL), generating acetic acid *in situ*. Reaction was stirred in acid for 48 hours, upon which time TLC (petrol:EtOAc, 2:1) indicated product (R_f 0.4) and complete consumption of starting material (R_f 0.95). Reaction was quenched with water and poured into a satd. $NaHCO_3$ soln. When gas ceased to evolve, reaction was extracted into DCM and concentrated *in vacuo*. Column chromatography (petrol:EtOAc, 10:1 to 1:1 gradient) yielded the desired product as a colorless oil (1.1 g, 30%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.40; $[\alpha]_D^{25}$ + 54.3 (c = 0.77, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 3.38 (1 H, dd, $J_{6a,6b}$ = 10.7, $J_{5,6b}$ = 2.2 Hz, H-6b), 3.51 (1 H, dd, $J_{6a,6b}$ = 10.4 Hz, $J_{5,6a}$ = 3.2 Hz, H-6a), 3.52 (1H, dd, $J_{6a,6b}$ = 9.6 Hz, $J_{5,6a}$ = 3.6 Hz, H-6a') 3.57 - 3.61 (4 H, m, H-2', H-2, H-4, H-6b'), 3.68 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4'), 4.03 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.5 Hz, H-3), 4.06 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.3 Hz, H-3'), 4.05 - 4.09 (1 H, m, H-5'), 4.15 (1 H, ddd, $J_{4,5}$ = 10.0 Hz, $J_{5,6a}$ = 3.5 Hz, $J_{5,6b}$ = 2.4 Hz, H-5), 4.38 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.46 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.54 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.65 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.69 (2 H, d, *J* = 10.0 Hz, 2 x OCH_2Ph), 4.71 (2 H, d, *J* = *2* x OCH_2Ph), 4.82 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.87 (1 H, d, *J* =11.0 Hz, 1 x OCH_2Ph), 4.88 (2 H, d, *J* = 10.7 Hz, 2 x OCH_2Ph), 4.99 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 5.00 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 5.18 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1), 5.19 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1'), 7.25-2.40 (35 H, m, Ar-H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 61.5 (C-6), 68.12 (C-6'), 70.6 (C-5), 71.1 (C-5'), 72.1 (OCH_2Ph), 72.8 (OCH_2Ph), 72.8 (OCH_2Ph), 73.5 (OCH_2Ph), 75.0 (OCH_2Ph), 75.5 (2 C, 2 x OCH_2Ph), 77.0 (C-4'), 77.6 (C-4), 79.3 (C-2 or C-2'), 79.5 (C-2 or C-2'), 81.5 (C-3), 81.7 (C-3'), 94.1 (C-1 or C-1'), 94.3 (C-1 or C-1'), 127.4, 127.5, 127.6, 127.6, 127.7, 127.8, 127.9, 128.1, 128.3, 128.3, 128.4, 128.4 (7x OCH_2Ph), 134.6, 137.7, 138.1, 138.1, 138.3, 138.7, 138.8 (7 x 1 C, 7 x OCH_2Ph); MS *m/z* (ESI^+): 990.4 $[M+NH_4]^+$.

6-O-hexanoyl-2,3,4,2',3',4,6'-hepta-O-benzyl-α,α-D-trehalose (211)

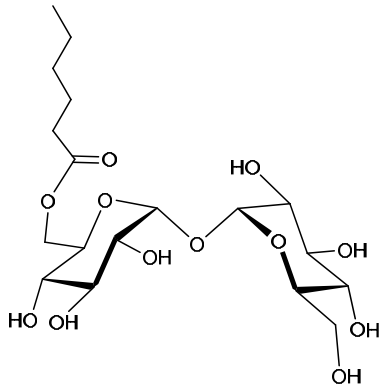


210 (116 mg, 0.119 mmol, 1 eq) was dissolved in anhydrous DCM (10 mL) with anhydrous pyridine (2 mL) and to this was added hexanoyl chloride (0.02 mL, 0.15 mmol, 1.25 eq), which, upon addition caused the reaction mixture to turn yellow. Reaction was stirred for 16 h at RT upon which time TLC (petrol:EtOAc, 2:1) indicated incomplete conversion from starting material (R_f 0.1) to product (R_f 0.6). An additional portion of hexanoyl chloride (0.05 mL, 0.375 mmol, 3.125 eq) was added and reaction was left at RT for an additional 24 h. Reaction was washed with satd $NaHCO_3$ solution and concentrated *in vacuo*. Column chromatography (petrol:EtOAc, 4:1 to 2:1 gradient) yielded the desired product as a clear oil (52.1 mg, 40%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.60; $[\alpha]_D^{25}$ + 30.4 (c = 0.56, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 0.86 (3 H, t, *J* = 6.8 Hz, $OCO(CH_2)_4C\underline{H}_3$), 1.20 - 1.30 (2 H, m $OCO(CH_2)_3C\underline{H}_2CH_3$), 2.16 - 2.31 (4 H, m, $OCOCH_2(C\underline{H}_2)_2CH_2CH_3$), 2.24 (2 H, t, *J* = 7.3 Hz, OCO $OCOC\underline{H}_2(CH_2)_3CH_3$), 3.38 (1 H, dd, $J_{6a,6b}$ = 10.7 Hz, $J_{5,6b}$ = 1.6 Hz, H-6b), 3.52 (1 H, dd, $J_{6a,6b}$ = 10.8 Hz, $J_{5,6a}$ = 3.7 Hz, H-6a), 3.53 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.2 Hz, H-4), 3.56 (1 H, dd, $J_{2,3}$ = 9.4 Hz, $J_{1,2}$ = 3.5 Hz, H-2), 3.60 (1 H, dd, $J_{2,3}$ = 9.8 Hz, $J_{1,2}$ = 3.8 Hz, H-2'), 3.68 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4'), 4.03 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.2 Hz, H-3), 4.06 (1 H, at, $J_{2,3}$= $J_{3,4}$ = 8.8 Hz, H-3'), 4.06 (1 H, dd, $J_{6a,6b}$ = 10.8 Hz, $J_{5,6b}$ = 2.2 Hz, H-6b'), 4.11 - 4.18 (2 H, m, H-6a', H-5'), 4.25 (1 H, ddd, $J_{4,5}$ = 9.9 Hz, $J_{5,6b}$ = 3.4 Hz, $J_{5,6a}$ = 2.2 Hz, H-5), 4.38 (1 H, d, *J* = 12.3 Hz, 1 x OCH_2Ph), 4.46 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.52 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.54 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.68 (2 H, d, *J* = 11.7 Hz, 2 x OCH_2Ph), 4.72 (2 H, d, *J* = 12.0 Hz, 2 x OCH_2Ph), 4.81 (1 H, d, *J* = 10.7 Hz 1 x OCH_2Ph), 4.86 (3 H, d, *J* = 11.1 Hz, 3 x OCH_2Ph), 4.99 (1 H, d, *J* = 10.8 Hz, 1 x OCH_2Ph), 5.01 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 5.20 (1 H, d, $J_{1,2}$ =3.3 Hz, H-1), 5.21 (1 H, d, $J_{1,2}$ = 3.2 Hz, H-1'), 7.21 - 7.35 (35 H, m, Ar-H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 13.9 ($OCO(CH_2)_4\underline{C}H_3$), 22.2 (1 C, 1 x $OCO(CH_2)_3\underline{C}H_2CH_3$), 24.5 (1 C, 1 x $OCO(CH_2)_2\underline{C}H_2CH_2CH_3$), 31.2 (1 C, $OCOCH_2\underline{C}H_2(CH_2)_2CH_3$), 34.0 (1 C, $OCO\underline{C}H_2(CH_2)_3CH_3$), 62.5 (C-6'), 68.1 (C-6), 69.0 (C-5), 70.7 (C-5'), 72.7 (OCH_2Ph), 72.9 (OCH_2Ph), 73.5 (OCH_2Ph), 75.1 (OCH_2Ph), 75.1 (OCH_2Ph), 75.6 (OCH_2Ph), 75.7 (OCH_2Ph), 77.5 (C-4 or C-4'), 77.6 (C-4 or C-4'), 79.2 (C-2 or C-2'), 79.4 (C-2 or C-2'), 81.6 (C-3 or C-3'), 81.8 (C-3

or C-3'), 94.0 (C-1 or C-1'), 94.4 (C-1 or C-1'), 127.3, 127.5, 127.5, 127.6, 127.6, 127.7, 127.8, 127.9, 127.9, 127.9, 128.1, 128.3, 128.4, 128.4, 128.5 (7 x OCH_2Ph), 137.7, 137.9 (2 C), 138.0, 138.3, 138.6, 138.8, (7 x 1 C = 7 x OCH_2Ph) 173.5 (C=O); IR (thin film): 3063, 3031, 2925, 2855 (C=CH), 1735 (C=O), 1586, 1496, 1454, 1360 (C=C), 1262, 1219, 1156, 1100, 772 cm^{-1}; MS *m/z* (ESI^+): 1093.4 $[M+Na]^+$; isotopic distribution: species observed $[M+Na]^+$, peaks observed 1093.47 (100%), 1094.48 (73.7 %) 1095.49 (27.4 %), 1096.50 (6.4%) peaks calculated 1093.51 (100%), 1094.51 (73.8%), 1095.51 (28.3%), 1096.52 (8.2%).

6-O-hexanoyl-α,α-D-trehalose (212)

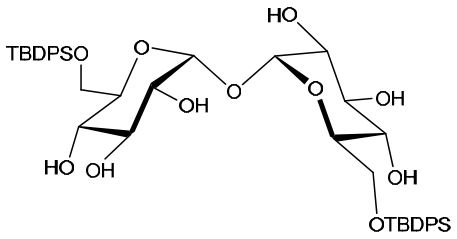


211 (6.2 mg, 0.019 mmol) was dissolved in 20 mL ethanol and circulated through a 10% Pd/C cartridge in the H-Cube® at 80 bar, 45 ˚C and 1 mL/min. Reaction was monitored by mass spec and after 3 hours TLC (EtOAc:MeOH, 2:1) indicated complete conversion from starting material (R_f 1) to product (R_f 0.4). Reaction was concentrated *in vacuo* and purified via column chromatography (EtOAc:MeOH, 3:1) to give the desired product as a white crystalline solid (2.6 mg, 93%).

TLC (EtOAc:MeOH, 2:1, v/v): R_f=0.34; M.p. = 135-137 ^oC; $[\alpha]_D^{25}$ = 111.3 (c = 1.0, MeOH). ^1H NMR (500 MHz, D_2O): δ 0.79 (3 H, t, *J* = 6.9 Hz, $OCO(CH_2)_4C\underline{H}_3$), 1.17 - 1.26 (4 H, m, $OCOCH_2CH_2(C\underline{H}_2)_2CH_3$), 1.54 (2 H, quin, *J* = 7.3 Hz, $OCOCH_2C\underline{H}_2(CH_2)_2CH_3$), 2.35 (2 H, t, *J* = 7.4 Hz, $OCOC\underline{H}_2CH_2(CH_2)_2CH_3$), 3.36 (1 H, at, $J_{3,4} = J_{4,5}$ = 9.5 Hz, H-4), 3.41 (1 H, at, $J_{3,4} = J_{4,5}$ = 9.6 Hz, H-4'), 3.55 (1 H, dd, $J_{2,3}$ = 9.9 Hz, $J_{1,2}$ = 3.8 Hz, H-2), 3.57 (1 H, dd, $J_{2,3}$ = 10.1 Hz, $J_{1,2}$ = 3.8 Hz, H-2'), 3.66 - 3.82 (5 H, m, H-6a, H-6b, H-5, H-3, H-3'), 3.94 (2 H, ddd, $J_{4,5}$ = 10.2 Hz, $J_{5,6a}$ = 4.8 Hz, $J_{5,6b}$ = 2.0 Hz, H-5'), 4.22 (2 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6a}$ = 5.0 Hz, H-6a'), 4.35 (2 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6b}$ = 2.2 Hz, H-6b'), 5.07 (1 H, d, $J_{1,2}$ = 3.9 Hz, H-1), 5.10 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1'); ^{13}C NMR (126 MHz, D_2O): δ 13.1 (1 C, $OCO(CH_2)_4\underline{C}H_3$), 21.6 (1 C, $OCO(CH_2)_3\underline{C}H_2CH_3$), 24.0 (1 C, $OCO(CH_2)_3\underline{C}H_2CH_2CH_3$), 30.5 (1 C, $OCOCH_2\underline{C}H_2(CH_2)_2CH_3$), 33.7 (1 C, $OCO\underline{C}H_2(CH_2)_3CH_3$), 60.5 (C-6), 62.9 (C-6'), 69.6 (C-5'), 69.6 (C-4 or C-4'), 69.9 (1 C-4 or C-4'), 70.9 (C-2 or C-2'), 70.9

(C-2 or C-2'), 72.2 (C-5), 72.3 (C-3 or C-3'), 72.6 (C-3 or C-3'), 93.2 (C-1'), 93.4 (C-1), 176.8 (C=O); HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{18}H_{32}O_{12}Na^+$, 463.1786; found, 463.1793.

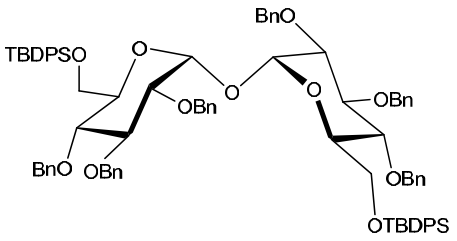
***6,6'-O-Di-tertbutyldiphenylsilyl-α,α-d-Trehalose* (207)**[31]



To a stirred suspension of D-trehalose (2g, 5.85 mmol) and imidazole (0.39g, 5.29 mmol) in dry DMF (10mL) was added tert-butyl diphenylchlorosilane (TBDPS-Cl) (3mL, 11.5mmol) at RT. After stirring for 36 hours, TLC (ETOAc:MeOH, 2:1), indicated the complete consumption of starting materials and the formation of the product (R_f 0.35) as well as the mono TBDPS-trehalose (R_f 0.2). The reaction mixture was concentrated *in vacuo* and purification attained via silica gel chromatography (*i*-PrOH:ETOAc, 2:1) to give the desired product as a white solid (2.315, 55%).

TLC (ETOAc:MeOH, 2:1, v/v): R_f=0.34; $[\alpha]_D^{17.3}$ +71.5 (1,4-dioxane); ^{1}H NMR (500MHz, DMSO); δ ppm 1.0 (18 H, s, 6 x CH_3), 3.23-3.32 (4 H, m, H-2, H-2', H-4, H-4'), 3.62 (2 H, dd, $J_{2,3}$ = 9.14 Hz, $J_{3,4}$ = 5.04 Hz, H-3, H-3'), 3.80 (2 H, dd, $J_{5,6b}$ = 5.67 Hz, $J_{6a,6b}$ = 11.21 Hz, H-6b, H-6b'), 3.85-3.87 (4 H, m, H-5, H-5', H-6a, H-6a'), 4.62 (2 H, d, $J_{2,OH}$ = 6.31 Hz, C-2 OH, C-2' OH), 4.84 (2 H, d, $J_{3,OH}$ = 4.73 Hz, C-3 OH, C-3' OH), 4.91 (2 H, d, $J_{4,OH}$ = 5.36 Hz, C-4 OH, C-4' OH), 5.06 (2 H, d, $J_{1,2}$ = 3.07 Hz, H-1, H-1'), 7.37-7.71 (40 H, m, Ar-H); ^{13}C NMR (126MHz, DMSO) δ ppm 18.99 (2 C, s, 2 x $\underline{C}(CH_3)_3$), 26.62 (6 C2 x C($\underline{C}H_3)_3$), 63.40 (2 C, C-6, C-6'), 69.97 (2 C, C-4, C-4'), 71.76 (2 C, C-2, C-2'), 72.30 (2 C, C-5, C-5'), 73.20 (2 C, C-3, C-3'), 92.78 (2 C, C-1, C-1'), 127.51, 127.70, 129.19, 129.64, 134.48, 135.18 (20C, 6 x s, ArC), 133.34, 133.53 (4C, 2 x s, 4º ArC); MS *m/z* (ES^+) 841.3 ($M+Na^+$); HRMS (ESI^+) cald. For $C_{44}H_{58}O_{11}Si_2$ 841.3415 Found 841.3410 $[M+Na]^+$..

6,6'-di-O-tert-butyldiphenylsilyl-2,3,4,2',3',4'-hexa-O-benzyl-α,α-D-trehalose (213)

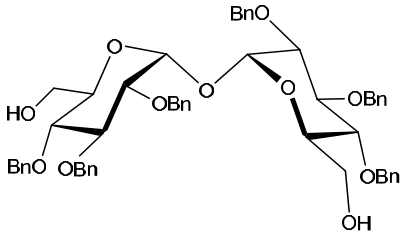


To a stirred suspension of D-trehalose **1** (2 g, 5.85 mmol, 1 eq) and imidazole (0.39 g, 5.29 mmol, 0.9 eq) in dry DMF (10 mL) was added *tert*-butyl diphenylchlorosilane (TBDPS-Cl) (3 mL, 11.5 mmol, 2.0 eq) at RT. After stirring for 36 hours, TLC (EtOAc:MeOH, 2:1), indicated the complete consumption of starting materials and the formation of the **207** (R_f 0.35) as well as **208** (R_f 0.2). The reaction mixture was concentrated *in vacuo* and purification attained via silica gel chromatography (EtOAc: *i*-PrOH, 2:1) to give the desired product as a mixture of mono and di-TBDPS protected compounds as a white solid. (2.3 g, 55%), which were used without further purification. This mixture was dissolved in anhydrous DMF (25 mL), and sodium hydride (60% dispersed in mineral oil) (700 mg, 29.1 mmol) was added portionwise for a period of 10 min at 0 ºC. Benzyl bromide (2 mL, 11.6 mmol, 6 eq) was then added dropwise and the mixture left to stir under an atmosphere of argon at room temperature. After an 18 h period, TLC (petrol:EtOAc, 5:1) indicated the formation of product (R_f 0.9) with complete consumption of the starting material (R_f 0). The reaction mixture was quenched by the slow addition of methanol (150 mL) and stirred for 30 min, at which point the resulting solution was concentrated *in vacuo*. The residue was dissolved in DCM (800 mL), washed with water and brine, filtered and concentrated *in vacuo*. Purification by column column chromatography (petrol:EtOAc, 10:1) afforded **213** and **210** as a 2:1 mixture of products (by NMR), which were not fully separated (2.987 g, 41% over two steps) as a viscous clear oil. Further column chromatography yielded pure **213** (300 mg, 5% over two steps).

TLC (petrol:EtOAc, 5:1, v/v): R_f = 0.90; $[\alpha]_D^{25}$ + 23.6 (c = 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.28 (18 H, s, 2 x C (CH_3)), 3.75 (2 H, dd, $J_{2,3}$ = 9.8 Hz, $J_{1,2}$ = 3.8 Hz, H-2, H-2'), 3.79 (2 H, dd, $J_{6a,6b}$ = 10.5 Hz, $J_{5,6b}$ = 1.8 Hz, H-6b, H-6b'), 3.92 (2 H, dd, $J_{6a,6b}$ = 10.6 Hz, $J_{5,6a}$ = 2.8 Hz, H-6a, H-6a'), 4.03 (2 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4, H-4'), 4.18 - 4.28 (4 H, m, H-5, H-5', H-3,H-3'), 4.72 (2 H, d, *J* = 11.8 Hz, 2 x OCH_2Ph), 4.80 (2 H, d, *J* = 11.9 Hz, 2 x OCH_2Ph), 4.89 (2 H, d, *J* = 10.9 Hz, 2 x OCH_2Ph), 5.04 (2 H, d, *J* = 10.6 Hz, 2 x OCH_2Ph), 5.10 (2 H, d, *J* = 10.6 Hz, 2 x OCH_2Ph), 5.15 (2 H, d, *J* = 10.6 Hz, 2 x OCH_2Ph), 5.39 (2 H, d, $J_{1,2}$ = 3.5 Hz, H-1, H-1'), 7.13 - 7.63 (40 H, m, Ar-H) 7.79 - 7.93 (10 H, m, Ar-H); ^{13}C NMR (101 MHz, $CDCl_3$): δ 19.5 (6 C, s, 3 x C($\underline{C}H_3)_3$), 27.0 (2 C, s, 2 x $\underline{C}(CH_3)_3$), 65.6 (2 C, C-6, C-6'), 71.7 (2 C, C-5, C-5'), 72.9 (2 C, OCH_2Ph), 75.3 (2 C, OCH_2Ph), 76.0 (2 C, OCH_2Ph), 77.7 (2 C, C-4, C-4'), 80.3 (2 C, C-2, C-2'), 82.1 (2 C, C-3, C-3'), 94.2 (2 C, C-1, C-1'), 126.1, 127.0, 127.2, 127.4, 127.7, 127.7, 127.8, 127.9, 128.0, 128.1, 128.3, 128.3, 128.6, 128.9,

129.2, 129.8, 129.9, 133.5, 133.6, 133.7, 135.6, 135.8, (6 x OCH_2Ph, 4 x OTBDPS), 137.9, 138.2, 138.4, 138.7, 138.9, (10 C, 6 x 1 C, OCH_2Ph, 4 x 1 C, OTBDPS); IR (thin film): 3069, 2930, 2856 (C=CH), 1454, 1428, (C=C), 1219, 111, 1069, 1027, 824, 773 cm^{-1}; MS *m/z* (ESI^+): 1376.6 $[M+NH_4]^+$; isotopic distribution: Species observed $[M+Na]^+$, peaks observed 1381.59 (91.5%), 1382.59 (100%), 1383.59 (48.3%), 1384.59 (17.8%), 1385.60 (5.8%), 1386.59 (1.7%) peaks calculated 1381.62 (95.5%), 1382.62 (100%), 1383.62 (59.8%), 1384.62 (25.3%), 1385.63 (8.0%).

***2,3,4,2',3',4'-hexa-O-benzyl-α,α-D-trehalose (4)*[50]**

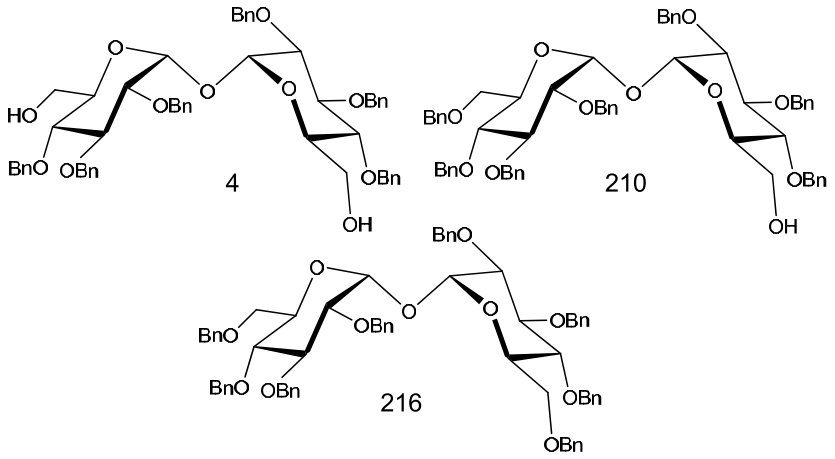


The 2:1 mixture of **213** and **210** (1.01g, 0.83 mmol) was dissolved in DMF (10 mL) and to this was added TBAF dropwise (70% in H_2O) (1 mL, 2.87 mmol, 3.5 eq). Reaction was stirred for 3 h at 60 ˚C. TLC (petrol:EtOAc, 2:1) indicated products (R_f 0.4) and (R_f 0.1) and complete consumption of starting material (R_f 1.0). Crude product was washed with satd brine and $NaHCO_3$, extracted into dichloromethane and concentrated *in vacuo.* Purification by column chromatography (petrol:EtOAc, 2:1 followed by 1:1 and then pure EtOAc) yielded **4** (202.5 mg, 36%, based on 2:1 ratio of starting material) as a clear oil and **210** (156 mg, 58% based on 2:1 ratio of starting material) and recovered starting material (130 mg, 7%). Combined yield of both products was 61%.

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.1; $[\alpha]_D^{25}$ + 88.1 (c = 1.0, $CHCl_3$) [Lit. $[\alpha]_D^{25}$ +104 (c = 1.6 in $CHCl_3$)][50]; ^{1}H NMR (400 MHz, $CDCl_3$): δ 3.55 (2 H, dd, $J_{2,3}$ = 9.8 Hz, $J_{1,2}$ = 3.1 Hz, H-2, H-2'), 3.59-3.64 (6 H, m, H-6a, H-6a', H-6b, H-6b'), 3.61 (2 H, at, $J_{3,4} = J_{4,5}$ = 10.1 Hz, H-4, H-4'), 4.05 - 4.13 (4 H, m, H-3, H-3', H-5, H-5'), 4.67 (2 H, d, *J* = 10.9 Hz, 2 x OCH_2Ph), 4.68 (2 H, d, *J* = 11.9 Hz, 2 x OCH_2Ph), 4.73 (2 H, d, *J* =11.9 Hz, 2 x OCH_2Ph), 4.90 (2H, d, *J* =10.9 Hz, 2 x OCH_2Ph), 4.91 (2 H, d, *J* =11.4 Hz, 2 x OCH_2Ph), 5.02 (2 H, d, *J* =10.8 Hz, 2 x OCH_2Ph), 5.16 (2 H, d, $J_{1,2}$ = 3.8 Hz, H-1, H-1), 7.15 - 7.52 (30 H, m, Ar-H); ^{13}C NMR (101 MHz, $CDCl_3$): δ 61.4 (2 C, C-6, C-6'), 71.4 (2 C, C-5, C-5'), 72.9 (2 C, 2 x OCH_2Ph), 75.0 (2 C, 2 x OCH_2Ph), 75.6 (1 C, 2 x OCH_2Ph), 76.7 (2 C, s, C-4, C-4'), 79.4 (2 C, C-2, C-2'), 81.6 (2 C, C-3, C-3'), 93.9 (2 C, C-1, C-1'), 123.8, 127.5, 127.6, 127.7, 127.9, 128.1, 128.4, 128.5 (6 x OCH_2Ph), 138.0, 138.2, 138.7 (6 x 1 C, 6 x OCH_2Ph); MS *m/z* (ESI^+): 900.4 $[M+NH_4]^+$.

8.2.16.1 Alternate synthesis of compounds 210 and 4

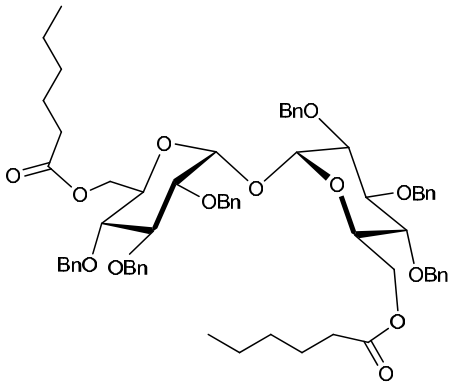
2,3,4,6,2',3',4',6'-Octyl-O-benzyl- α,α -D-trehalose (216)[50], 2,3,4,2',3',4'-hexa-O-benzyl-α,α-D-trehalose (4)[50] and 2,3,4,2',3',4',6'-hepta-O-benzyl-α,α-D-trehalose (210)[50]



D-trehalose (12.2 g, 35.7 mmol, 1 eq) and imidazole (7.3 g, 107 mmol, 3 eq) were dissolved in anhydrous DMF (80 250 mL) and to this was added TBDPSCl (23 mL, 83.9 mmol, 2.35 eq). Reaction mixture was stirred under Argon for 18 hours and upon complete dissappearance of starting trehalose by TLC (DCM:MeOH, 9:1), indicated the complete consumption of starting materials (R_f 0.0) and the formation of the diTBDPS-trehalose product (R_f 0.5) as well as the monoTBDPS-trehalose (R_f 0.1). Reaction was poured into ice water (400 mL) and the resulting white precipitate was filtered off and washed with ice water (400 mL) and ether (400 mL) and then concentrated under reduced pressure. This crude product was then redissolved in anhydrous DMF (200 mL) and cooled to 0 ˚C. To this was added benzyl bromide (35 mL, 204 mmol, 6.18 eq). Sodium hydride (6 g, 250 mmol, 7.4 eq) was added slowly to prevent uncontrolled reaction. Reaction was stirred under argon atmosphere fo 3 days, upon which TLC (petrol:EtOAc, 5:1) revealed complete converion to three products **213** (R_f 0.7), **209** (R_f 0.6) and **216** (R_f 0.25) and dissappearance of starting sugars (R_f 0). Reaction mixture was extracted into ether (3 x 500 mL) and washed with satd sodium bicarbonate (2 x 500 mL) and satd brine (2 x 500 mL) and concentrated under reduced pressure. Mixture was redissolved in DMF (300 mL) and to this was added TBAF (20 mL, 57.34 mmol, 1.6 eq) and reaction mixture was heated to 70 °C for 72 h, upon which time complete conversion to the desired **210** and **4** was detected by TLC (petrol:EtOAc, 1:1), which indicated products (R_f 0.4) and (R_f 0.1). Purification by column chromatography (petrol:EtOAc 8:1 followed by 6:1, 4:1, 2:1, 1:2 and ethyl acetate) yielded **4** (11 g, 34%) and **210** (9.5 g, 28%) and **216 (**6.23 g, 16%) as clear oils for a net yield of all three products over three steps of 79%.

Data for compound **216**; TLC (petrol:EtOAc, 5:1, v/v): R_f = 0.25; $[\alpha]_D^{25}$ + 79.3 (c = 1.0 in $CHCl_3$); ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 3.44 (2 H, dd, $J_{6a,6b}$ = 10.4 Hz, $J_{5,6b}$ = 1.5 Hz, H-6b), 3.58 (2 H, dd, $J_{6a,6b}$ = 10.6 Hz, $J_{5,6a}$ = 3.0 Hz, H-6a), 3.65 (2 H, dd, $J_{2,3}$ = 9.5 Hz, $J_{1,2}$ = 3.7 Hz, H-2, H-2'), 3.74 (2 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4, H-4'), 4.09 (2 H, t, *J* = 9.2 Hz), 4.23 (2 H, ddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 3.1 Hz, $J_{5,6b}$ = 1.5 Hz, H-5, H-5'), 4.43 (2 H, d, *J* = 12.1 Hz, 2 x OCH_2Ph), 4.51 (2 H, d, *J* = 10.6 Hz), 4.60 (2 H, d, *J* = 11.9 Hz, 2 x OCH_2Ph), 4.67 - 4.78 (4 H, m, 4 x OCH_2Ph), 4.87 (2 H, d, *J* = 10.9 Hz, 2 x OCH_2Ph), 4.91 (2 H, d, *J* = 10.9 Hz, 2 x OCH_2Ph), 5.05 (2 H, d, *J* = 10.9 Hz, 2 x OCH_2Ph), 5.29 (2 H, d, $J_{1,2}$ = 4.4 Hz, H-1, H-1'), 6.65 - 7.71 (40 H, m, Ar-H); MS *m/z* (ESI^+): 1085.9 $[M + Na]^+$.

6,6'-di-O-hexanoyl-2,3,4,2',3',4'-hexa-O-benzyl-α,α-D-trehalose (214)

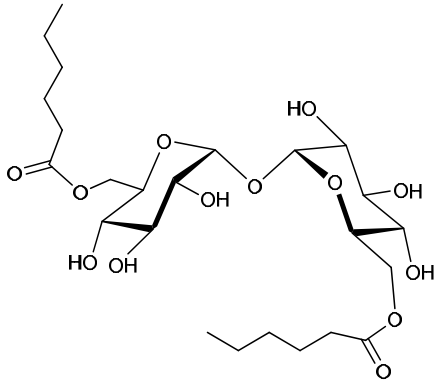


4 (1 g, 1.13 mmol, 1 eq) was dissolved in anhydrous pyridine (2 mL) and to this was added hexanoyl chloride (0.5 mL, 3.8 mmol, 3.3 eq). Reaction was stirred for 30 min at RT upon which time TLC (petrol:EtOAc, 2:1) indicated complete conversion from starting material (R_f 0.05) to product (R_f 0.9). Reaction was washed with satd. $NaHCO_3$ solution and concentrated *in vacuo*. Column chromatography (petrol:EtOAc, 10:1, 1% triethylamine) yielded the desired product as a clear oil (1.12g, 91%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.90; $[\alpha]_D^{25}$ + 75.6 (c = 0.39, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$): δ 0.86 (6 H, t, *J* = 6.9 Hz, 2 x $OCO(CH_2)_4C\underline{H}_3$), 1.23 - 1.29 (8 H, m, 2 x $OCO(CH_2)_2(C\underline{H}_2)_2CH_3$), 1.54 - 1.65 (8 H, m, 2 x $OCO(C\underline{H}_2)_2(CH_2)_2CH_3$), 3.54 (4 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4, H-4'), 3.57 (4 H, dd, $J_{2,3}$ = 9.6 Hz, $J_{1,2}$ = 3.5 Hz, H-2, H-2'), 4.06 (4 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.3 Hz, H-3, H-3'), 4.06 (4 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6a}$ = 2.0 Hz, H-6a, H-6a'), 4.16 (2 H, dd, $J_{6a,6b}$ = 12.3 Hz, $J_{5,6b}$ = 3.3 Hz H-6b, H-6b'), 4.24 (2 H, ddd, $J_{4,5}$ = 10.0 Hz, $J_{5,6b}$ = 3.4 Hz, $J_{5,6a}$ = 2.0 Hz, H-5, H-5'), 4.53 (2 H, d, *J* = 10.6 Hz, OCH_2Ph), 4.69 (2 H, d, *J* = 12.1 Hz, OCH_2Ph), 4.73 (2 H, d, *J* = 11.9 Hz, OCH_2Ph), 4.88 (4 H, d, *J* = 10.6 Hz, OCH_2Ph), 4.88 (3 H, d, *J* = 10.9 Hz, OCH_2Ph), 5.02 (2 H, d, *J* = 10.6 Hz, OCH_2Ph), 5.18 (2 H, d, $J_{1,2}$ = 3.5 Hz, H-1, H-1'), 7.04 - 7.51 (40 H, m, Ar-H); ^{13}C NMR (101 MHz, $CDCl_3$): δ 13.9 (2

C, $OCO(CH_2)_4\underline{C}H_3$), 22.3 (2 C, $OCO(CH_2)_4\underline{C}H_2CH_3$), 24.5 (2 C, $OCO(CH_2)_4\underline{C}H_2CH_2CH_3$), 31.3 (2 C, $OCOCH_2\underline{C}H_2$ $(CH_2)_2CH_3$), 34.0 (2 C, $OCO\underline{C}H_2(CH_2)_3CH_3$), 62.5 (2 C, C-6, C-6'), 69.1 (2 C, C-5, C-5'), 72.9 (2 C, OCH_2Ph), 75.2 (4 C, 2 x OCH_2Ph), 75.7 (2 C, OCH_2Ph), 77.4 (2 C, C-4, C-4'), 79.2 (2 C, C-2, C-2'), 81.6 (2 C, C-3, C-3'), 94.0 (2 C, C-1, C-1'), 127.5, 127.7, 127.8, 127.9, 128.1, 128.4, 128.4 (6 x OCH_2Ph); 137.8, 137.8, 138.5 (6 x 1 C = 6 x OCH_2Ph); 173.5 (2 C, 2 x C=O); IR (thin film): 3419 br (OH), 2956, 2925, 2854 (C=CH), 1733 (C=O), 1637, 1456 (C=C), 1219, 1177, 1150, 1101, 1078, 1053, 1026, 990, 772 cm^{-1}; MS *m/z* (ESI^+): 1096.5 $[M+NH_4]^+$; isotopic distribution: species observed $[M+Na]^+$, peaks observed 1101.53 (100%), 1102.53 (72.8%), 1103.54 (28.8%), 1104.54 (8.1%), 1105.54 (1.8%), peaks calculated 1101.52 (100%), 1102.52 (74.4%), 1103.53 (30.0%), 1104.53 (8.3%), 1105.53 (1.9%), 1105.49 (1.4%).

6,6'-di-O-hexanoyl-α,α-D-trehalose (215)



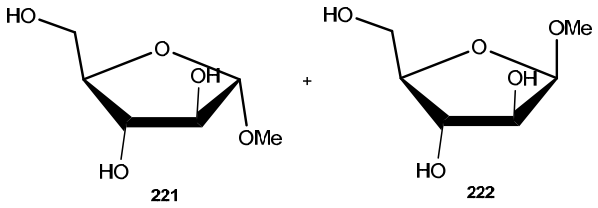
214 (474 mg, 0.44 mmol) was dissolved in 20 mL ethanol and the solvent was degassed under alternating reduced pressure and argon. 10% Pd/C (300 mg) was added to the solution, which was then activated through repeated vacuum, flush cycles with 2 hydrogen balloons. Reaction was stirred at RT for 16 h, upon which time ESI mass spec showed complete conversion to the desired product. TLC (EtOAc:MeOH, 2:1) indicated complete conversion from starting material (R_f 1) to product (R_f 0.6). Reaction was concentrated *in vacuo* to produce the desired product as a crystalline white solid (120.0 mg, 55%).

TLC (EtOAc:MeOH, 2:1, v/v): R_f = 0.6; M.p. = 157.7-159.0 ^oC; $[\alpha]_D^{25}$ + 152.3 (c = 0.13, H_2O); ^1H NMR (500 MHz, D_2O): δ 0.78 (6 H, t, J = 6.9 Hz, 2 x $OCO(CH_2)_4C\underline{H}_3$), 1.11 - 1.29 (8 H, m, 2 x $OCO(CH_2)_2(C\underline{H}_2)_2CH_3$)1.53 (4 H, quin, J = 7.3 Hz, 2 x $OCOCH_2(C\underline{H}_2)(CH_2)_2CH_3$), 2.33 (4 H, t, J = 7.3 Hz, 2 x $OCOC\underline{H}_2(CH_2)_4CH_3$), 3.39 (2 H, at, $J_{3,4} = J_{4,5}$ = 9.6 Hz H-4, H-4'), 3.54 (2 H, dd, $J_{2,3}$ = 9.8 Hz, $J_{1,2}$= 3.8 Hz, H-2, H-2'), 3.76 (2 H, at, $J_{2,3} = J_{3,4}$ = 9.5 Hz, H-3, H-3'), 3.92 (2 H, ddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 5.0 Hz, $J_{5,6b}$ = 2.2 Hz, H-5, H-5'), 4.21 (2 H, dd, $J_{6a,6b}$ =12.3, $J_{5,6a}$ = 5.4 Hz, H-6a), 4.33 (2 H, dd, $J_{6a,6b}$ =12.3 Hz, $J_{5,6b}$ 2.2 Hz, H-6b), 5.05 (2 H, d, $J_{1,2}$ = 3.8 Hz, H-1, H-1'); ^{13}C NMR

(126 MHz, D_2O): δ 13.1 (2 C, 2 x $OCO(CH_2)_4\underline{C}H_3$), 21.6 (2 C, 2 x $OCO(CH_2)_3\underline{C}H_2CH_3$), 24.0 (2 C, 2 x $OCO(CH_2)_2\underline{C}H_2CH_2CH_3$), 30.5 (2 C, 2 x $OCO\ CH_2\underline{C}H_2(CH_2)_2CH_3$), 33.8 (2 C, 2 x $OCO\underline{C}H_2(CH_2)_4CH_3$), 62.9 (2 C, C-6, C-6'), 69.7 (2 C, C-5, C-5'), 70.0 (2 C, C-4, C-4'), 70.92 (2 C, C-2, C-2'), 72.4 (2 C, C-3, C-3'), 93.3 (2 C, C-1, C-1'), 176.8 (2 C, 2 x C=O); HRMS *m/z* (ESI^+): $[M+Na]^+$ calcd. for $C_{24}H_{42}O_{13}Na^+$, 561.2518; found, 561.2513; analysis (calcd., found for $C_{24}H_{42}O_{13}$): C (53.52,53.56), H (7.86,7.84).

8.2.17 Synthesis of arabinofuranosides 221 and 222

Methyl α-D-arabinofuranoside (221) and Methyl β-D-arabinofuranoside (222)[51]



To a suspension of D-arabinose (0.74g, 4.9mmol) in anhydrous MeOH (60 mL), acetyl chloride (0.1mL) was added dropwise at 0 °C temperature under Ar atmosphere and the resulting solution was stirred at 0 °C. The reaction was monitored by TLC (DCM:MeOH, 5:1) and appeared to stop after 3 days with the disappearance of the starting sugar (R_f 0) and appearance of two product spots **221** (R_f 0.45) and **222** (R_f 0.4). The solution was allowed to attain room temperature and excess acid was neutralized by Na_2CO_3 and the resulting suspension was filtered through Celite. The combined filtrate was concentrated under reduced pressure and the crude residue thus obtained was dissolved in MeOH, and purified by column chromatography (MeOH:acetone:$CHCl_3$, 1:1:3) to yield the methyl glycosides **221** (159 mg, 20%) as a and methyl glycoside **222** (160 mg, 20%) as a colorless thick gum as well as a mixture of the desired products and the undesired pyranosides **227** and **228** (358mg, 44%).

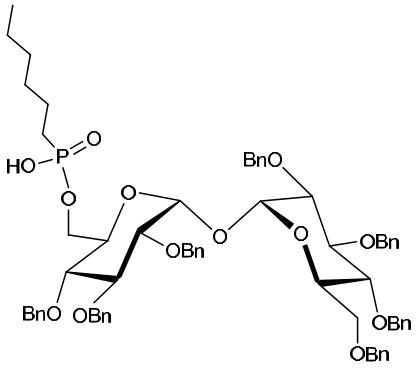
Data for compound **221**α: TLC (DCM:MeOH, 5:1, v/v): R_f = 0.45; $[\alpha]_D^{25}$ +107.2 (c = 1.0, H_2O) [Lit. $[\alpha]_D$ +125.2 (c = 1.0 in H_2O)[52]; ^{1}H NMR (500 MHz, MeOD-D_4) δ ppm 3.43 (3 H, s, -OMe), 3.57 (1 H, dd, $J_{5a,5b}$=11.7 Hz, $J_{4,5a}$ = 7.3, H-5a), 3.70 (1 H, dd, $J_{5a,5b}$ = 11.7 Hz, $J_{4,5b}$ = 3.6 Hz, H-5b), 3.79 (1 H, ddd, $J_{4,5a}$ = 7.3 Hz , $J_{3,4}$ = 6.9 Hz, $J_{4,5b}$ = 3.6 Hz, H-4), 3.92 (1 H, dd, $J_{2,3}$ = 7.6 Hz, $J_{3,4}$ = 6.9 Hz, H-3), 3.97 (1 H, dd, $J_{2,3}$ = 7.6 Hz, $J_{1,2}$ = 4.4 Hz, H-2), 4.76 (1 H, d, $J_{1,2}$ = 4.4 Hz, H-1); ^{13}C NMR (126MHz, MeOD-D_4) δ ppm 55.5 (O-Me), 65.5 (C-5), 76.9 (C-3), 79.1 (C-2), 84.5 (C-4), 104.1 (C-1); MS m/z (ESI^+): 187 $[M+Na]^+$;.

Data for compound **222**β: TLC (DCM:MeOH, 5:1, v/v): R_f = 0.4; $[\alpha]_D^{25}$ -78.0 (c = 0.2, H_2O) [Lit. $[\alpha]_D$ -119 (c = 3.0 in H_2O)[53]; ^{1}H NMR (500 MHz, MeOD-D_4) δ ppm 3.39 (3 H, s, -OMe), 3.66 (1 H, dd,

$J_{5a,5b}$=11.8 Hz, $J_{4,5a}$ = 5.4, H-5a), 3.77 (1 H, dd, $J_{5a,5b}$ = 11.8 Hz, $J_{4,5b}$ = 3.5 Hz, H-5b), 3.85 (1 H, dd, $J_{3,4}$ = 6.3 Hz, $J_{2,3}$ = 3.5 Hz, H-3), 3.93 (1 H, ddd, $J_{3,4}$ = 6.3 Hz , $J_{4,5a}$ = 5.4 Hz, $J_{4,5b}$ = 3.5 Hz, H-4), 3.96 (1 H, dd, $J_{2,3}$ = 3.5 Hz, $J_{1,2}$ = 1.5 Hz, H-2), 4.78 (1 H, d, $J_{1,2}$ = 1.5 Hz, H-1); ^{13}C NMR (126MHz, MeOD-D_4) δ ppm 55.3 (O-Me), 63.1 (C-5), 78.8 (C-3), 83.4 (C-4), 85.5 (C-2), 110.6 (C-1); MS *m/z* (ESI$^+$): 187 $[M+Na]^+$.

8.2.18 Synthesis of compound 60

6-O-(Hexylphosphoryl)-2,3,4,2',3',4,6'-hepta-O-benzyl-α,α-D-trehalose (240)

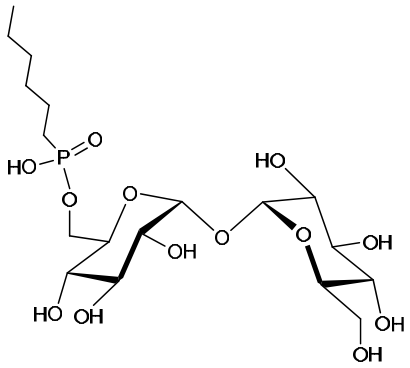


210 (50 mg, 0.53 mmol) was dissolved in anhydrous pyridine (2 mL). To this was added Hexyldichlorophosphate (0.1 mL, 0.5 mmol) and the reaction mixture was stirred for 16 hours at RT upon which time TLC (DCM:MeOH, 5:1) revealed complete conversion to product (R_f 0.75) and disappearance of starting material (R_f 1.0). Reaction mixture was concentrated under reduced pressure and purified by column chromatography (DCM:MeOH, 9:1) to yield the desired product as a white solid (49 mg, 85%).

TLC (DCM:MeOH, 5:1, v/v): R_f = 0.75; $[\alpha]_D^{25}$ + 53.8 (c, 0.79, $CHCl_3$); ^{1}H NMR (500 MHz, CHLOROFORM-*d*) δ ppm 0.80 (3 H, t, *J*=6.8 Hz, $(CH_2)_5C\underline{H}_3$), 1.04 - 1.31 (6 H, m, $(CH_2)_2(C\underline{H}_2)_3CH_3$), 1.42 - 1.63 (4 H, m, $(C\underline{H}_2)_2(CH_2)_3CH_3$), 3.32 (1 H, dd, $J_{6a,6b}$ = 10.4 Hz, $J_{5,6b}$ = 2.2 Hz, H-6b), 3.45 (1 H, dd, $J_{6a,6b}$ = 10.4, $J_{5,6a}$ = 4.6 Hz, H-6a), 3.50 (1 H, dd, *J* = 9.8, 3.9 Hz, H-2 or H-2'), 3.55 (1 H, dd, *J*=9.1, 3.8 Hz, H-2 or H-2'), 3.59 (1 H, t, $J_{3,4}$ = $J_{4,5}$ = 9.8 Hz, H-4 or H-4'), 3.63 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.5 Hz, H-4 or H-4'), 3.91 - 4.19 (6 H, m, H-6'a, H-6'b, H-5, H-5', H-3, H-3'),4.34 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.44 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.50 (1 H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.62 (5 H, br. S, 5 x OCH_2Ph), 4.79 (2 H, d, *J* = 10.4 Hz, 2 x OCH_2Ph), 4.83 (2 H, d, *J* = 11.0 Hz, 2 x OCH_2Ph), 4.96 (2 H, d, *J* = 11.0 Hz, 2 x OCH_2Ph), 5.17 (2 H, d, $J_{1,2}$ = 3.8 Hz, H-1 or H-1'), 5.19 (1 H, d, *J* = 3.8 Hz, H-1 or H-1'), 6.66 - 7.74 (35 H, m, Ar-H); ^{13}C NMR (126 MHz, CHLOROFORM-*d*) δ ppm 14.1 $((CH_2)_5\underline{C}H_3)$ 22.6 $((CH_2)_4\underline{C}H_2CH_3)$, 22.6 $((CH_2)_3\underline{C}H_2CH_2CH_3)$, 22.9

(d $J_{C,P}$ = 130 Hz, $\underline{C}H_2(CH_2)_4CH_3$), 30.5 (d, $J_{C,P}$ = 11.3 Hz, $CH_2\underline{C}H_2(CH_2)_3CH_3$), 31.4 ($(CH_2)_2\underline{C}H_2(CH_2)_2CH_3$), 31.5 ($(CH_2)_3\underline{C}H_2CH_2CH_3$), 62.9 (1 C, d, C-6', $J_{6,P}$ = 7.8 Hz,), 68.1 (C-6), 70.17 (1 C, s), 70.25 (d, $J_{5,P}$ = 5.2 Hz,C-5'), 70.6 (C-5), 72. 7 ($O\underline{C}H_2Ph$), 72.8 ($O\underline{C}H_2Ph$), 73.4 ($O\underline{C}H_2Ph$), 74.9 (2 C, 2 x $O\underline{C}H_2Ph$), 75.4 ($O\underline{C}H_2Ph$), 75.5 ($O\underline{C}H_2Ph$), 77.6 (2 C, C-3, C-3'), 79.2 (C-2 or C-2'), 79.7 (C-2 or C-2'), 81.3 (C-3 or C-3'), 81.6 (C-3 or C-3'), 93.4 (C-1 or C-1'), 93.8 (C-1 or C-1'), 127.1, 127.2, 127.4, 127.4, 127.6, 127.6, 127.8, 127.8,127.9, 127.9, 128.1, 128.3, 128.4, 128.4 (7x OCH_2Ph), 137.8, 138.0, 138.1, 138.2, 138.4, , 138.8, 138.9 (7 x 1 C, 7 x OCH_2Ph); MS *m/z* (ESI^-): 1119.5 $[M-H]^-$; Isotopic distribution: species observed $[M+Na]^+$, peaks observed 1143.50 (100%) 1144.50 (73.5%) 1145.51 (28.7%) 1146.51 (8.4%) 1147.52 (1.6%) peaks calculated 1143.50 (100%) 1144.50 (73.8%) 1145.51 (29.5%) 1146.51 (8.4%) 1147.51 (1.9%).

6-O-(Hexylphosphoryl)-α,α-D-trehalose **(241)**

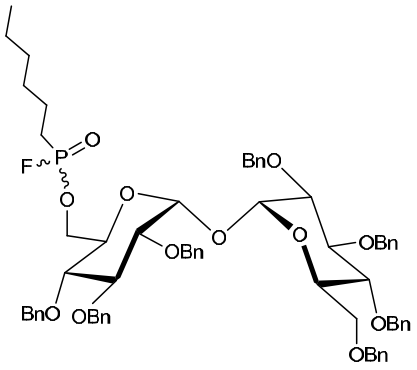


240 (16 mg, 0.014 mmol, 1 eq) was dissolved in (ethanol:water, 1:1, 5 mL) and to this was added 10% Pd/C (50 mg) and reaction was actived by repeated purge flush cycles with hydrogen gas and was then stirred under hydrogen gas overnight at RT. Upon completion TLC (H_2O:*i*-PrOH:EtOAc,1:2:2) revealed complete disappearance of starting material (R_f 1) and appearance of product (R_f 0.6). Palladium was removed by filtration and reaction was concentrated under reduced pressure and lyophilized to afford the desired product as a white, amorphous solid (7 mg, 99%).

TLC (H_2O:*i*-PrOH:EtOAc,1:4:4, v/v/v): R_f = 0.6; $[\alpha]_D^{25}$ + 72.6 (c = 0.34, water); ^{1}H NMR (500 MHz, DEUTERIUM OXIDE) δ ppm 0.79 (3 H, t, $J_{CH3,CH2}$ = 6.9 Hz, $(CH_2)_5C\underline{H_3}$), 1.17 - 1.25 (4 H, m, $(CH_2)_3(C\underline{H_2})_2CH_3$), 1.30 (2 H, quin, $J_{CH2,CH2}$ = 7.2 Hz, $(CH_2)_2$ $C\underline{H_2}(CH_2)_2CH_3$), 1.45 (2 H, m, $CH_2C\underline{H_2}(CH_2)_3CH_3$), 1.62 (2 H, m, $C\underline{H_2}(CH_2)_4CH_3$), 3.36 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.5 Hz, H-4), 3.45 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4'), 3.56 (1 H, dd, $J_{2,3}$ = 9.5 Hz, $J_{1,2}$ = 3.8 Hz, H-2), 3.58 (1 H, dd, $J_{2,3}$ = 9.5, $J_{1,2}$ = 3.8 Hz, H-2'), 3.68 (1 H, dd, $J_{6a,6b}$ = 12.0 Hz, $J_{5,6a}$ = 5.4 Hz, H-6a), 3.74 (1 H, ddd, $J_{4,5}$ = 9.8 Hz, $J_{5,6a}$ = 5.0, $J_{5,6b}$ 1.9 Hz, H-5), 3.77 (1 H, dd, $J_{6a,6b}$ = 12.1 Hz, $J_{5,6a}$ = 2.1 Hz, H-6b), 3.76 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.8 Hz, H-3), 3.78 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.5 Hz, H-3'), 3.84 (1 H, ddd, $J_{4,5}$ = 10.0, $J_{5,6a}$ =

2.9, $J_{5,6b}$ = 1.9 Hz, H-5'), 4.01 - 4.06 (2 H, m, H-6a',H-6b'), 5.11 (1 H, d, $J_{1,2}$ = 3.7 Hz, H-1 or H-1'), 5.11 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1 or H-1'); ^{13}C NMR (126 MHz, DEUTERIUM OXIDE) δ ppm 13.3 ($(CH_2)_5\underline{C}H_3$), 21.8 ($(CH_2)_4\underline{C}H_2CH_3$), 22.3 (1 C, d, $J_{C,P}$ = 5.4 Hz, $(CH_2)_2\underline{C}H_2(CH_2)_2CH_3$), 25,4 (1 C, d, $J_{C,P}$ = 135.4 Hz, $\underline{C}H_2(CH_2)_4CH_3$), 29.6 (1 C, d, $J_{C,P}$ = 16.2 Hz $CH_2\underline{C}H_2(CH_2)_3CH_3$), 30.6 ($(CH_2)_3\underline{C}H_2CH_2CH_3$), 60.5 (C-6), 62.7 (1 C, d, C-6', $J_{6,P}$ = 6.3 Hz), 69.2 (C-4'), 69.7 (C-4), 71.0 (C-2 or C-2'), 71.0 (C-2 or C-2'), 71.1 (d, $J_{5,P}$ = 6.3 Hz,C-5'), 72.2 (1 C, s), 72.3 (1 C, s), 72.5 (1 C, s), 93.3 (C-1 or C-1'), 93.3 (C-1 or C-1');^{31}P NMR (162 MHz, DEUTERIUM OXIDE) δ ppm 32.06; HRMS *m/z* (ESI$^-$): $[M-H]^-$ cald. for $C_{18}H_{35}O_{13}P$, 489.1743; found, 489.1749.

***6-O-(Hexylphosphorylfluoro)-2,3,4,2',3',4,6'-hepta-O-benzyl-α,α-D-trehalose* (242)**

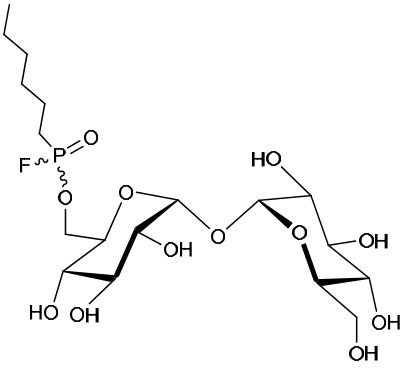


240 (200 mg, 0.17 mmol, 1 eq) and BTFFH (195 mg, 0.61 mmol, 3.6 eq) were dissolved in anhydrous DMF (3 mL) and to this was added DIPEA (0.1 mL, 0.6 mmol, 1 eq) and reaction was stirred overnight at RT. Upon completion the desired product was detected by TLC (petrol:EtOAc, 2:1) (R_f 0.35) with disappearance of starting sugar (R_f 0.0). Reaction was concentrated under reduced pressure and purified by column chromatography (petrol:EtOAc, 2:1) to yield the desired product as a 1:1 mixture of both phosphate stereoisomers and a clear oil (147 mg, 77%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.35; ^{1}H NMR (500 MHz, CHLOROFORM-*d*) δ ppm 0.89 (3 H, t, $J_{CH3,CH2}$ = 6.9 Hz, $(CH_2)_5\underline{C}H_3^{\dagger}$ or $(CH_2)_5\underline{C}H_3$), 0.88 (3 H, t, $J_{CH3,CH2}$ = 6.9 Hz, $(CH_2)_5\underline{C}H_3^{\dagger}$ or $(CH_2)_5\underline{C}H_3$), 1.27-41 (6 H, m, $(CH_2)_2(C\underline{H}_2)_3CH_3^{\dagger}$ and $(CH_2)_2(C\underline{H}_2)_3CH_3$), 1.79 - 1.89 (4 H, m, $(C\underline{H}_2)_2(CH_2)_3CH_3^{\dagger}$ and $(C\underline{H}_2)_2(CH_2)_3CH_3$), 3.38 (1 H, dd, *J*=10.1, 1.6 Hz, H-6b† or H-6b), 3.39 (1 H, dd, *J*=10.1, 1.3 Hz, H-6b† or H-6b), 3.49 - 3.55 (6 H, m, H-6a†, H-6a, H-4†, H-4, H-2'†, H-2'), 3.61 (2 H, dd, *J*=9.8, 3.8 Hz, H-2†, H-2), 3.69 (2 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4'†, H-4'), 3.68 (2 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4†, H-4), 4.05 (2 H, a, t, $J_{2,3}$ = $J_{3,4}$ = 9.5 Hz, H-3†, H-3), 4.06 (2 H, at, $J_{2,3}$ = $J_{3,4}$ = Hz, H-3'†, H-3'), 4.09 - 4.21 (8 H, m, H-5†, H-5'†, H-5, H-5', H-6a'†, H-6b'†, H-6a', H-6b'), 4.40 (1 H, d, *J* = 12.0 Hz, 1 x O$C\underline{H}_2Ph^{\dagger}$), 4.40 (1 H, d, *J* = 12.0 Hz, 1 x O$C\underline{H}_2$Ph), 4.47 (1 H, d, *J* = 10.7 Hz, 1 x O$C\underline{H}_2Ph^{\dagger}$), 4.48 (1 H, d, *J* = 10.9 Hz, 1 x O$C\underline{H}_2$Ph), 4.55 (2 H, d, *J* = 12.0 Hz, 1 x O$C\underline{H}_2Ph^{\dagger}$ and 1 x O$C\underline{H}_2$Ph), 4.59 (2 H, d, *J* = 12.0 Hz, 1 x O$C\underline{H}_2Ph^{\dagger}$ and 1 x O$C\underline{H}_2$Ph), 4.63 (2 H, d, *J* = 12.0 Hz, 1 x

OC$\underline{H}_2$Ph† and 1 x OC$\underline{H}_2$Ph) 4.64 (2 H, d, *J*=12.3 Hz, 1 x OC$\underline{H}_2$Ph† and 1 x OC$\underline{H}_2$Ph), 4.72 (2 H, d, *J* = 11.7 Hz, 1 x OC$\underline{H}_2$Ph† and 1 x OC$\underline{H}_2$Ph), 4.74 (1 H, d, *J* = 11.7 Hz, 1 x OC$\underline{H}_2$Ph), 4.82 (2 H, d, *J* = 10.7 Hz, 1 x OC$\underline{H}_2$Ph† and 1 x OC$\underline{H}_2$Ph), 4.87 (2 H, d, *J* = 10.7 Hz, 1 x OC$\underline{H}_2$Ph† and 1 x OC$\underline{H}_2$Ph), 4.87 (1 H, d, *J* = 11.0 Hz, 1 x OC$\underline{H}_2$Ph†), 4.88 (1 H, d, *J* = 11.3 Hz, 1 x OC$\underline{H}_2$Ph), 4.89 (2 H, d, *J* = 10.7 Hz, 1 x OC$\underline{H}_2$Ph† and 1 x OC$\underline{H}_2$Ph), 4.90 (2 H, d, *J* = 11.0 Hz, 1 x OC$\underline{H}_2$Ph† and 1 x OC$\underline{H}_2$Ph), 4.98 (1 H, d, *J* = 10.7 Hz, 1 x OC$\underline{H}_2$Ph), 4.99 (1 H, d, *J* = 11.0 Hz, 1 x OC$\underline{H}_2$Ph†), 5.01 (2 H, d, *J* = 10.7 Hz, 1 x OC$\underline{H}_2$Ph, 1 x OC$\underline{H}_2$Ph†), 5.19 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1' or H-1'†),), 5.21 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1' or H-1'†), 5.22 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1 or H-1†), 5.22 (1 H, d, *J* = 3.5 Hz, H-1 or H-1†), 7.20 - 7.37 (35 H, m, Ar-H, Ar-H†); ^{13}C NMR (126 MHz, CHLOROFORM-*d*) δ ppm 14.0, 14.0 ((CH$_2$)$_5\underline{C}$H$_3^{\dagger}$ and (CH$_2$)$_5\underline{C}$H$_3$), 21.8, 21.8 ((CH$_2$)$_4\underline{C}$H$_2$CH$_3^{\dagger}$ and (CH$_2$)$_4\underline{C}$H$_2$CH$_3$), 22.3, 22.3. (CH$_2\underline{C}$H$_2$(CH$_2$)$_3$CH$_3^{\dagger}$ and CH$_2\underline{C}$H$_2$(CH$_2$)$_3$CH$_3$), 24.2 (1 C, dd, $J_{C,P}$ = 144.0, $J_{C,F}$ = 22.8 Hz, C$\underline{H}_2$(CH$_2$)$_4$CH$_3$), 24.3 (1 C, dd, $J_{C,P}$ = 143.1, $J_{C,F}$ = 22.9 Hz, C$\underline{H}_2$(CH$_2$)$_4$CH$_3^{\dagger}$), 29.7, 29.9 ((CH$_2$)$_2\underline{C}$H$_2$(CH$_2$)$_2$CH$_3^{\dagger}$ and (CH$_2$)$_2\underline{C}$H$_2$(CH$_2$)$_2$CH$_3$), 30.0, 31.1 (CH$_2$)$_3\underline{C}$H$_2$CH$_2$CH$_3^{\dagger}$ and (CH$_2$)$_3\underline{C}$H$_2$CH$_2$CH$_3$), 64.9 (1 C, d, $J_{P,C6}$ = 36.7 Hz, C-6' phos†), 64.9 (1 C, d, $J_{P,C6}$ = 36.2 Hz, C-6' phos), 68.1, 68.1 (C-6† and C-6), 69.5 (1 C, d, $J_{P,C5}$ = 7.0 Hz, C-5'†), 69.5 (1 C, d, $J_{P,C5}$ = 6.7 Hz, C-5'), 70.6, 70.8 (C-5†, C-5'), 72.7, 72.7, 73.0, 73.1, 73.5, 73.5, 75.1, 75.1, 75.2, 75.2, 75.6, 75.6, 75.7, 75.7 (7 x O$\underline{C}$H$_2$Ph†, 7 x O$\underline{C}$H$_2$Ph), 76,8, 76.8 (C-4†, C-4, C-4'† and C-4'), 77.7, 77.7, 79.4, 79.4 (C-2'†, C-2', C-2† and C-2), 81.4, 81.5, 81.7, 81.8 (C-3'†,C-3', C-3†, C-3), 93.9, 94.1, 94.4, 94.6 (C-1'†, C-1',C-1† and C-1), 127.3, 127.4, 127.5, 127.5, 127.7, 127.7, 127.8, 128.0, 128.1, 128.1, 128.3, 128.4, 128.5 (14 x OCH$_2$Ph), 137.8, 137.8, 137.9, 137.9, 137.9, 138.0, 138.0, 138.0, 138.2, 138.2, 138.5, 138.5, 138.8, 138.8 (14 x 1 C, 14 x OCH$_2$Ph); ^{31}P NMR (162 MHz, CHLOROFORM-*d*) δ ppm 32.2 (1 P, d, $J_{P,F}$ = 1072.1 Hz), 31.8 (1 P, d, $J_{P,F}$ = 1072.4 Hz); ^{19}F NMR (377 MHz, CHLOROFORM-*d*) δ ppm -64.5 (1 F, d $J_{P,F}$ = 1072.4Hz), -63.9 (1, F, d, $J_{P,F}$ = 1072.4 Hz); MS *m/z* (ESI^{+}): 1145.5 [M+Na]$^{+}$; Isotopic distribution: species observed [M+Na]$^{+}$ peaks observed 1145.50 (100%) 1146.50 (74.5%) 1147.50 (29.0%) 1148.50 (4.0 %) 1149.49 (0.4%) peaks calculated 1145.50 (100%) 1146.50 (73.8%) 1147.50 (29.3%) 1148.51 (8.2%) 1149.51 (1.8%).

6-O-(Hexylphosphorylfluoro) -α,α-D-trehalose **(60)**



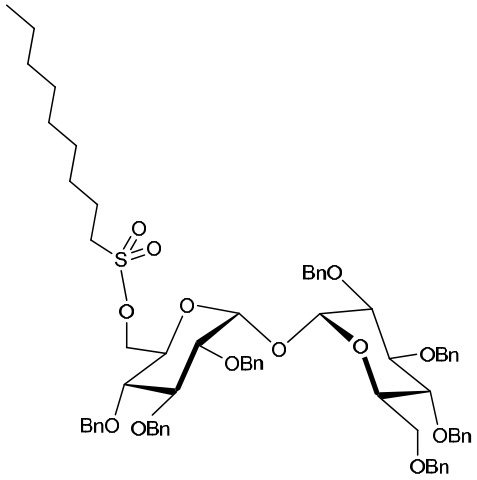
242 (56 mg, 0.049 mmol, 1 eq) was dissolved in ethanol (5 mL) and to this was added 10% Pd/C (100 mg) and reaction was actived by repeated purge flush cycles with hydrogen gas and was then stirred under hydrogen gas overnight at RT. Upon completion TLC (EtOAc:MeOH, 1:1) revealed complete disappearance of starting material (R_f 1) and appearance of product (R_f 0.55). Palladium was removed by filtration and reaction was concentrated under reduced pressure and lyophilized to afford the desired product as a white, amorphous hygroscopic solid (20 mg, 81%). Desired product was obtained as a white amorphous solid and a 1:1 mixture of both phosphoryl stereoisomers.

TLC (EtOAc:MeOH, 2:1, v/v): R_f = 0.55; ^{1}H NMR (500 MHz, DEUTERIUM OXIDE) δ ppm 0.80 (6 H, t, $J_{CH3,CH2}$ = 6.3 Hz ($(CH_2)_5C\underline{H}_3^{\dagger}$ and $(CH_2)_5C\underline{H}_3$), 1.52 - 1.65 (8 H, m, $(CH_2)_3(C\underline{H}_2)_2CH_3^{\dagger}$ and $(CH_2)_3(C\underline{H}_2)_2CH_3^{\dagger}$), 1.69 - 1.80 (4 H, m, $CH_2\underline{C}H_2(CH_2)_3CH_3^{\dagger}$ and $CH_2\underline{C}H_2(CH_2)_3CH_3$), 1.99 - 2.15 (4 H, m, $C\underline{H}_2(CH_2)_4CH_3^{\dagger}$, $C\underline{H}_2(CH_2)_4CH_3$), 3.36 (2 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.5 Hz, H-4†, H-4), 3.43 (2 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4'†, H-4'), 3.53 - 3.59 (4 H, m, H-2†, H-2',H-2† and H-2), 3.67 (2 H, m, $J_{6a,6b}$ = 12.0 Hz, $J_{5,6a}$ = 5.0 Hz, H-6a†, H-6a), 3.69 - 3.89 (8 H, m, H-6b†, H-6b, H-5, H-5, H-3'†, H-3', H-3†, H-3), 3.89 - 3.95 (2 H, m, H-5'†, H-5'), 4.33 - 4.44 (4 H, m, H-6a'†, H-6a', H-6a†, H-6a), 5.10 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1†), 5.11 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1), 5.13 (2 H, d, $J_{1,2}$ = 3.8 Hz, H-1'†, H-1'); ^{13}C NMR (126 MHz, DEUTERIUM OXIDE) δ ppm 13.3, 13.3 ($(CH_2)_5\underline{C}H_3^{\dagger}$ and $(CH_2)_5\underline{C}H_3$), 20.9 (1 C, d, $J_{C,P}$ = 8.0 Hz, $CH_2\underline{C}H_2(CH_2)_3CH_3^{\dagger}$), 20.9 (1 C, d, $J_{C,P}$ = 8.0 Hz, $CH_2\underline{C}H_2(CH_2)_3CH_3$), 21.7, 21.8 ($(CH_2)_4\underline{C}H_2CH_3^{\dagger}$ and $(CH_2)_4\underline{C}H_2CH_3$) 22.6 (1 C, dd, $J_{C,P}$ = 139.2 Hz, $J_{C,F}$ = 19.1 Hz, $C\underline{H}_2(CH_2)_4CH_3^{\dagger}$), 22.5 (1 C, dd, $J_{C,P}$ = 139.2 Hz, $J_{C,F}$ = 18.1 Hz $C\underline{H}_2(CH_2)_4CH_3$), 28.8 (1 C, d, $J_{C,P}$ =2.9 Hz, $(CH_2)_2\underline{C}H_2(CH_2)_2CH_3^{\dagger}$), 28.9 (1 C, d, $J_{C,P}$ =2.9 Hz, $(CH_2)_2\underline{C}H_2(CH_2)_2CH_3$), 30.3, 30.3 ($(CH_2)_3\underline{C}H_2CH_2CH_3^{\dagger}$ and $(CH_2)_3\underline{C}H_2CH_2CH_3$), 60.5, (C-6†, C-6), 65.8 (1 C, d, $J_{C,P}$ = 5.04 1 C, C-6'†), 65.8 (1 C, d, $J_{C,P}$ = 5.04 1 C, C-6'), 69.0, 69.0, 69.1, 69.1 (C-4'†, C-4',C-4† and C-4), 70.3, 70.4 (C-5'† and C-5'), 70.9, 70.9, 71.0, 71.0 (C-2'†, C-2',C-2† and C-2), 72.2, 72.2 (C-5† and C-5), 72.3, 72.3, 72.5, 72.5 (C-3'†, C-3',C-3† and C-3), 93.4, 93.4, 93.5, 93.5 (C-1'†, C-1',C-1† and C-1); ^{19}F NMR

(377 MHz, DEUTERIUM OXIDE) δ ppm -66.5 (1 F, d, $J_{F,P}$ = 1077.3 Hz), -66.7 (1 F, d, $J_{F,P}$ = 1079.6 Hz); ^{31}P NMR (162 MHz, DEUTERIUM OXIDE) δ ppm 37.9 (1 P, d, $J_{P,F}$ = 1080.0 Hz), 38.2 (1 P, d, $J_{P,F}$ = 1076.0 Hz); HRMS *m/z* (ESI$^+$): [M+Na] $^+$ calcd. for $C_{18}H_{34}FO_{12}PNa^+$, 515.1664; found, 515.1666.

8.2.19 Synthesis of octanesulfonyl-trehalose 61

6-O-Octanesulfonyl-2,3,4,2',3',4,6'-hepta-O-benzyl-α,α-D-trehalose (239)

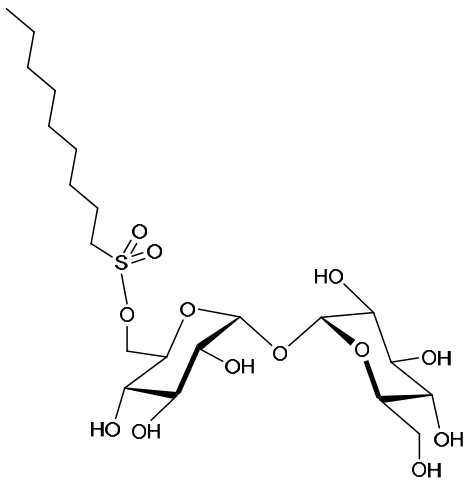


210 (200 mg, 0.22 mmol, 1 eq) was dissolved in anhydrous pyridine (2 mL) and to this was added octanesulfonyl chloride (0.2 mL, 5 eq), which, upon addition caused a yellow color change of the reaction mixture. Reaction was stirred for 2 hours at RT upon which time TLC (petrol:EtOAc, 2:1) indicated conversion from starting material (R_f 0.1) to product (R_f 0.6). Reaction was washed with satd $NaHCO_3$ solution and concentrated *in vacuo*. Column chromatography with a gradient of solvents (petrol:EtOAc, 6:1 to 1:1) yielded the desired product as a clear oil (94 mg, 40%).

TLC (petrol:EtOAc, 2:1, v/v): R_f = 0.6; $[\alpha]_D^{25}$ + 73.7 (c, 0.57, $CHCl_3$); ^{1}H NMR (500 MHz, CHLOROFORM-*d*$_3$) δ ppm 0.87 - 0.91 (5 H, m, $OSO_2(CH_2)_6$ $\underline{CH_2CH_3}$), 1.21 - 1.40 (10 H, m, $OSO_2CH_2CH_2(C\underline{H_2})_5CH_3$), 1.78 (2 H, quint, $J_{CH2,CH2}$ = 8.2 Hz, $OSO_2CH_2C\underline{H_2}(CH_2)_5CH_3$), 2.99 (2 H, td, $J_{CH2,CH2}$ = 8.2, $J_{CH2,S}$ = 1.3 Hz, $OSO_2C\underline{H_2}(CH_2)_6CH_3$), 3.39 (1 H, dd, $J_{6a,6b}$ = 11.0 Hz, $J_{5,6b}$ = 1.6, H-6b), 3.51 - 3.56 (3 H, m, H-6a, H-4, H-2'), 3.61 (1 H, dd, $J_{2',3'}$ = 9.6 Hz, $J_{1,2}$ = 3.6 Hz, H-2), 3.69 (1 H, at, $J_{3,4}$ = $J_{4,5}$ = 9.6 Hz, H-4'), 3.96 (1 H, dd, $J_{6a',6b'}$ = 11.0 Hz, $J_{5',6b'}$ 1.6 Hz, H-6b'), 4.02 (1 H, at, $J_{2,3}$ = $J_{3,4}$ = 9.5 Hz, H-3'), 4.03 (1 H, at, $J_{2,3}$ = $J_{3,4}$ 9.3 Hz, H-3), 4.10 - 4.17 (3 H, m, H-5, H-6a'), 4.21 (1 H, dt, $J_{4,5}$ = 9.8, $J_{5,6a}$ = 2.2, $J_{5,6b}$ = 1.6 Hz, H-5), 4.40 (1 H, d, *J* = 12.3 Hz, 1 x OCH_2Ph), 4.47 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.55 (1 H, d, *J* = 12.0 H 1 x OCH_2Ph), 4.62 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.63 (2 H, d, *J* = 12.0 Hz, 2 x OCH_2Ph), 4.66 1 H, d, *J* = 10.2 Hz, 1 x OCH_2Ph), (4.76 (1

H, d, *J* = 12.0 Hz, 1 x OCH_2Ph), 4.82 (2 H, d, *J* = 10.7 Hz, 2 x OCH_2Ph), 4.87 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 4.89 (1 H, d, *J* = 10.6 Hz, 1 x OCH_2Ph), 4.89 (1 H, d, *J* = 10.7 Hz, 1 x OCH_2Ph), 4.99 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 5.01 (1 H, d, *J* = 11.0 Hz, 1 x OCH_2Ph), 5.17 (1 H, d, $J_{1,2}$ = 3.8 Hz, H-1), 5.19 (1 H, d, $J_{1',2'}$ = 3.8 Hz, H-1'), 7.17 - 7.38 (35 H, Ar-H)**;** ^{13}C NMR (126 MHz, CHLOROFORM-d_3) δ ppm 14.06 ($OSO_2(CH_2)_7\underline{C}H_3$), 22.57, 23.3, 26.8, 28.1, 28.9, 28.9, 31.7 ($OSO_2CH_2CH_2(\underline{C}H_2)_5CH_3$), 50.4 ($OSO_2\underline{C}H_2(CH_2)_6CH_3$), 67.5 (C-6), 68.09 (C-6'), 68.97 (C-5'), 70.79 (C-5), 72.72 (OCH_2Ph), 73.18 (OCH_2Ph), 73.49 (OCH_2Ph), 75.11 (OCH_2Ph), 75.20 (OCH_2Ph), 75.6 (OCH_2Ph), 75.6 (OCH_2Ph), 76.7 (C-4), 77.4 (C-4'), 79.3 (C-2'), 79.4 (C-2), 81.5 (C-3), 81.8 (C-3'), 94.2 (C-1'), 94.7 (C-1), 127.34, 127.4, 127.5, 127.6, 127.6 127.7, 127.8, 127.9, 127.9, 128.0, 128.1, 128.1, 128.3, 128.4, 128.4 (7 x OCH_2Ph), 137.74, 137.88, 137.9, 138.0, 138.2, 138.5, 138.7 (7 x 1C, 7 x OCH_2Ph); IR (thin film): ν = 3062.3, 3029.5, 2923.7, 2854.3, 1604.4, 1496.5, 1453.8, 1360.0, 1209.0, 1157.7, 1097.1, 1071.1(sulfonyl), 1027.6, 733.9; MS *m/z* (ESI^+): 1171.5 $[M+Na]^+$; Isotopic distribution: species observed $[M+Na]^+$, peaks observed 1171.5137 (100%), 1172.5125 (56.4 %) 1173.5205 (20.5 %), 1174.5276 (5.1%), 1175.5508 (2.6%). peaks calculated 1171.5212 (100%), 1172.5245 (76.8%), 1173.5277 (36.3%), 1174.5311 (12.7%), 1175.5238 (3.5%), 1176.5271 (0.8%).

6-O- Octanesulfonyl -α,α-D-trehalose (61)

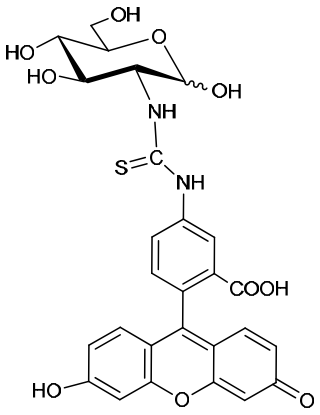


239 (38 mg, 0.033 mmol) was dissolved in 20 mL ethanol and to this was added 10% Pd/C (100 mg). Palladium was activated by repeated purge, flush sequence with hydrogen gas. Reaction was stirred for 16 hours at RT upon which time complete conversion to the desired product was detected by TLC (EtOAc:MeOH, 2:1) indicated complete conversion from starting material (R_f 1) to product (R_f 0.4). Palladium was removed by filtration and reaction was concentrated *in vacuo* and lyophilized to yield the desired product as a white amorphous solid (12.2 mg, 71%).

TLC (EtOAc:MeOH, 2:1, v/v): R_f = 0.4; $[\alpha]_D^{25}$ = 106 (c = 0.2, MeOH); ^{1}H NMR (500 MHz, MeOD) δ ppm 0.92 (3 H, t, $J_{CH2,CH2}$ = *6.2 Hz*, $OSO_2(CH_2)_7\underline{CH_3}$), 1.27 - 1.39 (10 H, m, $OSO_2CH_2CH_2(C\underline{H_2})_5CH_3$), 1.84 (2 H, quin, *J*=7.7 Hz, $OSO_2(CH_2)_6C\underline{H_2}CH_3$), 3.24 (2 H, dt, $J_{CH2,CH2}$ = 12.1, $J_{CH2,S}$ = 3.7 Hz, $OSO_2C\underline{H_2}(CH_2)_6CH_3$), 3.50 (1 H, dd, $J_{2,3}$ = 9.8 $J_{1,2}$ = 3.6 Hz, H-2 or H-2'), 3.51 (1 H, dd, $J_{2,3}$ = 9.7 $J_{1,2}$ = 3.5 Hz, H-2 or H-2'), 3.70 (1 H, dd, dd, $J_{6a,6b}$ = 11.8 Hz, $J_{5,6a}$ = 5.2 Hz, H-6a), 3.77 - 3.84 (3 H, m, H-3, H-3', H-6b), 3.86 (1 H, ddd, $J_{4,5}$ = 12.0, $J_{5,6a}$ = 5.4, $J_{5,6b}$ = 2.2 Hz, H-5), 4.07 (1 H, ddd, $J_{4,5}$ = 10.1 Hz, $J_{5,6a}$ = 4.4 Hz, $J_{5,6b}$ = 1.9 Hz, H-5'), 4.39 (1 H, dd, $J_{6a,6b}$ = 11.3 Hz, $J_{5,6b}$ = 4.7 Hz, H-6a'), 4.46 (1 H, dd, $J_{6a,6b}$ = 11.0 Hz, $J_{5,6b}$ = 1.9 Hz, H-6b'), 5.10 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1 or H-1'), 5.14 (1 H, d, $J_{1,2}$ = 3.5 Hz, H-1 or H-1'); ^{13}C NMR (126 MHz, MeOD) δ ppm 14.4 ($OSO_2(CH_2)_7\underline{C}H_3$), 23.7, 24.7, 29.2, 30.2, 30.2, 32.9($OSO_2CH_2(\underline{C}H_2)_6CH_3$), 50.8 ($OSO_2\underline{C}H_2(CH_2)_6CH_3$), 62.6 (C-6', SO), 70.3 (C-6), 71.4 (C-5'), 71.5 (C-4'), 71.9 (C-4), 73.1 (C-2 or C-2'), 73.1 (C-2 or C-2'), 73.9 (C-5), 74.4 (C-3 or C-3'), 74.6 (C-3 or C-3'), 95.3 (C-1 or C-1'), 95.43 (C-1 or C-1'); HRMS *m/z* (ESI$^+$): $[M+Na]^+$ calcd. for $C_{20}H_{38}O_{13}SNa^+$, 541.1931; found, 541.1923.

8.2.20 Synthesis of FITC-glucose 254.

2-deoxy-2-(N'-(Fluorescein-5-yl)-thioureido) -D-glucose **(254)**



Glucosamine hydrochloride (21.5 mg, 0.1 mmol) and fluorescein isothiocyanate (FITC, 58.4 mg, 0.15 mmol) were dissolved in a mixture of 75 mM solution of $NaHCO_3$ (3mL) and of acetonitrile (1.5 mL) . The mixture was heated at + 55 °C for 2 hours until TLC (H_2O:*i*-PrOH:EtOAc,1:4:8) indicated consumption of starting materials and formation of two products α (R_f 0.64) and β (R_f 0.52). The mixture was concentrated under reduced pressure and purified by silica column chromatography (H_2O:*i*-PrOH:EtOAc,1:4:8) to give the title compound (29.7 mg, 0.052 mmol, 52%) as an orange solid.

TLC (H_2O:*i*-PrOH:EtOAc,1:4:8, v/v/v): **254αα** R_f = 0.64; **254β** R_f = 0.52; $[\alpha]_D^{25}$ = +32.4 (c = 0.25, CH_3OH); ^{1}H NMR (500 MHz, DMSO-d_6): δ **254α** 3.2-3.6 (3H, m, H-3α, H-4, H-6b), 3.62 (1H, m, H-6a), 3.72-3.82 (2H, m, H-5, H-2), 5.72 (1H, bs, H-1), 6.60 (4H, m, H-10' and H-12'), 6.70 (2H, m, H-9'), 7.29 (1H, d, $J_{7'-8'}$=8.3 Hz, H-7'), 8.03 (1H, dd, $J_{8'-7'}$=8.3 Hz, $J_{8'-2'}$=2.0Hz, H-8'), 8.19 (1H, d, $J_{2'-8'}$=2.0Hz, H-2'), 8.80 (1H, s, NHCS attached to C-2); **254β** 3.2-3.6 (H-3α, H-4, H-6b), 3.62 (1H, m, H-6a), 3.72-3.82 (3H, m, H-5, H-2, H-3), 4.19 (1H, d, J_{4-3}=2.4 Hz, H-4), 4.25 (1H, d, J_{2-1}=6.3 Hz, H-2), 6.16 (1H, d, J_{1-2}=6.3 Hz, H-1), 6.60 (4H, m, H-10' and H-12'), 6.70 (2H, m, H-9'), 7.28 (1H, d, $J_{7'-8'}$=8.3 Hz, H-7'), 8.02 (1H, dd, $J_{8'-7'}$=8.3 Hz, $J_{8'-2'}$=2.0Hz, H-8'), 8.17 (1H, d, $J_{2'-8'}$=2.0Hz, H-2'), 9.46 (1H, s, NHCS attached to C-2); ^{13}C NMR (126 MHz, DMSO-d6): δ **254α** 63.2 (C-6), 65.2 (C-2), 68.9 (C-5), 70.0 (C-4), 71.3 (C-3), 83.1 (C-5'), 87.0 (C-1), 102.2 (C-12'), 109.4 (C-14'), 112.7 (C-10'), 120.6 (C-2'), 123.7 (C-3'), 126.1 (C-7'), 129.1 (C-9'), 133.1 (C-8'), 140.8 (C-1'), 149.4 (C-6'), 151.8 (C-13'), 159.5 (C-11'), 168.5 (C=O C-4'), 179.2 (NHC=SNH); δ **254β** 63.6 (C-6), 65.2 (C-2), 68.1 (C-5), 73.5 (C-4), 79.6 (C-3) , 83.1 (C-5'), 94.4 (C-1), 102.3 (C-12'), 109.3 (C-14'), 112.7 (C-10'), 120.6 (C-2'), 123.9 (C-3'), 126.2 (C-7'), 129.1 (C-9'), 133.1 (C-8'), 140.7 (C-1'), 149.7 (C-6'), 151.8 (C-13'), 159.5 (C-11'), 168.3 (C=O C-4'), 180.4 (NHC=SNH); IR (KBr):3333 (OH), 1595 , 1459, 1305, 1210, 1172, 1117, 850 cm^{-1}; HRMS (m/z): $[M-H]^-$ calcd. for $C_{27}H_{23}N_2O_{10}S^-$, 567.1079; found, 567.1087.

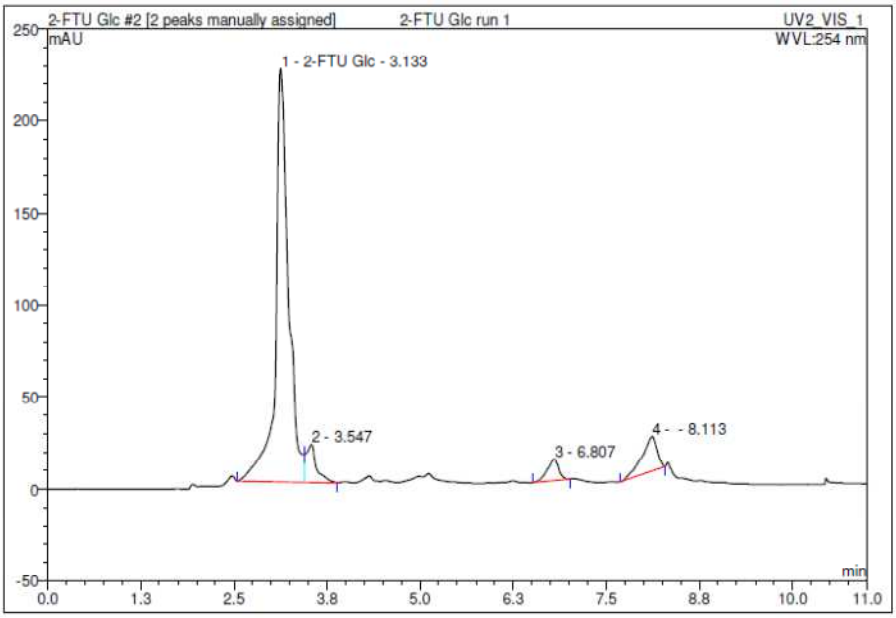


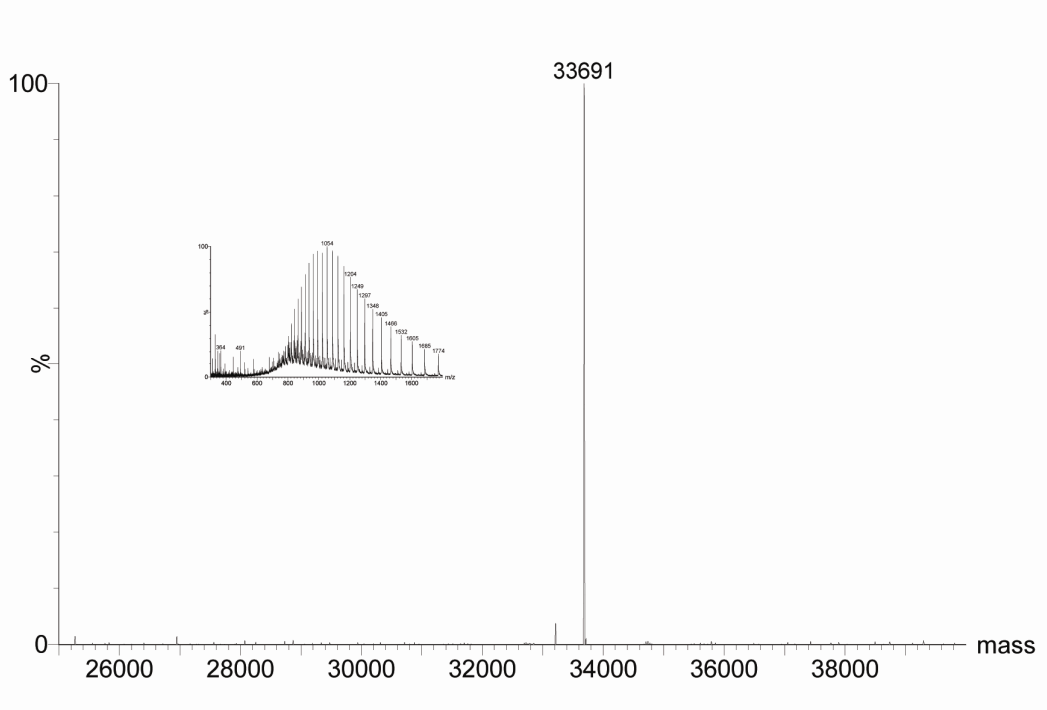
Figure 8.17. Preparative HPLC trace for the purification of 254. Conditions: Phenomenex Synergi Hydro C18 column (150 mm x 21.2 mm, 4 µm), eluting with water with 0.1 % TFA (A) and acetonitrile (B) eluting at a flow rate of 20 mL/min; Gradient: 0 min B=%0, 1 min B=20%, 5.5 min B=40%, 6 min B=100%, 8 min B=0%; Compound 28 elutes at 3.1 min (labeled 254) while starting materials FITC elute at 6.8 and 8.1 min.

8.2.21 Protein Modifications

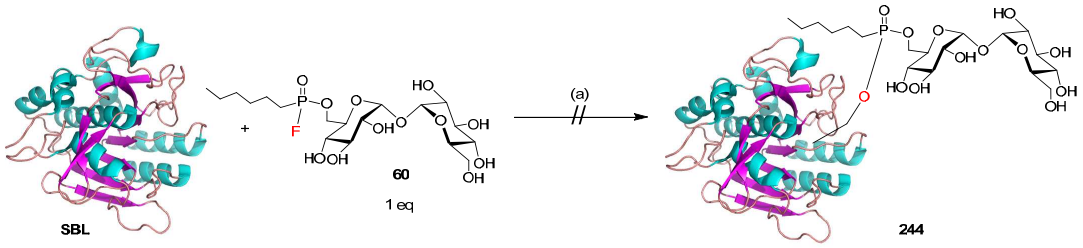
8.2.21.1 Reaction of Ag85C with 60

Ag85C-124-phosfluortre (243)

To antigen 85 (200 µl, 1.6 mg/mL, 9.6 µmoles, 1 eq) in TEA buffer (20 mM, pH 7.2) was added **60** as a solution in water (12 µl, 0.5 mg/mL, 12 µmoles, 1.7 eq). Reaction was shaken at RT in an eppendorf tube for 80 minutes. Then 50 µl of the reaction mixture was evaluated by LCT MS, which revealed complete disappearance of starting protein (mass 33217) and appearance of acylated protein (mass 33691).



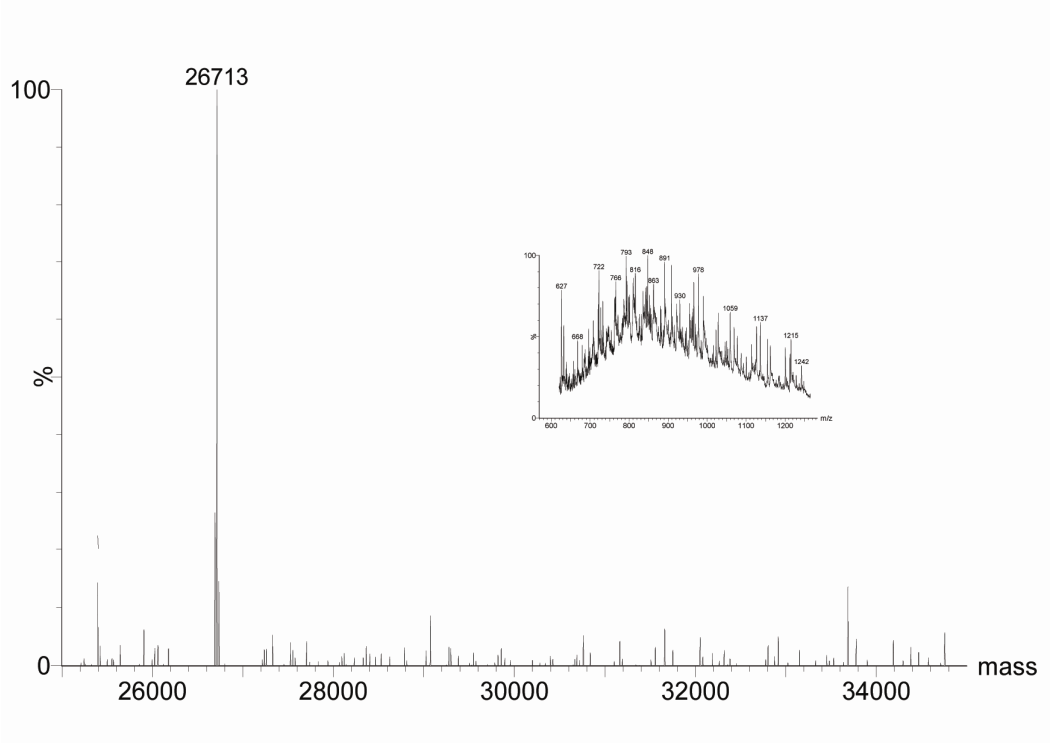
8.2.21.2 Reaction of Subtilisin Bacillus lentus (SBL) with 60



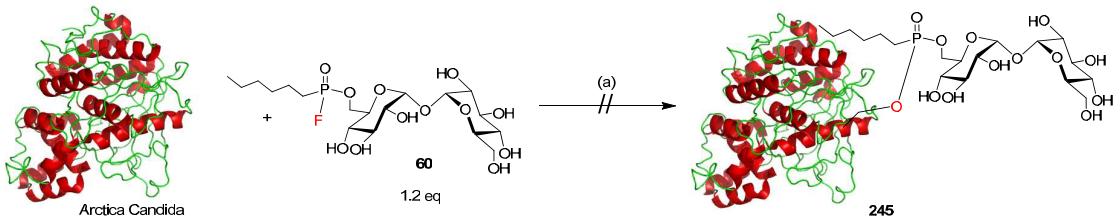
Scheme 8.2. SBL shows no reactivity towards fluorophosphonate 60 (a) 20 mM TEA, pH 7.2, RT, 20 min.

To SBL (200 µl, 0.7 mg/mL, 5.2 µmoles, 1 eq) in TEA buffer (20 mM, pH 7.2) was added **60** as a solution in water (5 µl, 0.5 mg/mL, 5 µmoles, 1 eq). Reaction was shaken at RT in an eppendorf tube for 6 hours. Then 50 µl of the reaction mixture was evaluated by LCT MS, which revealed no reaction of the starting protein (mass 26714).

SBL treated with 60



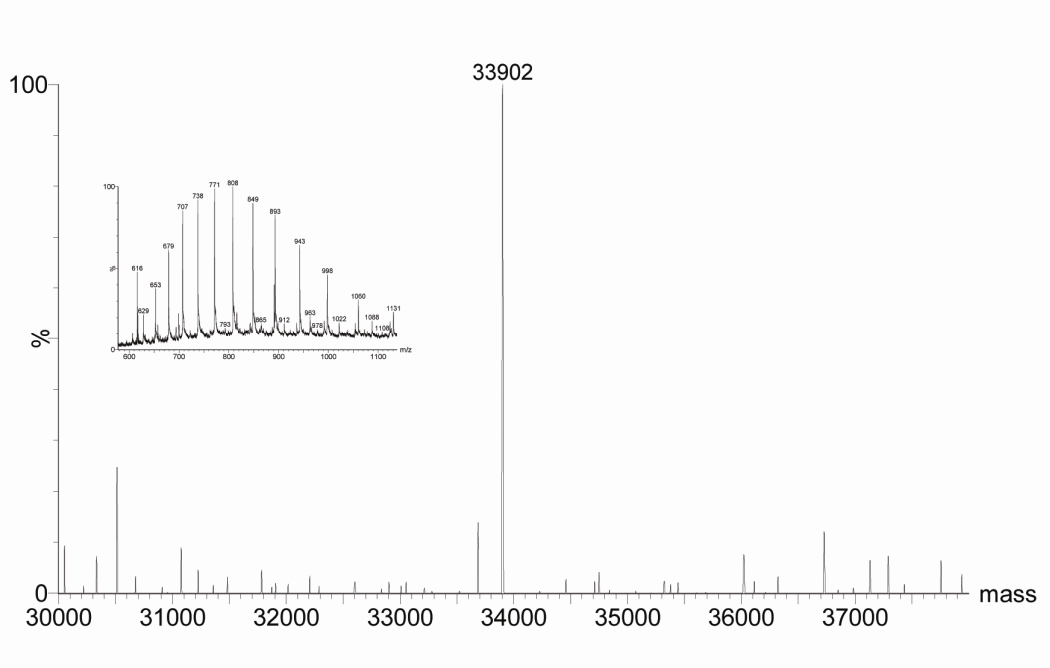
8.2.21.3 Reaction of Arctica candida lipase with 60



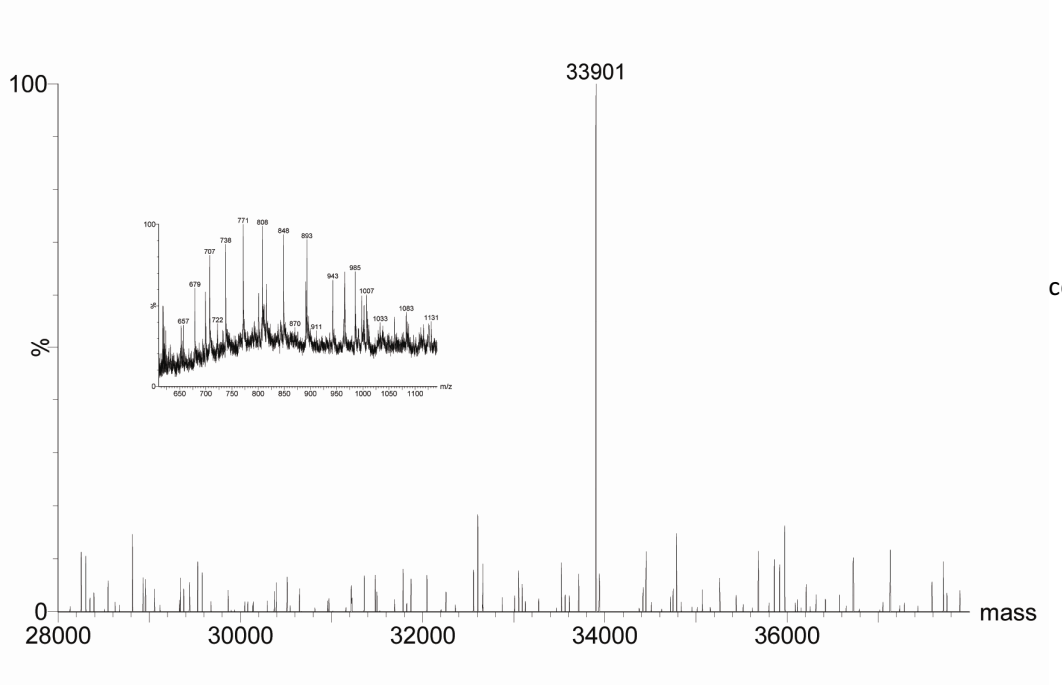
Scheme 8.3. Arctica Candida did not react with fluorophosphonate 60 (a) 20 mM TEA, pH 7.2, 20 min.

To Arctica candida lipase (200 µl, 1.4 mg/mL, 8.3 µmoles, 1 eq) in TEA buffer (20 mM, pH 7.2) was added 60 as a solution in water (10 µl, 0.5 mg/mL, 10 µmoles, 1 eq). Reaction was shaken at RT in an eppendorf tube for 35 minutes. Then 50 µl of the reaction mixture was evaluated by LCT MS, which revealed no reaction of the starting protein (mass 33902).

Unreacted Arctica candida

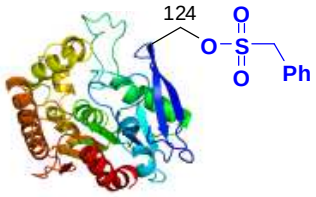


Arctica candida treated with 60



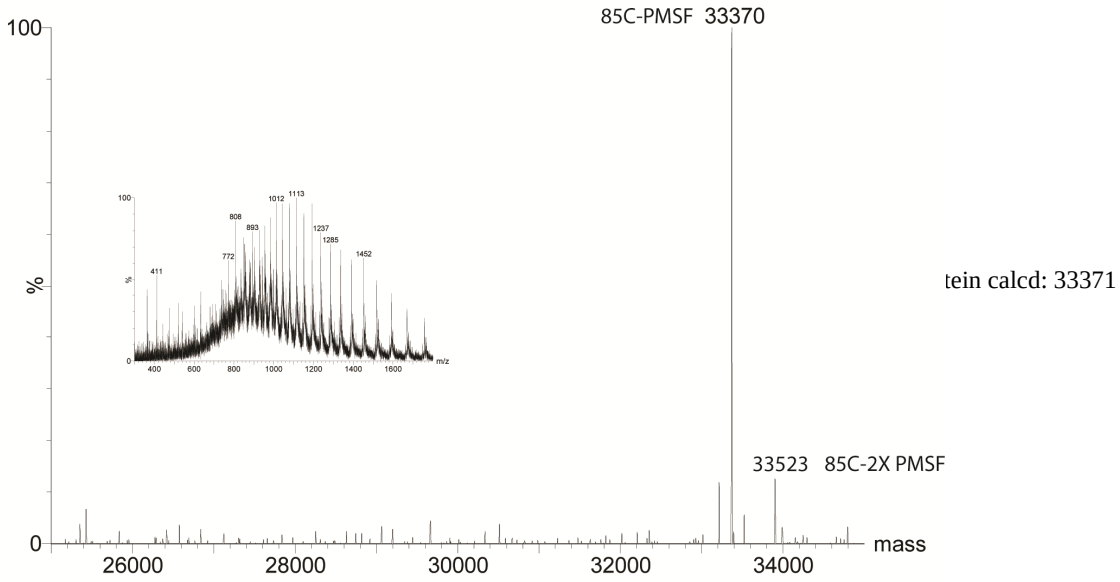
8.2.21.4 Reaction of Ag85C with PMSF 230

Ag85C-124PMS (230)



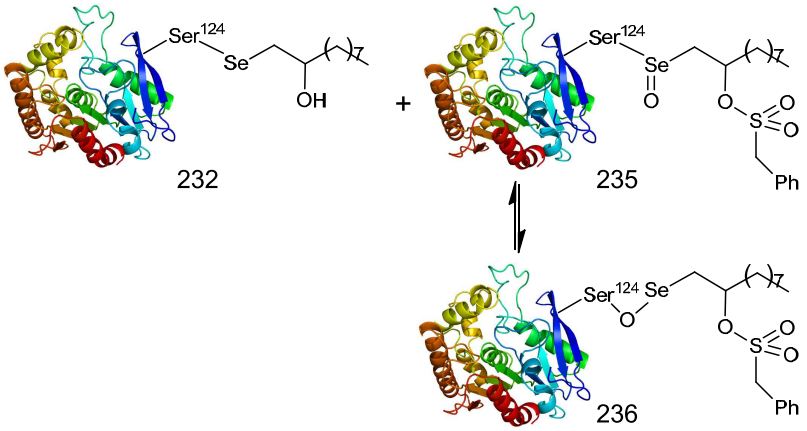
To a solution of Ag85C (1.5 mL of 0.8 mg/mL, 36.1 nmol) in 20 mM TEA buffer (pH 7.2), phenylmethanesulphonyl fluoride (PMSF) in acetonitrile (1.92 mg in 150 µL) was added and the resulting mixture vortexed for 30 seconds. After 22 h of shaking at 22 ºC, the reaction was analyzed directly by LC–MS and >95% conversion to ag85C-124PMS **230** (calculated mass, 33371; observed mass, 33375) was observed. Small molecules were removed from the reaction mixture by loading the sample onto a PD10 desalting column (GE Healthcare) previously equilibrated with 10 column

volumes of pH 7.2 TEA buffer (20 mM) and eluting with 3.5 mL of the same buffer. The solution was concentrated on a vivaspin membrane concentrator (10 kDa molecular weight cut off) to 1.5 mL. The collected sample was flash frozen with liquid nitrogen and stored at –80 ºC.

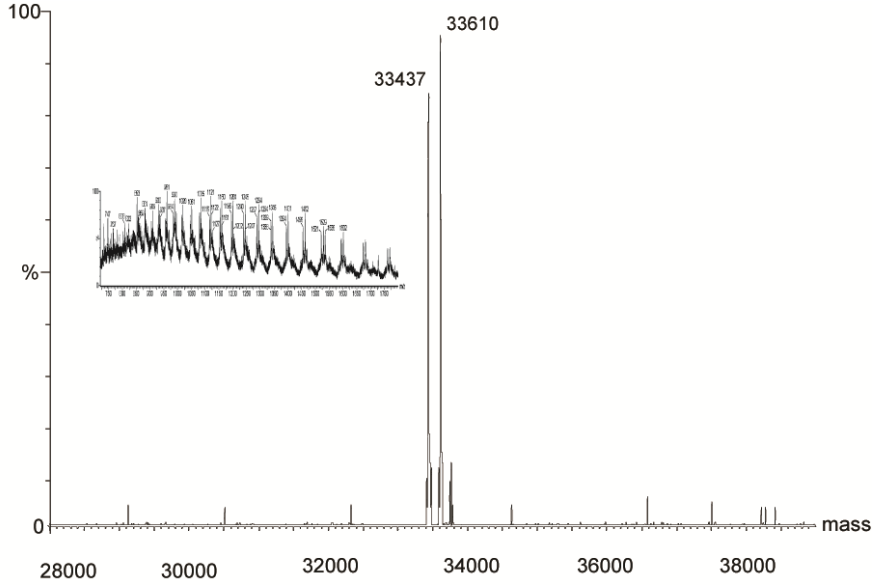


8.2.21.5 Reaction of Ag85-124PMS with Selenide 231.

Ag85C-124(2-hydroxydecyl-1-seleno) (232) and Ag85C-124(2-hydroxydecyl-1-seleno)-PMS (235,236)



$NaBH_4$ (0.18 mg, 4.79 μmol) was added portionwise to a solution of bis(2-hydroxydecyl)-1,1'-diselenide **231** (2.27 mg, 4.79 μmol) in MeOH (30 μL) at 22 ºC and the resulting mixture vortexed. Gas evolution occurred, and the solution turned colorless. After 1 minute, this mixture was added dropwise to a solution of Ag85C-124PMS **230** (200 μL of ca. 0.8 mg/mL, 4.79 nmol) in 20 mM TEA buffer (pH 7.2). After 12 h of shaking at 22 ºC, the reaction was analyzed directly by LC–MS and 50% conversion to Se ag85C **232** (calculated mass, 33436; observed mass, 33437) together with 50% conversion to Se-PMS ag85C **235,236** (calculated mass, 33606; observed mass, 33610) was observed. The protein sample was flash frozen with liquid nitrogen and stored at –80 ºC.



8.3 References

1. Thompson, J., Robrish, S.A., Pikis, A., Brust, A. & Lichtenthaler*, F.W. Phosphorylation and metabolism of sucrose and its five linkage-isomeric α-glucosyl-fructoses by *Klebsiella pneumoniae*. *Carbohydr. Res.* **331**, 149-161 (2001).
2. Veibel. *Biochemische Zeitschrift* **239**, 350,358, 372 (1931).
3. Wright, T., Brunt and Brownsey. Synthesis of 3-deoxy-3-fluoro-D-glucose 1- and 6-phosphates and their intreaction with phosphoglucomutase and UDPG-pyrophosphorylase. *Journal of the Chemical Society, Chemical Communication* **11**, 2 (1972).
4. Drueckhammer, D.G. & Wong, C.H. Chemoenzymic syntheses of fluoro sugar phosphates and analogs. *Journal of Organic Chemistry* **50**, 5912-5913 (1985).
5. Chenault, H.K., Mandes, R.F. & Hornberger, K.R. Synthetic Utility of Yeast Hexokinase. Substrate Specificity, Cofactor Regeneration, and Product Isolation. *Journal of Organic Chemistry* **62**, 331-336 (1997).
6. Gibson, R.P., Turkenburg, J.P., Charnock, S.J., Lloyd, R. & Davies, G.J. Insights into Trehalose Synthesis Provided by the Structure of the Retaining Glucosyltransferase OtsA. *Chemistry & Biology* **9**, 1337-1346 (2002).
7. Gauthier, C., Legault, J., Lebrun, M., Dufour, P. & Pichette, A. Glycosidation of lupane-type triterpenoids as potent in vitro cytotoxic agents. *Bioorganic & Medicinal Chemistry* **14**, 6713-6725 (2006).

8. Trujillo, M., Morales, E.Q. & Vazquez, J.T. Tetra-O-benzoylglucosylation: A New 1H Nuclear Magnetic Resonance Method for Determination of the Absolute Configuration of Secondary Alcohols. *Journal of Organic Chemistry* **59**, 6637-6642 (1994).
9. Willard. *Canadian Journal of Chemistry* **42**, 2560 (1964).
10. Ferrier, R.J. & Furneaux, R.H. Photobromination of carbohydrate derivatives. VI. Functionalization at C6 of 2,3,4-Tri-O-acetyl-1,6-anhydro-β-D-glucopyranose. *Aust. J. Chem.* **33**, 1025-1036 (1980).
11. Bassily, R.W., el-Sokkary, R.I., Silwanis, B.A., Nematalla, A.S. & Nashed, M.A. An improved synthesis of 4-azido-4-deoxy- and 4-amino-4-deoxy-α,α-trehalose and their epimers. *Carbohydr. Res.* **239**, 197-207 (1993).
12. Richardson, G.B.a.A.C. Chemical modification of trehalose. Part II. Synthesis of the galacto-analogue of trehalose. *J. Chem. Soc.*, 3 (1970).
13. Dondoni, A. & Scherrmann, M.-C. Thiazole-Based Synthesis of Formyl C-Glycosides. *J. Org. Chem.* **59**, 6404-6412 (1994).
14. Li, X., Ohtake, H., Takahashi, H. & Ikegami, S. A facile synthesis of 1'-C-alkyl-α-disaccharides from 1-C-alkyl-hexopyranoses and methyl 1-C-methyl-hexopyranosides. *Tetrahedron* **57**, 4297-4309 (2001).
15. Namme, R., Hideyo, T.M. & Ikegami, T.S. Development of Ketoside-Type Analogues of Trehalose by Using α-Stereoselective O-Glycosidation of Ketose. *Eur. J. Org. Chem.* **2007**, 3758-3764 (2007).
16. Yamanoi, T., Inoue, R., Matsuda, S., Katsuraya, K. & Hamasaki, K. Synthesis of trehalose mimics by bismuth(III) triflate or bis(trifluoromethane)sulfonimide-catalyzed 1-C-methyl-d-hexopyranosylation. *Tetrahedron: Asymmetry* **17**, 2914-2918 (2006).
17. Zhu, X., Jin, Y. & Wickham, J. Efficient Synthesis of Methylene exo-Glycals: Another Use of Glycosylthiomethyl Chlorides. Vol. 72 2670-2673 (2007).
18. RajanBabu, T.V. & Reddy, G.S. 1-Methylene sugars as C-glycoside precursors. *J. Org. Chem.* **51**, 5458-5461 (1986).
19. Li, X., Ohtake, H., Takahashi, H. & Ikegami, S. An efficient synthesis of new 1'-C-methyl-α-O-disaccharides using 1-methylenesugars as the glycosyl donors. *Tetrahedron* **57**, 4283-4295 (2001).
20. Kuzuhara, H. & Fletcher, H.G. Syntheses with partially benzylated sugars. VIII. Substitution at carbon-5 in aldose. The synthesis of 5-O-methyl-D-glucofuranose derivatives. *J. Org. Chem.* **32**, 2531-2534 (1967).
21. Glanzer, B.I. & Csuk, R. Reaction of pyranoid and furanoid aldonolactones with chloromethyltrimethylsilane-derived reagents. *Carbohydrate Research* **220**, 79-92 (1991).
22. Griffin, Frank K., Paterson, Duncan E., Murphy, Paul V. & Taylor, Richard J.K. A New Route to exo-Glycals Using the Ramberg-Bäcklund Rearrangement. Vol. 2002 1305-1322 (2002).
23. Hanessian, S. & Kagotani, M. Novel methods for the preparation of partially acetylated carbohydrates. *Carbohydrate Research* **202**, 67-79 (1990).
24. Tiwari, P. & Misra, A.K. An Efficient Stereoselective Dihydroxylation of Glycals using a Bimetallic System, RuCl3/CeCl3/NaIO4. Vol. 71 2911-2913 (2006).
25. Nicolaou, K.C., Snyder, S.A., Nalbandian, A.Z. & Longbottom, D.A. A New Method for the Stereoselective Synthesis of α and β Glycosylamines Using the Burgess Reagent. Vol. 126 6234-6235 (2004).
26. Plé, K., Chwalek, M. & Voutquenne-Nazabadioko, L. Synthesis of l-arabinopyranose containing hederagenin saponins. *Tetrahedron* **61**, 4347-4362 (2005).
27. Seeberger, P.H., Eckhardt, M., Gutteridge, C.E. & Danishefsky, S.J. Coupling of Glycal Derived Thioethyl Glycosyl Donors with Glycal Acceptors. An Advance in the Scope of the Glycal Assembly. *Journal of the American Chemical Society* **119**, 10064-10072 (1997).
28. Zhang, Y.-M., Mallet, J.-M. & Sinaÿ, P. Glycosylation using a one-electron-transfer, homogeneous reagent. Application to an efficient synthesis of the trimannosyl core of N-glycosylproteins. *Carbohydrate Research* **236**, 73-88 (1992).
29. Jain, R.K., Liu, X.-G., Oruganti, S.R., Chandrasekaran, E.V. & Matta, K.L. Synthesis of oligosaccharide substrates for N-linked glycoprotein processing enzymes. *Carbohydrate Research* **271**, 185-196 (1995).
30. Ogawa, T. & Yamamoto, H. Synthesis of linear -mannotetraose and -mannohexaose, partial structures of the cell-surface -mannan of *Candida albicans* and *Candida utilis*. *Carbohydr. Res.* **104**, 271-283 (1982).
31. Patel, M. Part II thesis *Oxford University* (2007).
32. Khan, R., Konowicz, P.A., Gardossi, L., Matulova, M. & Degennaro, S. Regioselective Deacetylation of Fully Acetylated Mono- and Di-Saccharides With Hydrazine Hydrate. *Australian Journal of Chemistry* **49**, 293-298 (1996).
33. Grynkiewicz, G.K., Miroslawa; . *Pol. J. Chem* **61**, 147-153 (1987).

34. Khan, K.M., Perveen, S., Shah, S.T.A., Shekhani, M.S. & Voelter, W. Sodium hydride/hexamethylphosphoric triamide: a new and efficient reagent towards the synthesis of protected 1,2- and 5,6-enopyranosides. *New Journal of Chemistry* **25**, 896-898 (2001).
35. Szarek, W.A., Hay, G.W., Doboszewski, B. & Perlmutter, M.M. Reaction of 1,2-anhydro-3,4:5,6-di-O-isopropylidene-1-C-nitro--mannitol with potassium hydrogenfluoride in ethylene glycol: a synthesis of 2-deoxy-2-fluoro--glucose. *Carbohydrate Research* **155**, 107-118 (1986).
36. Albert, M., Dax, K. & Ortner, J. A novel direct route to 2-deoxy-2-fluoro-aldoses and their corresponding derivatives. *Tetrahedron* **54**, 4839-4848 (1998).
37. Shelling, J.G., Dolphin, D., Wirz, P., Cobbledick, R.E. & Einstein, F.W.B. 2'-fluoromaltose: Synthesis and properties of 4-O-(2-deoxy-2-fluoro-α-glucopyranosyl)-glucopyranose, and the crystal structure of 2,3-di-O-acetyl-1,6-anhydro-4-O-(3,4-tri-O-acetyl-2-deoxy-2-fluoro-α-glucopyranosyl)-[beta]--glucopyranose. *Carbohydr. Res.* **132**, 241-259 (1984).
38. Withers, S.G., MacLennan, D.J. & Street, I.P. The synthesis and hydrolysis of a series of deoxyfluoro-glucopyranosyl phosphates. *Carbohydr. Res.* **154**, 127-144 (1986).
39. Foster, A.B.H., R.; Hall, L.D. Fluorinated carbohydrates. Part V. Nuclear magnetic resonance data on pyranose and furanose derivatives of 3-deoxy-3-fluoro-D-glucose. *Can. J. Chem.* **24**, 8 (1970).
40. Chittenden, G.J.F. A simplified synthesis of α-galactopyranose 1,3,4,6-tetraacetate. *Carbohydr. Res.* **183**, 140-143 (1988).
41. Tang, C.-J., Shen, X. & Wu, Y. Synthesis of An Advanced Precursor to the Amino Epimer of (+)-Pyloricidins. *Monatshefte für Chemie / Chemical Monthly* **134**, 1617-1622 (2003).
42. Umezawa, S.e.a. *Bull. Chem. Soc. Jpn* **48**, 556-559 (1975).
43. Myszka, H., Bednarczyk, D., Najder, M. & Kaca, W. Synthesis and induction of apoptosis in B cell chronic leukemia by diosgenyl 2-amino-2-deoxy-β-glucopyranoside hydrochloride and its derivatives. *Carbohydr. Res.* **338**, 133-141 (2003).
44. Boullanger, P., Jouineau, M., Bouammali, B., Lafont, D. & Descotes, G. The use of N-alkoxycarbonyl derivatives of 2-amino-2-deoxy--glucose as donors in glycosylation reactions. *Carbohydrate Research* **202**, 151-164 (1990).
45. Saito, H., Yoshikawa, H. & Nishimura, Y. Studies on lignan lactone antitumor agents. I. Synthesis of aminoglycosidic lignan variants related to podophyllotoxin. *Chem. Pharm. Bull. (Tokyo)* **34**, 3733-3740 (1986).
46. Janiak, A.M., Hoffmann, M., Milewska, M.J. & Milewski, S. Hydrophobic derivatives of 2-amino-2-deoxy--glucitol-6-phosphate: A new type of -Glucosamine-6-phosphate synthase inhibitors with antifungal action. *Bioorg. Med. Chem.* **11**, 1653-1662 (2003).
47. Onodera, K. & Komano, T. Debenzyloxycarbonylation of 1,3,4,6-Tetra-O-acetyl-2-benzyloxycarbonylamino- 2-deoxy-D-hexopyranoses in the Conversion of Acetoxy to Glycosyl Bromide. *The Journal of Organic Chemistry* **26**, 3932-3933 (1961).
48. Leatherdale, C.A., Woo, W.K., Mikulec, F.V. & Bawendi, M.G. On the Absorption Cross Section of CdSe Nanocrystal Quantum Dots. *The Journal of Physical Chemistry B* **106**, 7619-7622 (2002).
49. Wang, M., Tu, P.-F., Xu, Z.-D., Yu, X.-L. & Yang, M. Design and Synthesis of Guanidinoglycosides Directed against the TAR RNA of HIV-1. *Helvetica Chimica Acta* **86**, 2637-2644 (2003).
50. Baer, H.H., Breton, R.L. & Shen, Y. Synthesis of a trehalose homolog, 6-deoxy-α-gluco-heptopyranosyl 6-deoxy-α-gluco-heptopyranoside, and the corresponding bis(heptosiduronic acid). *Carbohydr. Res.* **200**, 377-389 (1990).
51. Sanki, A.K. et al. Synthesis of methyl 5-S-alkyl-5-thio-D-arabinofuranosides and evaluation of their antimycobacterial activity. *Bioorganic & Medicinal Chemistry* **16**, 5672-5682 (2008).
52. Wright, K. et al. Kanamycin - New Tuberculostatic Agent. *American Review of Tuberculosis and Pulmonary Diseases* **78**, 286-286 (1958).
53. Augestad, I. & Berner, E. Chromatographic Separation of Anomeric Glycosides .2. New Crystalline Methylfuranosides of Galactose, Arabinose and Xylose. *Acta Chemica Scandinavica* **8**, 251-256 (1954).

9 Appendix 2: Biological materials and methods

9.1 General Experimental

Green Amber Red Screen and kinetics were conducted on a Waters Quattro Micro-MS with electrospray ionization operating in negative mode (ESI-) interfaced with a Waters 1525µ HPLC system and Waters 2777 sampler fitted with a 4-valve injector module. MS analysis was under the control of Micromass Masslynx 4.1 software and data processed using Masslynx4.1, QuantLynx, Microsoft Excel 2003, SigmaPlot 11.0 and Origin 7.5.

Protein concentrations were calculated using standard BCA assay or with a Labtek ND-1000 Nanodrop. Protein purification was performed on an AKTA Prime FPLC system (GE Healthcare).

High Performance Liquid Chromatography was conducted on a Dionex UltiMate 3000 HPLC system at ambient temperature, with an in-line variable UV absorbance detector, or a Varian PLS400 Evaporative Light Scattering detector (ELSD) parallel to the main flowpath.

Fluorescence and radioactivity of lipid extractions was read on a Phosphorimager: Typhoon 9410 Variable Mode Imager by GE Healthcare Bio-Sciences.

Scintillation counting was conducted on an LS 6500 Multi-purpose scintillation counter by Beckman Coulter.

Fluorescence readings of *Mtb* were conducted on a FLUOstar Optima by BMG Labtech.

Images of stained cells were obtained by confocal microscopy (Leica SP5, Leica Microsystems, Exton, PA) using a 63× oil immersion objective NA 1.4. All data were processed by Leica LAS AF Lite and Adobe Illustrator Software. DAPI was first excited with 405 diode laser and emission spectra were recorded at 413-470 nm prior to other fluorophores. FITC-Tre was excited at 488 nm and emission spectra were recorded at 497-557 nm. RFP was excited at 594 nm and emission spectra recorded at 602-651 nm. Alexafluor 633 was excited at 633 nm and emission spectra recorded at 651-759 nm**.**

Polymerase chain reactions (PCR) were conducted on an icycler system (BioRad, Hercules, CA).

Cryosections were produced on a Reichert Leica CM3000 Cryostat with 10-20 µm slice thickness.

Sequence Alignment and virtual cloning was performed with DNAStar Lasergene software suite. MegAlign software was used for sequence alignment and SeqBuilder software was used for plasmid design.

Cells were electroporated in a Gene Pulser Xcell Electroporation system (BioRad, Hercules, CA).

DNA and Protein Gels were imaged on an Alpha Imager HP System (ProteinSimple, Santa Clara, CA).

PET scans were conducted on a MicroPET Focus 220, Siemens Preclinical Solutions, Knoxville, TN). CT Scans were conducted on a BioVet CT System, (m2m imaging, Cleveland, OH)

Circular Dichroism was performed on a Jasco J-715 spectropolarimeter (Jasco Products Company).

Data was analyzed and graphed in Graphpad Prism 5, OrigenPro 8.0 or Microsoft Excel where indicated.

9.2 Protein Expression

9.2.1 Ag85 general protein expression methods

Plasmids were obtained from Colorado State tuberculosis research contract HHSN266200400091c as Rv3804/Ag85A Rv1886c/Ag85B Rv0129c/Ag85C in pET23b vector containing His-tags and were expressed as has been described previously[1]. LB $Amp^{100}Cam^{38}$ (30 mL) was inoculated with desired plasmid in BL21 DE3 pLys strain and grown overnight at 37 ˚C with shaking to generate a starter culture, which was used to inoculate 4L of autoclaved $Amp^{100}Cam^{38}$ LB broth. Culture was grown at 37 ˚C with shaking until it reaches an OD_{600} of 0.5 and removed from 37 ˚C and place in 25 ˚C incubator with shaking and allowed to cool to 25 ˚C. Expression was induced with isopropyl-β-D-thiogalactopyranoside (IPTG) (0.25mM) for 16 hours. Cells were harvested and resuspended in buffer (10mM Tris-HCl, pH 8) containing one tablets of Complete® protease inhibitor cocktail (Roche), lysozyme and DNAse I. Pellet was lysed by sonication and pelleted at 8,000 rpms for 60 min. Purification was carried out through fast protein liquid chromatography (FPLC) via nickel affinity chromatography. Two 5 mL Ni NTA columns were connected and equilibrated with binding buffer (10mM Tris-HCl, 500mM NaCl, 5mM imidazole, pH 8). Lysate was applied to column slowly via 50 mL super loop and protein purification was conducted via FPLC using a stepwise gradient of elution buffer (10mM Tris-HCl, 1M Imidazole, pH 8). Eluant was then dialyzed overnight into 20 mM TEA buffer pH 7.4. SDS-page was performed on all protein fractions and gels were stained with InstantBlue stain (Gentaur). PD-10 desalting cartridges (Amersham) were used to remove endotoxins and to buffer exchange into 1mM TEA pH 7.4 for mass spectrometry assay. Protein concentrations were calculated by standard BCA page assay and by absorbance on a nanodrop. Mass spectrometry and sequencing analysis confirmed the correct identity of the proteins.

9.3 Expression of Ag85A

Ag85A was expressed as above and, upon dialysis, was obtained at a concentration of 260 μg/mL. During expression, it was found that both a protein of the expected mass as well as a protein of mass + 178 Da were obtained in the purified protein **(Figure 9.1- Figure 9.4)**. Analysis indicates that this mass is caused by His-tag phosphonogluconoylation[2] of a GSS–His sequence present in the protein His-tag. Furthermore, when the His-tag mixture of Ag85A was cleaved with thrombin, a single protein of the correct mass was obtained, confirming that the modification was in the His-tag region. In separate expressions a mixture of the gluconoylated and non-gluconoylated proteins and purely gluconoylated protein was obtained. No significant difference in activity for the gluconoylated and non-gluconoylated protein was observed.

Sequence of Rv3804/Ag85A (PDB code **1SFR**)

(Met)G S S H H H H H H S S G L V P R G S H M F S R P G L P V E Y L Q V P S P S M G R D I K
V Q F Q S G G A N S P A L Y L L D G L R A Q D D F S G W D I N T P A F E W Y D Q S G L S
V V M P V G G Q S S F Y S D W Y Q P A C G K A G C Q T Y K W E T F L T S E L P G W L Q A
N R H V K P T G S A V V G L S M A A S S A L T L A I Y H P Q Q F V Y A G A M S G L L D P S
Q A M G P T L I G L A M G D A G G Y K A S D M W G P K E D P A W Q R N D P L L N V G K
L I A N N T R V W V Y C G N G K P S D L G G N N L P A K F L E G F V R T S N I K F Q D A Y
N A G G G H N G V F D F P D S G T H S W E Y W G A Q L N A M K P D L Q R A L G A T P N
T G P A P Q G A

Calculated Masses:

Calculated average isotopic mass with glucose: 33995

Found average isotopic mass: 33995 or 33989 depending on calibration

Calculated average isotopic mass without N terminal methionine and gluconoylation: 33814

Found average isotopic mass: 33817

Calculated average isotopic mass after thrombin cleavage: 32063

Found average isotopic mass: 32061

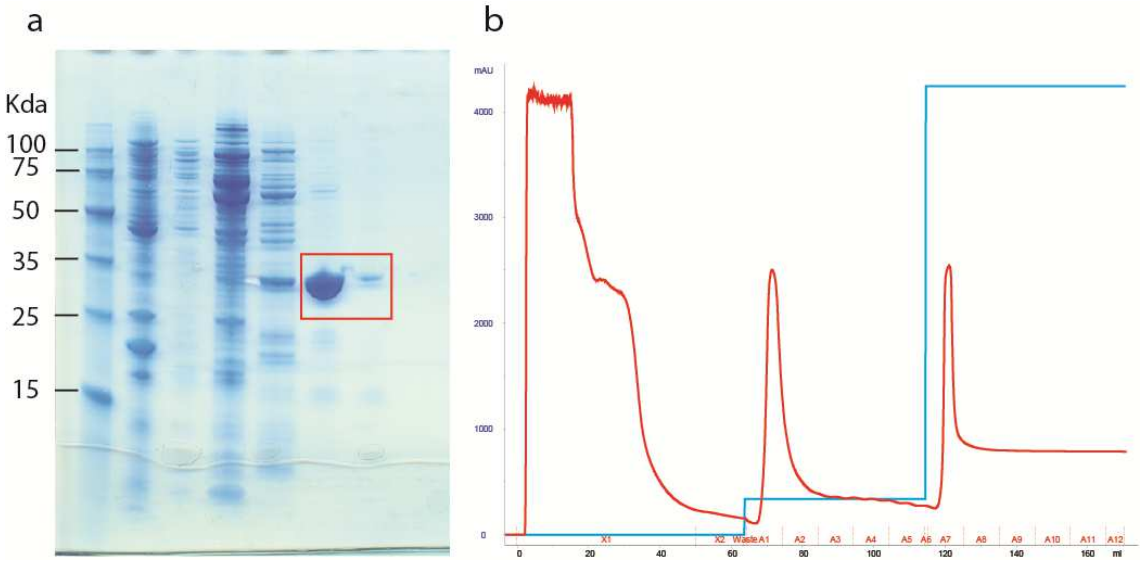


Figure 9.1. (a) SDS page gel of Ag85A, lane 1- novagen perfect protein makers 15-150 Kda, box indicates fractions collected (b) FPLC trace of protein expression red- protein absorbance, blue- gradient of imidazole (5 mM to 1 M).

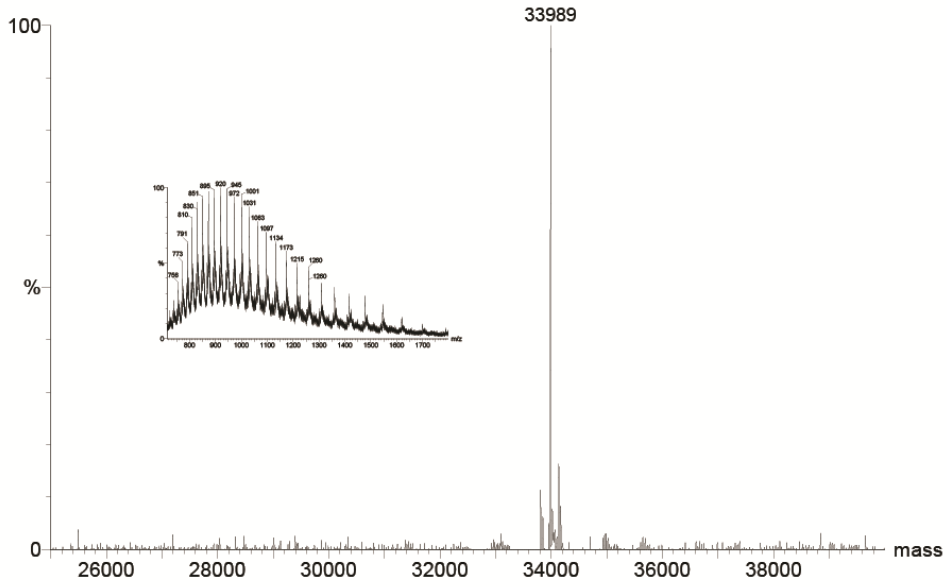


Figure 9.2. Mass spectrum of Ag85A with gluconoylation. Inset shows protein spectrum before deconvolution.

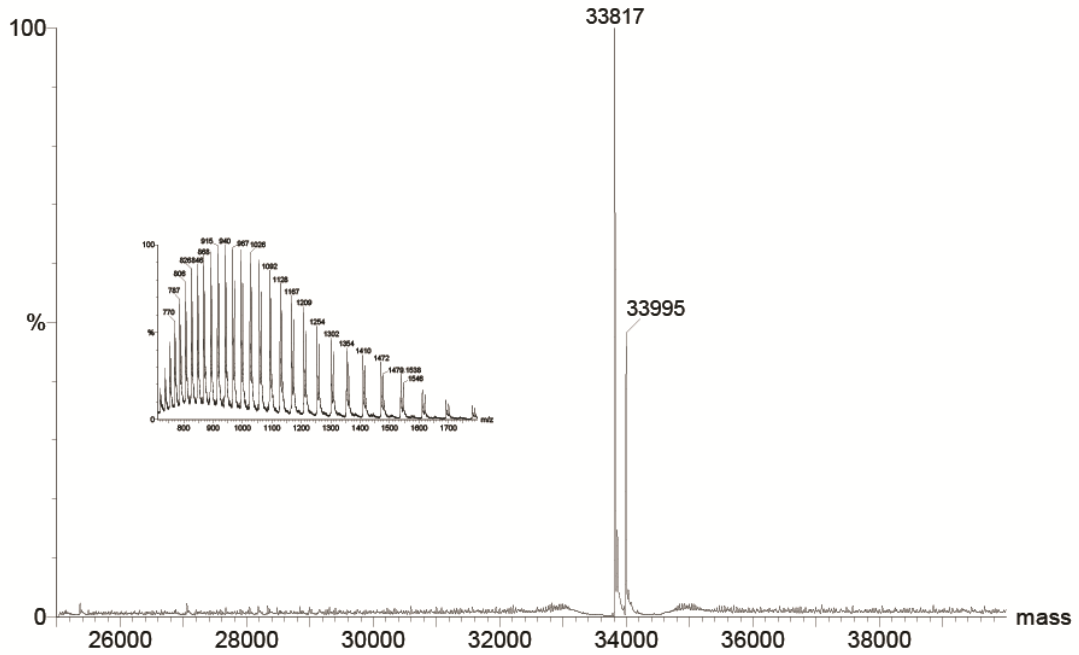


Figure 9.3. Mass spectrum of Ag85A as a mixture with and without gluconoylation. Inset shows protein spectrum before deconvolution.

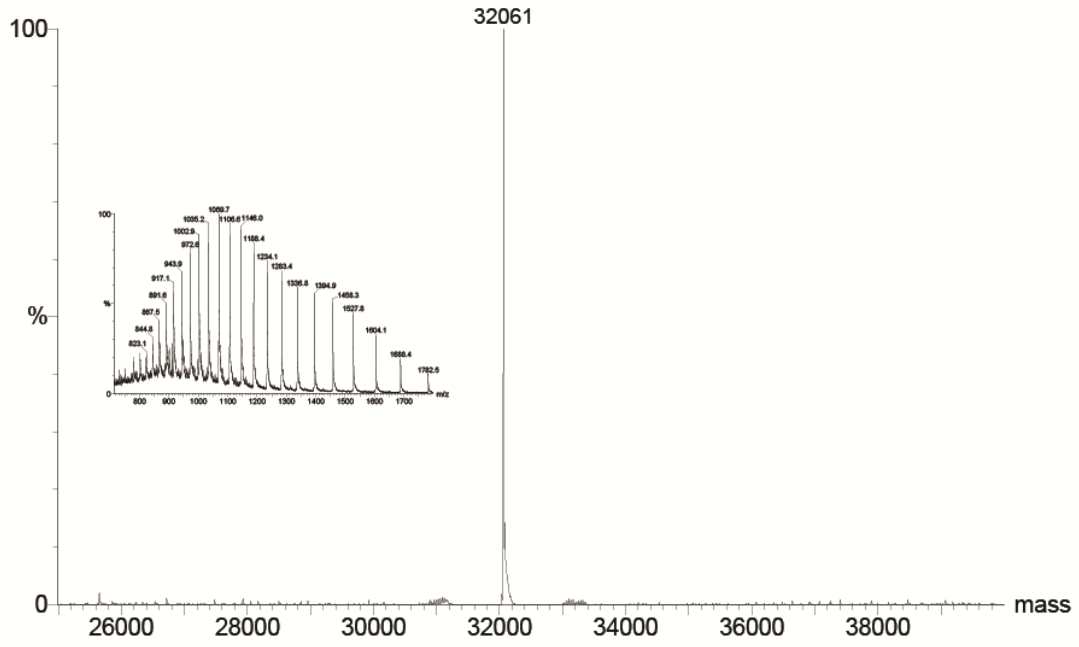


Figure 9.4. Mass spectrum of Ag85A treated with thrombin. Inset shows protein spectrum before deconvolution.

9.3.1 Expression of Ag85B

Ag85B was expressed as above and, upon dialysis, was obtained at a concentration of 350 µg/mL **(Figure 9.5)**

Sequence of Rv1886c/Ag85B (PDB codes 1F0N and 1F0P)

M F S R P G L P V E Y L Q V P S P S MG R D I K V Q F Q S G G N N S P A V Y L L D G L R A Q D D Y N G W D I N T P A F E W Y Y Q S G L S I V M P V G G Q S S F Y S D W Y S P A C G K A G C Q T Y K W E T F L T S E L P Q W L S A N R A V K P T G S A A I G L S M A G S S A M I L A A Y H P Q Q F I Y A G S L S A L L D P S Q G M G P S L I G L A M G D A G G Y K A A D M W G P S S D P A W E R N D P T Q Q I P K L V A N N T R L W V Y C G N G T P N E L G G A N I P A E F L E N F V R S S N L K F Q D A Y N A A G G H N A V F N F P P N G T H S W E Y W G A Q L N A M K G D L Q S S L G A G K L A A A L E H H H H H H

Calculated Masses:

Calculated average isotopic mass: 32313

Found average isotopic mass: 32310

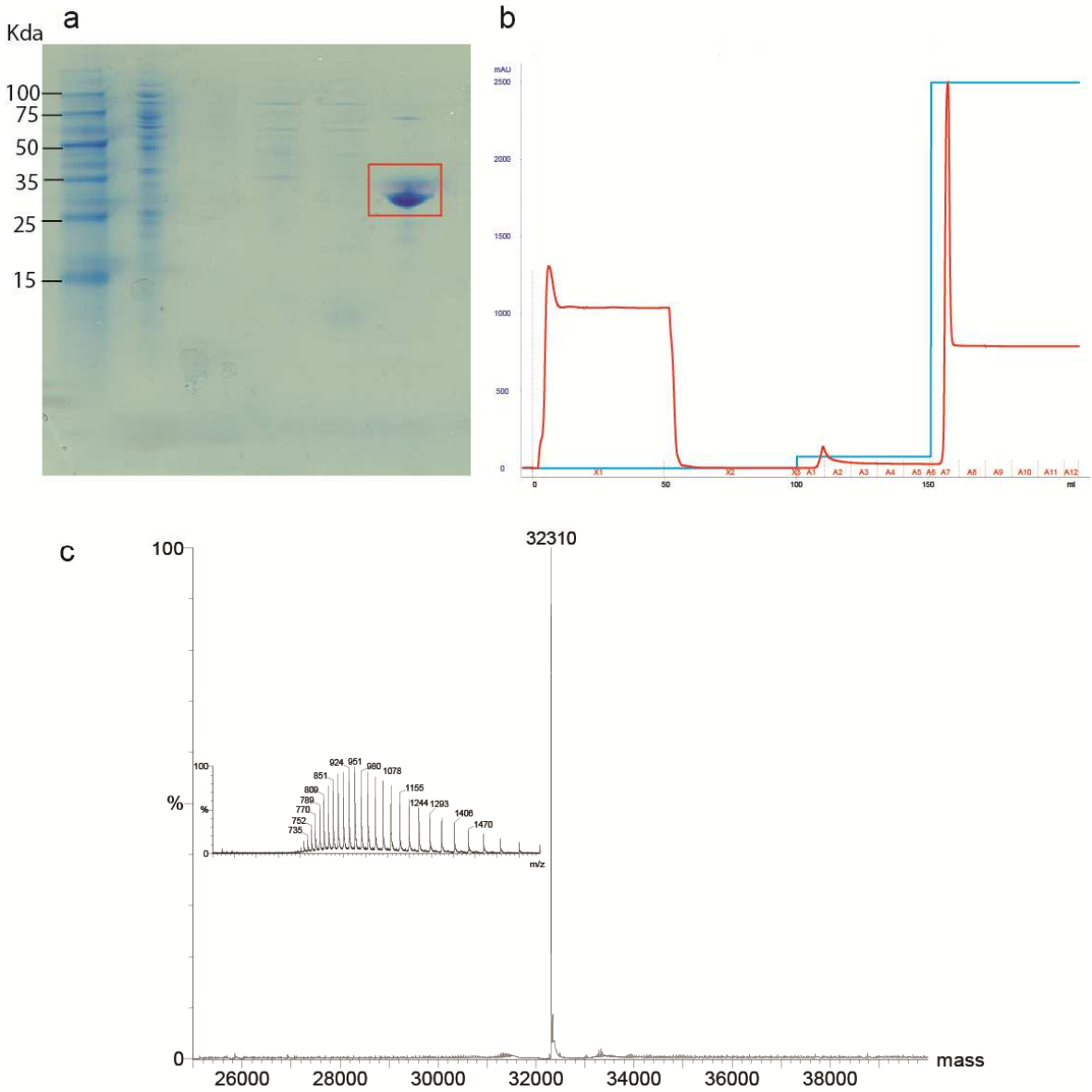


Figure 9.5. (a) SDS page gel of Ag85B, lane 1- novagen perfect protein makers 15-150 Kda, box indicates fractions collected (b) FPLC trace of protein expression red- protein absorbance, blue- gradient of imidazole. (5 mM to 1 M) (c) Mass spectrum of Ag85 B. Inset shows protein spectrum before deconvolution.

9.3.2 Expression of Ag85C

Ag85C was expressed as above and, upon dialysis, was obtained at a concentration of 1.6 mg/mL **(Figure 9.6).**

Sequence of Rv0129c/Ag85C (PDB codes 1DQY, 1DQZ)

M F S R P G L P V E Y L Q V P S A S M G R D I K V Q F Q G G G G P H A V Y L L D G L R A Q
D D Y N G W D I N T P A F E E Y Y Q S G L S V I M P V G G Q S S F Y T D W Y Q P S Q S N G
Q N Y T Y K W E T F L T R E M P A W L Q A N K G V S P T G N A A V G L S M S G G S A L I
L A A Y Y P Q Q F P Y A A S L S G F L N P S E S W W P T L I G L A M N D S G G Y N A N S
M W G P S S D P A W K R N D P M V Q I P R L V A N N T R I W V Y C G N G T P S D L G G D
N I P A K F L E G L T L R T N Q T F R D T Y A A D G G R N G V F N F P P N G T H S W P Y W
N E Q L V A M K A D I Q H V L N G A T P P A A P A A P A A L E H H H H H H

Calculated Masses:

Calculated average isotopic mass: 33217

Found average isotopic mass: 33218

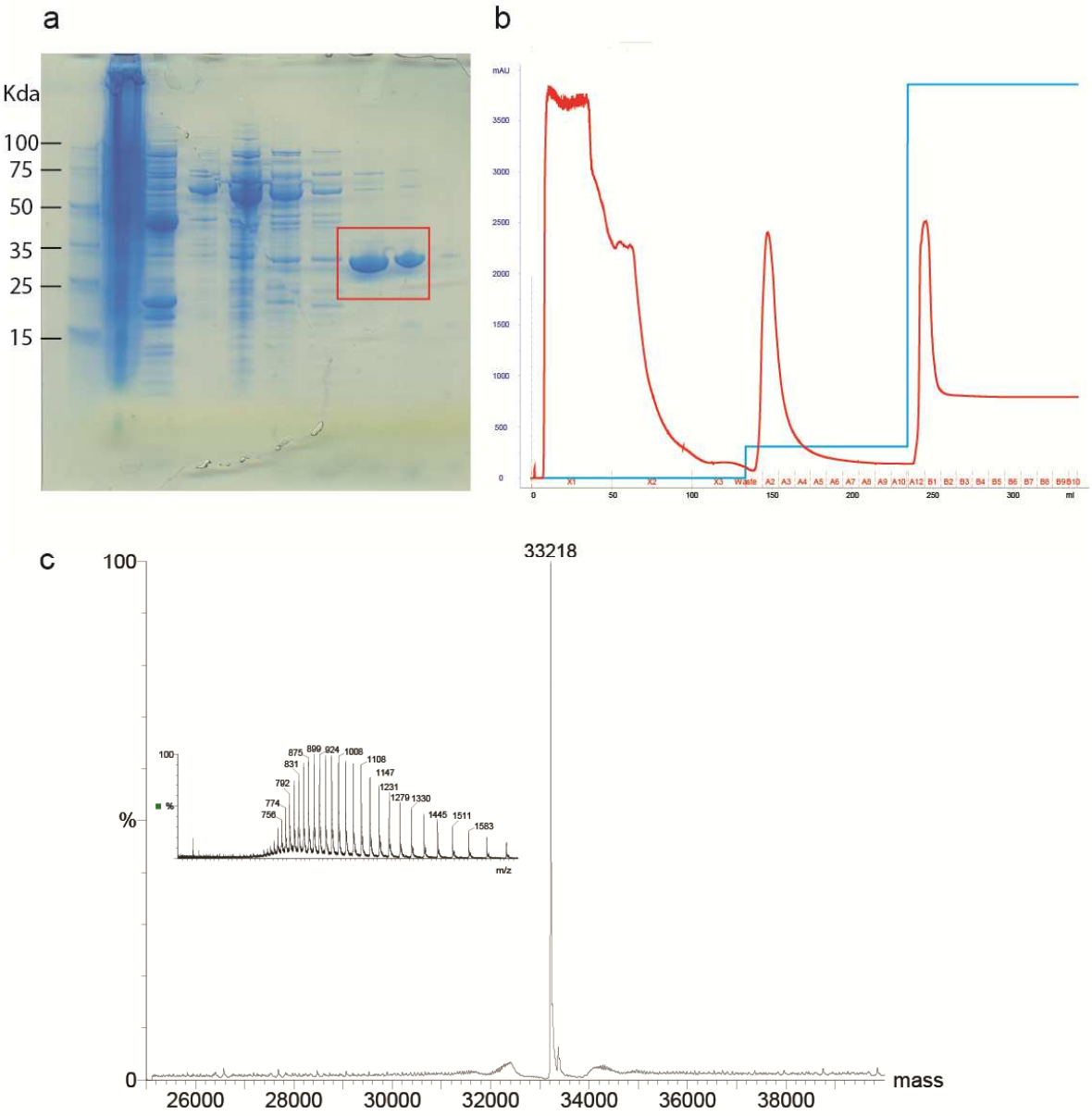


Figure 9.6. (a) SDS page gel of Ag85 C = lane 1- novagen perfect protein makers 15-150 Kda, box indicates fractions collected (b) FPLC trace of protein expression red- protein absorbance, blue- gradient of imidazole. (5 mM to 1 M) (c) Mass spectrum of Ag85C. inset shows protein spectrum before deconvolution.

9.3.3 Expression of Trehalose-6-phosphate synthase

The plasmids coding for trehalose-6-phosphate synthase (OtsA), the pET22b:otsA (kindly gifted by Prof. Gideon Davies) was used to transform *E.coli* BL21 DE3 cells. Transformed cells were grown overnight in 50 ml of LB broth containing 100 µg ampicillin. One-liter cultures were then inoculated to an A_{600} of ~0.05. The cells were grown at 37°C to an A_{600} of 0.45. Protein expression was subsequently induced by the addition of 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and left to grow for a further 16 h at 30°C. Expression of the protein was confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE). All protein purification steps were carried out at 4°C. The cell pellets were resuspended in 50 mL 20 mM tris(hydroxymethyl)aminomethane (Tris), 200 mM NaCl, 15 mM imidazole, pH 7.8, containing one tablets of Complete® protease inhibitor cocktail (Roche), lysozyme and DNAse I. The cells were disrupted on ice by sonication. The suspension was then centrifuged (30,000 g for 40 min) to remove cell debris. The supernatant was then filtered using a 0.4 µm syringe filter. The filtered supernatant was applied to a preequilibrated (Buffer A; 20 mM Tris, 200 mM NaCl, 10 mM imidazole, pH 7.8) Pharmacia 5ml HisTrap Ni-NTA column. The column was washed at 1 ml/min with 20 column volumes of the same buffer and then eluted with a linear gradient of imidazole (10 mM to 500 mM over 25 column volumes) in Buffer A. Protein was detected with an on-line detector monitoring A_{280} and column fractions were collected and analyzed by SDS-PAGE. Fractions containing the ca. 54 KDa protein was pooled and dialyzed twice against 4 liters of 100 mM Hepes, 200 mM KCl, 20 mM $MgCl_2$, pH 7.5 and stored at 4°C.

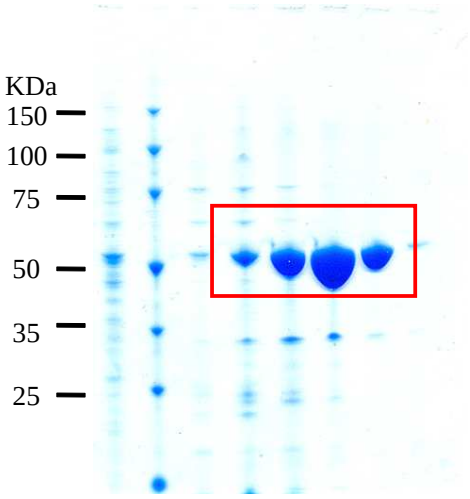
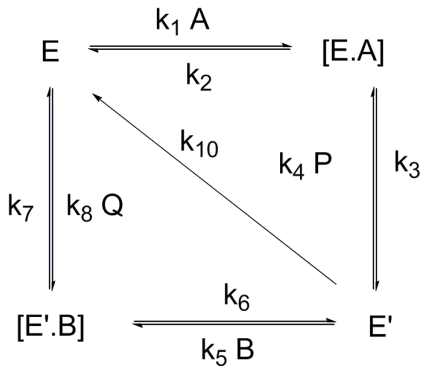


Figure 9.7. (a) SDS page gel of OtsA, lane 1- novagen perfect protein makers 15-150 Kda, box indicates fractions collected

9.4 Kinetic Analysis[i]

Derivation of Rate Equations

A Bi-Bi Ping-Pong reaction mechanism, generalized to include hydrolysis of the acyl enzyme intermediate, is represented below:



Performing a King-Altman analysis[1] of the mechamistic scheme and assuming initial rate conditions of [P]=0, [Q]=0, the four enzyme forms can be represented by the following distribution equations:

$$[E]/[E_t] = (k_2k_5k_7B + k_3k_5k_7B + k_2k_7k_{10} + k_3k_6k_{10} + k_3k_7k_{10} + k_2k_6k_{10})/\Delta \quad \text{(Eq. 1)}$$

$$[E.A]/[E_t] = (k_1k_5k_7AB + k_1k_7k_{10}A + k_1k_6k_{10}A)/\Delta \quad \text{(Eq. 2)}$$

$$[E']/[E_t] = (k_1k_3k_6A + k_1k_3k_7A)/\Delta \quad \text{(Eq. 3)}$$

$$[E'.B]/[E_t] = (k_1k_3k_5AB)/\Delta \quad \text{(Eq. 4)}$$

In all equations 1-4 the denominator term (Δ) is the sum of all the numerator terms in the four equations i.e.:

$$\Delta = k_{10}(k_{10} + k_3)(k_6 + k_{10}) + k_1(k_6 + k_7)(k_3 + k_{10})A + k_5k_7(k_2 + k_3)B + k_1k_5(k_3 + k_7)AB \quad \text{(Eq. 5)}$$

[i] All kinetic analysis was conducted by Dr. Conor Barry and for full details of all Ag85 ESI MS assays please see 3. Barry, C.S., Backus, K.M., Barry, C.E. & Davis, B.G. ESI-MS Assay of *M. tuberculosis* Cell Wall Antigen 85 Enzymes Permits Substrate Profiling and Design of a Mechanism-Based Inhibitor. *J. Am. Chem. Soc.*, ASAP (2011). Include here for completeness are experimental details from the enzyme kinetics conducted.

TMH is produced in two steps of the catalytic cycle and so the overall initial rate of production of TMH (v_{TMH}) is given by:

$$v_{TMH} = k_3[\text{E.A}] + k_7[\text{E'.B}] \quad \text{(Eq.6)}$$

Substituting to express [E.A], [E'.B] and Δ in terms of rate constants gives the following equation:

$$v_{TMH} = \frac{\text{E}_t[2(k_1k_3k_5k_7)\text{AB} + k_1k_3k_{10}(k_6 + k_7)\text{A}]}{k_{10}(k_2 + k_3)(k_6 + k_7) + k_1(k_3 + k_{10})(k_6 + k_7)\text{A} + k_5k_7(k_2 + k_3)\text{B} + k_1k_5(k_3 + k_7)\text{AB}} \quad \text{(Eq.7)}$$

The kinetic constants were defined by Clelands nomenclature[2] as follows:

$$V_t = \frac{numAB}{coeffAB} = \frac{k_3k_7}{k_3 + k_7} \quad \text{(Eq.8)}$$

$$V_h = \frac{numA}{coeffA} = \frac{k_3k_{10}}{k_3 + k_{10}} \quad \text{(Eq.9)}$$

$$K_A = \frac{coeffB}{coeffAB} = \frac{k_7(k_2 + k_3)}{k_1(k_3 + k_7)} \quad \text{(Eq.10)}$$

$$K_B = \frac{coeffA}{coeffAB} = \frac{(k_6 + k_7)(k_3 + k_{10})}{k_5(k_3 + k_7)} \quad \text{(Eq.11)}$$

$$K_{iA} = \frac{const}{coeffA} = \frac{k_{10}(k_2 + k_3)}{k_1(k_3 + k_{10})} \quad \text{(Eq.12)}$$

Writing Equn.7 in terms of these definitions gives:

$$v_{TMH} = \frac{2V_t\text{AB} + V_hK_B\text{A}}{K_{iA}K_B + K_B\text{A} + K_A\text{B} + \text{AB}} \quad \text{(Eq.1)}$$

The presence of a constant term ($K_{iA}K_B$) in the denominator contrasts with a normal Bi-Bi Ping-Pong mechanism in which no such term exists. This offers a useful diagnostic handle to differentiate

between a Ping-Pong mechanism with a stable intermediate and a mechanism with an unstable intermediate. A double reciprocal plot of $1/v_o$ against 1/[A] will show a series of parallel lines for a normal ping-pong but the lines will intersect to the left of the y-axis in the case of an unstable intermediate.

King-Altman analysis as above decrees that the rates of hexanoate formation and acyl-transfer are given by Equn.14 and Equn.15 respectively:

$$v_{hydrolysis} = \frac{V_h K_B \mathrm{A}}{K_{iA} K_B + K_B \mathrm{A} + K_A \mathrm{B} + \mathrm{AB}} \qquad \text{(Eq.14)}$$

$$v_{transfer} = \frac{V_t \mathrm{AB}}{K_{iA} K_B + K_B \mathrm{A} + K_A \mathrm{B} + \mathrm{AB}} \qquad \text{(Eq.15)}$$

The rate of acyl-transfer may be dissected from the overall rate of TMH formation measuring the initial rates of both hydrolysis and TMH formation and applying the relationship described by Equn.16:

$$v_{transfer} = (v_{TMH} - v_{hydrolysis})/2 \qquad \text{(Eq.16)}$$

9.4.1 Ag85 ESI MS assays

Sample delivery: The HPLC/auto-sampler control was divided into two stages: **a)** injection of the internal standard into valve 1; and **b)** injection of the analyte solution into valve 2. Both valves were switched simultaneously under the control of MassLynx 4.1 software and the analyte and standard mixed in a 186uL peek loop before elution directly onto the source. Conditions: Mobile phase: ACN : H_2O (50 : 50); Flow rate: 0.2 mL/min; isocratic method for 3 min; Injection volume: 10.0 µL, ESI- (single ion monitoring) for 3 min.

Standard curve determination: Eight samples were analysed with one containing 20 uM GlcNAc (in TEA buffer) only, and the other seven containing fixed concentrations of trehalose (5, 10, 20, 40, 80, 120 and 160uM), TDH (5, 10, 20, 40, 80, 120 and 160uM), TMH (0.5, 1, 2, 4, 8, 12 and 16uM) and hexanoic acid (0.5, 1, 2, 4, 8, 12 and 16uM) in 1mM TEA, pH 7.2. The samples were injected as described above. The peak area was measured for each compound after normalisation for ionization efficiency using the standard (GlcNAc). Plots were constructed using QuantLynx of response against concentration and the slope was used to obtain concentration information during subsequent reactions.

Ag85C kinetic parameter determination: six vials were prepared containing a fixed concentration of TDH and varying concentrations of trehalose at 37^{o}C . Reactions were initiated by A85c (20uL)

addition to give final concentrations of 10nM Ag85C, 25uM TDH and 5, 10, 25, 50, 100 and 150uM trehalose. Samples of reaction mixtures and standard (20uM GlcNAc) were delivered to the mass spectrometer as described above and concentrations of hexanoate, TMH, trehalose and TDH were monitored simultaneously. The reactions were monitored every 26.7 min over a 6 hour period. The total ion count (TIC) for each analyte and the standard were recorded for each injection. Reference to the standard curve for each compound allowed the concentration to be determined. Initial rates of TMH formation were calculated from concentration *vs* time plots using Microsoft Excel 2003 and Origin 7.5 software. The entire procedure was repeated for other fixed concentrations of TDH (75, 125, 200, 400 and 600 uM). The resulting matrix of initial rates was fit to equation 13 using SigmaPlot 11.0 to determine the kinetic parameters of Ag85C activity.

Pseudo-single-substrate kinetics: Six vials were prepared containing a fixed concentration of TDH and varying concentrations of trehalose at 37^{o}C. Reactions were initiated by Ag85 (20 uL) addition to give final concentrations of 50nM Ag85C (500 nM Ag85A/B), 500uM TDH and 10, 15, 25, 50, 75, 100, 175 and 250 uM trehalose. Samples of reaction mixtures and standard (20 uM GlcNAc) were delivered to the mass spectrometer as described above and concentrations of hexanoate, TMH, trehalose and TDH were monitored simultaneously. The reactions were monitored every 26.7 min over an 8 hour period. The total ion count (TIC) for each analyte and the standard were recorded for each injection. Reference to the standard curve for each compound allowed the concentration to be determined. Initial rates were calculated from concentration *vs* time plots using Microsoft Excel 2003 and Origin 7.5 software. Initial rates of acyltransfer were determined from initial rates of TMH and hexanoate production according to Eq.16 (*vide supra*). The acyltransfer initial rate data were fit to the Michaelis-Menten equation using Origin 7.5 to determine K_M(app) and k_{cat}(app) values.

GAR screen: A 96-well plate was set up with each well containing TDH and one of the screen compounds in 1 mM TEA buffer (pH = 7.2) at 37^{o}C. To each well was added by automated injection (with mixing), 20 uL of either Ag85A/B/C or buffer to give final concentrations of 500uM of each substrate and 2 uM of Ag85 (or 0 uM Ag85 in the case of negative control wells). The samples were incubated at 37^{o}C for 2 h 40 min before injection of a 10 uL aliquot directly into the ESI source. The resulting mass spectrum was measured in ESI^{-} continuum mode (150-1000 Da). The mass spectra were corrected for baseline subtraction and smoothed. The peak intensities for mono-hexanoylated product and the substrate were measured and the product/substrate ratio calculated for each well.

Ag85C Inhibition Assay*:* Reaction mixtures were prepared containing trehalose, TDH and inhibitor (or buffer in the case of the control) at 37 °C. Reactions were initiated by Ag85C (20 uL) addition to give final concentrations of 20nM Ag85C, 500 μM trehalose, 500 μM TDH and 150 μM inhibitor in 1 mM TEA buffer (pH 7.2). Samples of reaction mixtures and standard (20 uM GlcNAc) were delivered to the mass spectrometer as described above and concentrations of hexanoate, TMH, trehalose and

TDH were monitored simultaneously. The reactions were monitored every 26.7 min. The total ion count (TIC) for each analyte and the standard were recorded for each injection. Reference to the standard curve for each compound allowed the concentration to be determined. Initial rates of acyltransfer were determined from initial rates of TMH and hexanoate production according to Eq.16 (*vide supra*).

Kinetic analysis of Ag85C inactivation by 60: The inactivation of Ag85C by **60** was initiated by addition of an aqueous solution of **60** to a solution of Ag85C in 10 mM TEA (pH=7.3) at 22 °C to give final concentrations of ~10 μM enzyme and 1, 2, 2.5, 3.5, 5 or 10 μM **60**. At timed intervals (10 sec – 4 min) after initiation, 0.25 μL aliquots were removed and added to 250 μL of a solution containing 250 μM trehalose and 250 μM TDH in 1mM TEA buffer (pH=7.3). The initial rate of TMH production was determined as described above and used to plot residual activity against time. The data were fitted to a first order exponential decay model and the first order observed rate constant (kobs) was determined for each concentration of **60**. No saturation was observed at concentrations up to 10μM **60** whereupon the rate of Ag85C inactivation became too rapid to accurately measure measure. A plot of the kobs versus [**60**] data yielded a straight line in the observed concentration range from which a second order rate constant kobs/Ki was determined to be 1420 ± 90 min^{-1} mM^{-1}.

9.5 Minimum Inhibitory Concentration (MIC) Determination

Analysis was conducted as previously reported[4]. *Mtb* cells were grown in 7H9 media to an OD 0.2-0.3 (then dilute sample 1:1000). On a 96 well round-bottom plate, 50 μl/well of medium was added to all wells except the first row. In the first row, 100 μl drug of choice was added at twice the initial desired concentration. Using a multichannel pipettor, 50 μL was transferred to each next row starting with row 1 and ending with row 11. 50 uL was discarded after row 11. To each well was added 50 μl of the 1:1000 culture dilution, approximating 1 x 10^4 bacteria per well. The plate was incubated for 2 weeks at 37 °C inside a small zip-lock bag. The plate was read with an inverted plate reader and graded as either growth, no growth, or growth/no growth if <50%. The well contents were then resuspended and the OD_{600} of individual wells was read on a fluorostar plate reader. MIC_{50} was calculated by plotting concentration (μg/ml) versus percentage growth (absorbance MIC well/ absorbance wildtype well) and fitting curves in Origin Pro 8 to determine MIC_{50} values.

9.6 Trehalase assay

Procedure was according to the manufacturer's instruction. Substrates were prepared to a final concentration of 10 mM in 0.05 M citric acid, pH 5.7. 5 μL of porcine trehalase (Sigma) stock solution (final 0.3 units/mL) was added into 50 μL of substrate solution. The reaction mixture was

incubated at 37 °C for 10 minutes. At 10 minutes, 60 µL of 0.1 M Tris, pH 7.5 was added to the solution to quench the reaction. Released glucose was measured with Glucose assay kit (Sigma, Glucose (HK) Assay Kit).

9.7 Mycobacterial cell culture

M. tuberculosis (H37RV strain from ATCC) was cultured in middlebrook 7H9 or Glycerol Alanine Salts with Tween (GAST medium) where appropriate. Cultures were monitored by optical density (OD) at 650 nm and were maintained at OD_{650} ~0.2 unless otherwise noted. *M. smegmatis* mc^2155 strain from ATCC was treated in an identical fashion.

9.7.1 Middlebrook 7H9

7H9 was made following the manufacterer's instructions. 4.7 g of the middlebrook 7H9 powder was suspended in 900 mL of purified water (containing 2 mL glycerol and 0.5 mL Tween 80). 100 mL of Middlebrook ADC Enrichment was added to the medium, which was sterilized by filtration (Nalgene 0.2 µm vacuum filtration unit). Samples were tested for sterility prior to mycobacterial culture.

9.7.2 GAST

GAST (glycerol-alanine-salts (GAS) medium was made per liter of the following: 0.3 g of Bacto Casitone (Difco), 4.0 g of dibasic potassium phosphate, 2.0 g of citric acid, 1.0 g of L-alanine, 1.2 g of magnesium chloride hexahydrate, 0.6 g of potassium sulfate, 2.0 g of ammonium chloride, 1.80 ml of 10 sodium hydroxide, and 10.0 ml of glycerol, 0.05 g of ferric ammonium citrate Tween 80 was added to 0.05%. Media was adjuted to pH 6.5 with 10 N sodium hydroxide solution and sterilized by filtration (Nalgene 0.2 µm vacuum filtration unit).

9.8 ^{14}C-Trehalose, ^{14}C-glycerol and ^{14}C-glucose uptake into H37Rv

H37Rv *M tuberculosis* cells were grown in Middlebrook 7H9 media and were harvested at an OD_{650} of 0.5 by centrifugation (1250 *g* at 4°C for 10 min), washed once with buffer (Hepes 25 mM and 0.05M Tween at pH 7.2) and then resuspended in the same buffer. Radio-labeled ^{14}C-trehalose (0.1 µCi/mL), glycerol (1 µCi/mL), and corresponding nonlabeled sugars (50 mM) were mixed and added to the 16ml cell suspension. 1 mL aliquots were removed in at 15, 30, 45, 60 and 120 min and were filtered through GA-4 Metricel filter membranes (diameter, 2.54 cm; pore size, 0.8 µm; Gelman Instrument Co., Ann Arbor, Mich.) The membranes were washed (3 × 2 mL) with HEPES/Tween buffer and LiCl (0.1 M, 1 × 2mL). Filter papers were then placed in Ultimax gold scintillation fluid. Scintillation counting was conducted for 2 min per vial. All experiments were conducted in triplicate and

samples without sugar were used as control. Uptake was measured over a 24 h time period with ^{14}C-glucose (0.9 µCi/mL), ^{14}C-Trehalose (0.6 µCi/mL) in 7H9 media, which contains 10 mM glucose. No unlabeled trehalose was added to the experiment. 1 mL aliquots were removed and counted as before with time points at 40 min, 90 min, 4 h and 24 h.

9.8.1 ^{14}C- and FITC-Tre lipid extractions

Lipid extractions from bacteria treated with ^{14}C-trehalose and acetate and were conducted based on previous reports[5] and analyzed by radiographic TLC. H37Rv *Mtb* cells were grown in 7H9 media to an OD_{650} of 0.8 whereupon ^{14}C-trehalose (10 µCi) was added to 15 mL of culture. ^{14}C-acetate (30 µCi/tube) or dual ^{14}C-acetate (30 µCi/tube) and FITC-trehalose (to final concentration of 100 µM) were added to 15 ml of culture and incubated overnight at 37 °C. Cells were harvested by centrifugation (3500 rpm for 10 min) and extracted for 8 hours into 2 mL (2:1 chloroform:methanol). The organic layer was removed and concentrated under a stream of argon and the residue was resuspended in 200 µL (80:20 chloroform:methanol) and 50 µL was spotted onto silica TLC plates, which were developed for 1 hour (75:25:4 or 80:20:2 Chloroform:methanol:water). Plates were scanned for fluorescence and exposed to a phosphor storage plate, which was scanned for radioactivity.

9.8.2 Isoniazid (InH) inhibition of mycolate production

H37Rv was grown in 7H9 media (15 mL) to an OD_{650} of 0.8 and InH was added to a concentration of 0.2 µg/mL (10 × MIC) or 0.02 µg/mL (1 × MIC). Culture was exposed to INH for 2 hours prior to addition of ^{14}C-acetate (150 µCi/tube). Cells were labeled with ^{14}C-acetate for 45 minutes and then harvested by centrifugation and extracted into 2 mL (2:1 chloroform:methanol). The organic layer was removed, concentrated and the residue was resuspended in 200 µL (80:20 chloroform:methanol) and 50 µL was spotted onto silica TLC plates, which were developed for 1 hour (75:25:4 chloroform:methanol:water). Plates were scanned for fluorescence and exposed to a phosphor storage plate, which was scanned for radioactivity.

9.8.3 Mycolate analysis

Mycolate analysis were conducted based on previous reports[5]. A full size image of the TDM TLC plate was printed onto transparency film and was used to identify each spot on the plate. Each spot was then scraped and the silica was placed in a vial with 0.5 mL 20% TBAH and heated to 100 °C for 18 hours. To this was added dichloromethane (0.5 mL) and methyl iodide (50 µL) and reaction was shaken vigorously at room temperature of 2 hours. Organic layer was removed to a new vial and concentrated under a stream of argon. The resulting white crystals were extracted with diethyl ether (2

× 250 μL) and the extract was placed in a fresh vial and concentrated under a stream of argon. The resulting clear oil was re-dissolved in 50 μL (2:1 toluene:acetonitrile) and the entire reaction product was spotted on silica TLC plate and developed with three successive runs (3 × petroleum ether:diethyl ether 85:15). Plate was then exposed to storage phosphor plate, which was scanned for radioactivity.

9.9 FITC-Tre and FITC-Glc uptake into *Mtb*

To CDC1551 *Mtb* or the Ag85C transposon mutant of CDC1551(BEI resources, TBVTRM, NR-13498) in Middlebrook 7H9 medium (10 mL) at an OD_{650} of 0.44 was added FITC-Tre **9** in ethanol to a final concentration of 100μM. Heat-killed (80 ˚C, 60 min) bacteria (2 × 0.5 mL) at an OD_{650} of 0.6 were used as control. *Mtb* were incubated at 37 °C with shaking. After 8 hours, the culture (4 × 400 μL) was harvested by centrifugation (1 min, 12000 rpm) and washed (3 x 1 mL 7H9 medium) and resuspended (200 μL 7H9). Fluorescence measurements were conducted in 96-well format in appropriate plates (Nunc, Cat No 137103 (Roskilde, Denmark)). The background of the culture was obtained from cells treated in an identical fashion in the absence of FITC-Tre. Experiments were conducted in quadruplicate. Uptake was normalized for increase in OD_{650}. FITC-Glc **254** uptake (100μM) into CDC1551 and FITC-Tre **40** uptake into *H. influenza*e, *P. aeruginosa* and *S. aureus* was conducted in an analogous manner.

9.10 FITC-Tre uptake into other bacteria

H. influenzae (AMC 36-A-1, ATCC 10211) was rehydrated in standard Blood Heart Infusion (BHI) media with supplemented hemoglobin and IsoVitaliX (BD 211876) and plated on chocolate agar plates and grown at 37 °C under 5% CO_2. Single colony was picked and grown in supplemented BHI media for 8 hours in the presence of 100 μm FITC-Tre. After 8 hours cells were harvested in quadruplicate in a manner identical to FITC-Tre *Mtb* uptake.

*Pseudomonas auruginosa (P*AO1-LAC, ATCC 47085) was grown on standard blood agar. Growth in standard BHI in the presence of 100 μm FITC-tre was conducted for 7 hours to an OD of 0.55 and harvest and fluorescence analysis was conducted as with *Mtb*.

Staphylococcus aureus (ATCC 13801) grown on standard blood agar. Growth in standard BHI in the presence of 100 μm FITC-tre was conducted for 6 hours to an OD of 1.8 and cells were harvested and diluted to and OD of 0.55 and fluorescence analysis was conducted as with *Mtb*. Background fluorescence was measured in the absence of probe in an identical fashion, and was averaged for all three bacteria.

9.11 Trehalose inhibition of labeling *in vitro*

In vitro, H37Rv *Mtb* expressing RFP were grown in 7H9 containing OADC to an OD 0.25, at which point FITC-Tre was added to a concentration of 200 µM, and trehalose was added to a concentration of 100 mM. The culture was incubated with shaking for 24 h. Culture was then harvested and washed (3 X 1 mL PBS) by centrifugation, fixed in 1:1 PBS: buffered formalin (10% in PBS) for 15 minutes. Culture was pelleted and resuspended in PBS (200 µL). Aliquots (10 µL) were placed on glass coverslips and the PBS was allowed to evaporate, following which 1 drop mounting media (invitrogen prolong gold) was added and the coverslip was placed on a glass slide. Z stacks (0.13 10 µm step size), zoom 8 were collected and images were deconvolved with Hugyens software and visualized with Imaris software. Bacteria grown in the absence of trehalose were treated in an identical fashion.

9.12 Polar labeling *in vitro*

H37Rv *Mtb*, expressing RFP were labeled and harvested in an identical fashion to the '*in vitro* trehalose inhibition labeling.' ROIs for the poles and mid-sections were quantitated using the 'marching cubes' tool in Imaris to manually generate surfaces from which statistics were calculated for the regions of interest. Data were collected from 150 bacteria.

9.13 Test of growth inhibition due to FITC-Tre in macrophages

J774 macrophages were grown to confluency in DMEM media and plated in black nunc 96 well plates (200 µL/well). Cells were infected with RFP expressing H37Rv strain with either 100 µL or 50 µL of bacteria in DMEM (OD 0.04). This infection afforded a MOI of approximately 5-10 bacteria per macrophage. Infection was allowed to proceed for three hours upon which time, macrophages were washed with fresh DMEM and FITC-Tre (10 mM stock in ethanol) was added to a concentration of 100 µM. Uninfected macrophages were treated in an identical fashion. As a second control, infected macrophages were treated with the equivalent volume of ethanol. RFP fluorescence was monitored (excitation 594 nm, emission 612 nm) at 0, 24, 48 and 72 hours. Error bars reflect the standard deviation for 4 experiments.

9.14 *Mtb* antibody labeling

Primary *Mycobacterium tuberculosis* antibody (ab905, AbCam, Cambridge, MA, USA) was labeled utilizing Zenon Rabbit Polyclonal IgG labeling kit (Z-25302, Molecular Probes, Eugene, OR, USA). Antibody and Zenon Fab fragments were prepared according to kit specifications and diluted to 1 mL in PBS with goat serum. Coverslips were incubated with antibody solution for 30 min at 37 °C, were then washed (1 mL, 3 × 5 min, 0.1% Tween 20 in PBS) and were fixed in 5% formalin for 15 min. Coverslips were drained and mounted in anti-fade media (Invitrogen Prolong Gold). Cells treated

with FITC-Tre and no antibody were exposed to the same mounting procedure, except second fixation step was omitted. *Alexa 594* labeled *Mtb* antibody and QDs were excited at 556 nm and emission spectra recorded at 594-665 nm.

9.15 Tre-Quantum dots labeling of *Mtb* infected macrophages

Tre Quantum dots (Tre-QDs) were treated in an analogous fashion to FITC-Tre and were added to give a final concentration of 7 nM (110 sugars/dot, 1 μM effective sugar concentration). Uncoated QDs were used as a control. Fixation, permeabalization, blocking, mounting and microscopy were conducted as with FITC-Tre.

9.16 FITC-Glc labeling of Mtb infected macrophages

Macrophages were grown to confluence on coverslips and infected with RFP expressing *Mtb* or BCG, using our standard protocol. FITC-Glc (1:1 ethanol/PBS stock solution) was added to a 100 μM concentration, and allowed to label overnight, upon which time cells were fixed, blocked and mounted, using the standard protocol developed. FITC-Tre treated cells were analyzed in parallel to confirm that bacteria were labeling with the trehalose compound. For both RFP BCG and RFP H37Rv, no specific labeling was observed with FITC-Glc.

9.17 Trehalose inhibition of FITC-Tre labeling in vivo

In vivo, A sterile solution of D-trehalose in PBS was added to J774 macrophages infected with RFP BCG to a final concentration of 10 mM. Concurrent with the addition of trehalose, FITC-Tre was added to a final concentration of 200 μM. Fixation, permeabalization, blocking, mounting and microscopy were conducted as with FITC-Tre.

9.18 Primary bone marrow macrophages (BMM) isolation[6]

Six-week-old C57BL/6 female mice were used for these experiments. Mice were sacrificed and their femur bones and shoulders were removed by dissetction. Intact bones were then washed in 70% ethanol and placed in Dubelco's MEM media (DMEM) containing 20% fetal bovine serum (FBS), 1% HEPES, 1% pyruvate and 1 mmol/L L-glutamine on ice until use. Ball joints were removed and marrow was flushed from bones use a 25 gauge needle and syringe filled with ice cold DMEM media. The sterile needle was inserted at both ends of the femur to ensure that openings were present and approximately 2 mL DMEM was used to flush marrow. An additional 2 mL was applied to the other end of the bone to ensure full removal of marrow and process was repeated for all bones. Cell suspensions on ice were pipetted multiple times to break up aggregated cells. Then cell suspensions were passed through a 70 micron filter and resulting suspensions were counted on standard

hematocytometer. DMEM supplemented with 20% FBS, pyruvate, HEPES and glutamine was added to the cell suspension. DMEM also contained a 2× concentration recombinant mouse macrophage colony stimulating factor (M-CSF) Prospec (CYT-439) or Biovision (4238-50) so that the final working concentration was 100 U/mL. DMEM al so contained 2× concentration penicillin-streptomycin and 2 × mercaptoethanol (10^{-4}M) so that final concentration of cell suspension was 1× for all additives. 15 mL cell suspension was plated in untreated Greiner (100 × 20 mm) dishes. On day 3, 10 mL of DMEM complete with 20% FBS and M-CSF (100 ng/mL) was added to each dish. After 5 or 6 days plates were washed with cold PBS (5 mL), followed by cold complete DMEM containing 10% Fetal bovine serum, 1% HEPES, 1% pyruvate and 1 mmol/L l-glutamine (5 mL). Cells were then gently scraped in one direction and resulting cell suspension was centrifuged (8 min at 800 rpm). Cells were counted on standard hemocytometer and 1.6 $x10^6$ cells were added per petri dish. Cell pellets were resuspended in 20 mL DMEM and plated on sterile coverslips. INF-γ was added to a concentration of 4 ng/mL. BMMs were then infected with RFP H37Rv, labeled with FITC-Tre (200 µM, 40 h), treated with various antibodies, mounted and imaged in the same manner as with J774 macrophage cell line.

9.19 Immunocytochemistry (ICC) procedure

Media was removed and cells were washed in PBS (1 mL, 3 × 2 min). Cells were fixed in 5% formalin (1:1 in PBS) for 15 minutes and then washed again with PBS (1 mL , 2 × 2 min). Cells were permeabalized (1 mL, 0.1% triton X-100 in PBS) for 6 minutes at room temperature and then washed with PBS (1 mL, 3 × 2 min). Non-specific protein interactions were blocked with protein blocker (1 mL) with 1 drop goat serum for 1 hour at RT. The following primary antibodies were employed: Rabbit polyclonal to *LAMP1* (ab25170), Rabbit polyclonal to EEA1 (ab2900-100), rabbit polyclonal to Rab5 (ab18211) rabbit polyclonal to Rab14 (ab28639), chicken polyclonal to Ag85A (ab14073), rabbit polyclonal to Ag85B (ab43019) were obtained from Abcam. Rabbit polyclonal to Pro-cathepsin D (sc-10725) was obtained from Santa Cruz biotechnologies. Rabbit polyclonal to Ag85 was obtained from BEI resources (NR-13800). Antibodies were diluted 1:100 or 1:50 in PBS with 2 drops goat serum. Blocking agent was removed and coverslips were incubated with diluted antibodies (200 µL) for 16 hours at 4 °C. Coverslips were then washed (1 mL, 2 × 5 min, 0.1% Tween 20 in PBS) and then (1 mL, 5 min) in PBS. Secondary antibodies Goat anti rabbit IGG 633 (A21070) and goat anti chicken IGG 633 (A21103) were obtained from Invitrogen. Fresh 1:75 solutions of were made in PBS with 2 drops goat serum. Coverslips were labeled for 50 minutes with secondary antibody and then were washed (1 mL, 2 × 5 min 0.1% Tween 20 PBS) followed by DAPI (200 µL, 1 × 5 min, in PBS) and then PBS (1 mL, 5 min). Coverslips were then washed once further in PBS (1 mL, 5 min), drained and mounted in anti-fate media (Prolong GOLD, Invitrogen).

9.20 Image collection and processing

Multiple fields were sampled, and representative images were recorded. DAPI was first excited with a 405 nm diode laser, and emission spectra were recorded at 413-470 nm prior to other fluorophores. FITC-Tre was excited at 488 nm and emission spectra were recorded at 497-557 nm. RFP was excited at 594 and emission spectra recorded at 602-651 nm. Alexafluor 633 was excited at 633 nm and emission spectra recorded at 651-759nm. Essential sequential Z sections of stained cells were also recorded for generation of stacked images through cell 63X objective with a zoom of 4 and 512x512 image size. A 3-D volume was constructed from sequential z-sections of a cells assembled into a 3D volume in Imaris software (version 7.0.0, Bitplane AG, Zurich, Switzerland). All collected images for analyses were deconvolved by Huygens Essential software (Version 3.4, Scientific Volume Imaging BV, Hilversum, The Netherlands). Identical deconvolution parameters were employed for each antibody dataset. Percentage colocalization was calculated using the colocalization function in Imaris, with thresholding by RFP channel. RFP channel was used to generate regions of interests (ROIs) as surfaces for all bacteria and statistic function was utilized to export for mean fluorescence and volume of all channels within the ROI. These values were plotted as mean antibody florescence (ex. 633 em. 651-759 nm) versus mean FITC-Tre fluorescence (ex. 488 em. 497-557 nm).

9.21 Correction for 63X blue offset

For some image collection the 63X blue objective was mistakenly used, rather than the 63X regular objective. This error offset the colors slightly in the Z direction and offset the far red colors more than the green or near red. This offset was subtle and difficult to visually portray (**Figure 9.8**). The use of the blue objective caused pixels that should have colocalized perfectly to be slightly set off from one another, pink on the bottom, then red then green on top with the far red shifted the most. Shifting the colors affected the colocalization results, particularly for the Rab5. Because the green was no longer in the same plane as the pink, lower colocalization between the antibody and the green channel, rather than the antibody and the red channel, was observed.

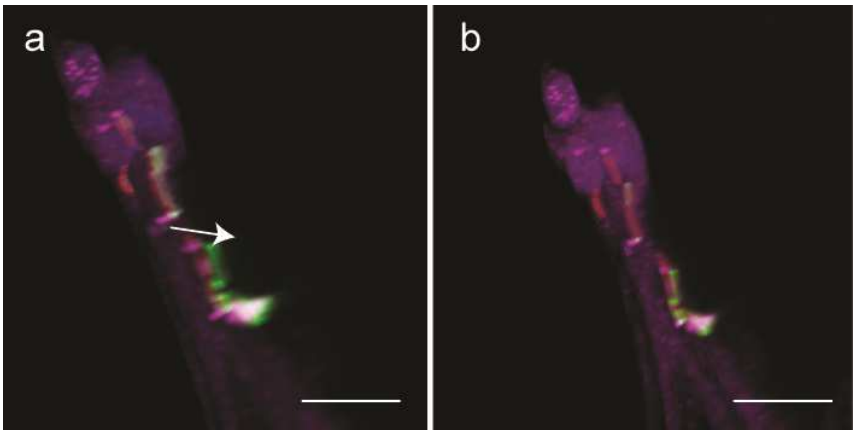


Figure 9.8. Color shift caused by blue objective (a) uncorrected. Arrow indicates direction of shift. (b) corrected. Scale bar 5 µm.

To correct this problem a tri-colored bead of known size was employed and the offset was measured. Then the batch function was used to offset all of the images by 2 voxels for pink, and 1 voxel for red. All of the colocalization studies were redone to obtain more meaningful data and additional images were gathered, utilizing the 63X objective, to validate the offset method. No significant difference in colocalization data was observed between the batch offset images and the images collected with the correct objective.

9.22 Data compilation

To expedite the data analysis a series of MATLAB programs were developed in order to compile the data into matrices from CSV files. These programs facilitated the extraction and combinination of a large volume of csv files without actually having to open and copy-paste from each file.

Function *Ugly1* imports data from CSV files that do not get delimited clearly by matlab, meaning files where the commas are not automatically recognized and placed into separate boxes. Matlab did not recognize the data in the ROI files properly and instead placed the data in the text portion. Program *Ugly1* solved this problem.

```
function data = ugly1(file).
A = importdata(file);
for i = 5:length(A.textdata)
end
```

A second function, *importother2*, was used to look at the bacteria as individual regions of interest and compile all of the data for the individual bacteria. To do this, first all of the data for each image collected was compiled. Images were numbered to refer to which slide they were collected from and

then lettered to distinguish the images. *Importother2* combines mean values for bugs into one matrix and resaves it in a new directory, as specified.

```
function importother2(file, file3, file4, file5, file2)
addpath ('C:\Documents and Settings\keriann\Desktop\deconvolution backup may 27')
newData1 = importdata(file, ';'); %loads file
vol = ugly1(file)
greenmean = ugly1(file3)
redmean = ugly1(file4)
pinkmean = ugly1(file5)
  M = [vol; greenmean; redmean; pinkmean]
save (file2, 'M')
cd ('C:\Documents and Settings\keriann\Desktop\deconvolution backup may 27\Cath D extracted data\bug stats')
save (file2, 'M')
end
```

Function *Cathd* extracts from percentage coloc files, allowing for the compilation of all percentage colocalization data. This function was renamed and rewritten for each antibody to save the data in an antibody specific directory, so as not to accidentially overwrite previous files.

```
function CathD(file, file3, file4, file5, file2)
newData1 = importdata(file, ';'); %loads file
percentvolumegreen = newData1.data(7) %extracts value from line 7
percentmaterialgreen = newData1.data(9) %extracts value form line 9
%save the two variables as a .mat file using the same name as the original
%file
newData2 = importdata(file3, ';'); %loads file
percentvolumered = newData2.data(7) %extracts value from line 7
percentmaterialred = newData2.data(9) %extracts value form line 9

newData3 = importdata(file4, ';'); %loads file
percentvolumegreennodecon = newData3.data(7) %extracts value from line 7
percentmaterialgreennodecon = newData3.data(9) %extracts value form line 9

newData4 = importdata(file5, ';'); %loads file
percentvolumerednodecon = newData4.data(7) %extracts value from line 7
percentmaterialrednodecon = newData4.data(9) %extracts value form line 9

%save the two variables as a .mat file using the same name as the original
%file
save (file(1:end-4), 'percentvolumegreen', 'percentmaterialgreen') %saves in current folder
save (file3(1:end-4), 'percentvolumered', 'percentmaterialred') %saves in current folder
save (file4(1:end-4), 'percentvolumegreennodecon', 'percentmaterialgreennodecon') %saves in current folder
save (file5(1:end-4), 'percentvolumerednodecon', 'percentmaterialrednodecon') %saves in current folder
cd ('C:\Documents and Settings\keriann\Desktop\deconvolution backup may 27\Cath D extracted data\cathD green coloc decon')
save (file2, 'percentvolumegreen', 'percentmaterialgreen') %saves in folder specified by path - currently C:\Users\sbackus\Documents
cd ('C:\Documents and Settings\keriann\Desktop\deconvolution backup may 27\Cath D extracted data\cathd red coloc decon')
save (file2, 'percentvolumered', 'percentmaterialred') %saves in folder specified by path - currently C:\Users\sbackus\Documents
cd ('C:\Documents and Settings\keriann\Desktop\deconvolution backup may 27\Cath D extracted data\cathD green coloc no decon')
save (file2, 'percentvolumegreennodecon', 'percentmaterialgreennodecon') %saves in folder specified by path - currently C:\Users\sbackus\Documents
```

```
cd ('C:\Documents and Settings\keriann\Desktop\deconvolution backup may 27\Cath D extracted data\cathD red coloc no decon')
save (file2, 'percentvolumerednodecon', 'percentmaterialrednodecon') %saves in folder specified by path - currently C:\Users\sbackus\Documents
cd ('C:\Documents and Settings\keriann\Desktop\deconvolution backup may 27\cathD second round')
end
```

Then to combine all of the matrices into a master matrix, a script, rather than a program, was used because it allowed for the results to be reviewed prior to saving. The following script was written and implemented:

```
%%
ii = 1
for i = 1:4 %where n is the number of files
   if exist([num2str(i) 'A.mat']) == 2
      load([num2str(i) 'A.mat'])
      mat(ii) = percentmaterialgreen
      vol(ii) = percentvolumegreen
      clear percentmaterialgreen percentvolumegreen
      ii = ii+1
   else
   end
   if exist([num2str(i) 'B.mat']) == 2
      load([num2str(i) 'B.mat'])
      mat(ii) = percentmaterialgreen
      vol(ii) = percentvolumegreen
      clear percentmaterialgreen percentvolumegreen
      ii = ii+1
   else
   end
   if exist([num2str(i) 'C.mat']) == 2
      load([num2str(i) 'C.mat'])
      mat(ii) = percentmaterialgreen
      vol(ii) = percentvolumegreen
      clear percentmaterialgreen percentvolumegreen
      ii = ii+1
   else
   end
   if exist([num2str(i) 'D.mat']) == 2
      load([num2str(i) 'D.mat'])
      mat(ii) = percentmaterialgreen
      vol(ii) = percentvolumegreen
      clear percentmaterialgreen percentvolumegreen
      ii = ii+1
   else
   end
      if exist([num2str(i) 'E.mat']) == 2
      load([num2str(i) 'E.mat'])
      mat(ii) = percentmaterialgreen
      vol(ii) = percentvolumegreen
      clear percentmaterialgreen percentvolumegreen
      ii = ii+1
   else
   end
    if exist([num2str(i) 'F.mat']) == 2
      load([num2str(i) 'F.mat'])
      mat(ii) = percentmaterialgreen
      vol(ii) = percentvolumegreen
```

```
    clear percentmaterialgreen percentvolumegreen
    ii = ii+1
  else
  end

end
```

9.23 *In vivo* labeling experiments in rabbits and mice

9.23.1 Animals and ethics assurance

This study was carried out in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The Committee on the Ethics of Animal Experiments of the National Institute of Allergy and Infectious Disease approved the experiments described herein under protocol LCID-3 and all efforts were made to minimize suffering. New Zealand White (NZW) rabbits and C57BL/6 female mice were individually housed in cages in a biological level 3 animal facility approved for the containment of *Mtb*.

9.23.2 Labeling of mouse and rabbit tissue

Tissue was obtained from female C57BL/6 mice (3 weeks post aerosol infection) as well as NZW rabbit (120 days post infection).

9.23.3 *In vivo* labeling with FITC-Tre in infected mice

Mice were treated with the FITC-Tre compound at four separate doses: double high dose was 4 mg at 8 am, followed by a boost of 4 mg at 2 pm. High dose was 4 mg. Double low dose was 2 x 0.4 mg and low dose was 0.4 mg. Uninfected mouse was treated with the double high dose as control. Likewise, one infected, untreated mouse was taken as an additional control.

All doses were in administered in 200 μL PBS with 10 μL ethanol. All mice were sacrificed after 24 hours by snapping their necks. Their organs were mounted on plates using Reichert-Jung mounting medium (Cambridge Instruments, Nussloch, Germany) and frozen in isopentane/dry ice. Ten to fifteen micrometer thick cryosections were obtained using a Reichert Leica CM3000 Cryostat.

Slices were fixed in 5% formalin for 20 minutes washed 2 x 3 minutes in PBS, permeabalized for 25 minutes in 0.1% Triton X80, washed 3 x 2 minutes in PBS and blocked overnight at 4 °C. Several slices were treated with antibody, utilizing the same protocol as the macrophages were subjected to and mounted with DAPI anti-fade media (Prolong Gold® Mounting Medium**)**

9.23.4 *In vitro* labeling: FITC-Tre labeling of tissue slices with and without Ag85.

10-15 µm thick cryosections were obtained using a Reichert Leica CM3000 Cryostat. In a humidity box, tissue slices from rabbits infected with *Mtb* were incubated at room temperature for 12 h with FITC-Tre in PBS buffer at 0.5 mg/ml, 0.05 mg/ml and 0.005 mg/ml.

For enzyme labeling FITC-Tre was used at 0.05 mg/ml and a 100 µM stock solution of enzyme was used and either 2 µL or 20 µL of enzyme was applied. Reactions were left for 12 hours at room temperature at which point FITC-Tre was removed with 3 x 2 min PBS wash and tissue was fixed in formalin 5% for 15 minutes.

9.23.5 Auramine stain

Tissue was incubated for 15 minutes with auramine stain following the manufacturers protocol (Auramine O stain Kit, Polyscience, Inc, 24665), following which the auramine was washed away extensively with running water. Slices were mounted with DAPI anti-fade media (VECTASHIELD® **Mounting Medium).**

9.23.6 PET experiment with ^{18}F-fluorobenzyl trehalose[ii]

9.23.6.1 Anesthesia and parameters for PET/CT scanning

In preparation for scanning, the rabbit was removed from its home cage, weighed and moved into the scan room at least 24 hours prior to the scan. Food intake prior to the scan was reduced. For intravenous (IV) access, a gauge 22 – 24 indwelling catheter was inserted into the marginal ear vein after local anesthesia was applied, a small blood sample collected and blood glucose was measured (Accu-Chek glucose meter). The normal reported was range 75-150 mg/dl and measurements fell within this range. The [^{18}F] fluorobenzyl-trehalose ([^{18}F] FBTre) metabolic marker was injected IV (1 mCi/kg) and allowed to distribute into the organs for 50-60 min. The actual dose delivered was determined by measuring radioactivity content of the syringe filled with the [^{18}F] FBTre solution immediately before and after injection into the catheter. This value was entered into the microPET Manager software (Siemens Preclinical Solutions, Knoxville, TN) along with the weight of the animal. Upon follow up scans, the uptake period and dose were matched within five minutes and 10%, respectively to facilitate quantitative assessment of [^{18}F] FBTrePET uptake parameters.Thirty minutes into the uptake period, a saline infusion was connected and secured to the IV access.

[ii] All rabbit infections and PET scanning experiments were conducted by Dan Schimel, Danielle Weiner, Dr. Laura Via and Emmanuel Dayo. Experimental procedures are reproduced here with their permission.

Ketamine 3.3 mg/kg and Xylazine 1 mg/kg combination was injected slowly IV through the infusion line followed by a 15 mL/hr Saline infusion drip. Upon sedation, the animal was transferred to the heated (BioVet CT System, m2m Imaging Corp., Newark, NJ) scanning bed (Focus 220 CereTom bed Agile Engineering, Knoxville, TN) designed for use with both the PET and CT scanners without repositioning the animal. Fiduciary markers containing 0.1 μCi of Ge-68 were placed over the upper thoracic vertebrae and below the sternum. The animal was placed sternal, feet first prone and maintained under 1 – 5 % isoflurane at 1 liter/min flow of medical O_2 delivered though a cuffed lubricated laryngeal mask (LMA). Anesthesia was maintained and monitored for the duration of the scans. EKG leads, pulse oximeter sensor, and thermister probes were connected to monitor vital signs to the duration of the anesthesia.

Once the animal was stable under anesthesia, the bed was advanced through the microPET scanner (MicroPET Focus 220, Siemens Preclinical Solutions, Knoxville, TN). and into the HR-CT scanner. An 8-slice helical CT scanner (CereTom, NeuroLogica, Danvers, MA) with a Scanner Drive System (SDS) was used in this study. TheSDS enables the scanner to move during acquisition and allows the animal bed to remain stationary for the duration of the scan. Beginning at the base of the animal's head, a 250 mm long scan though the lung and into the abdominal cavity was acquired (~25 seconds) with technical parameters of 120 kV, 5mA/second, 1.25mm slice thickness and 0.4949 mm pixel spacing. The animal's breathing was held 5 seconds after the scan was started for 15 seconds at 8-10 cm H2O with a small animal ventilator (TOPO Dual Mode Ventilator, Kent Scientific Corp., Torrington, CT) to capture full inflation of the lungs for optimal visulation of the lesions. A second, contrast-enhanced CT scan was collected immediately after injection of a non-ionic, water soluble, tri-iodinated x-ray contrast agent (2 mL/kg) containing 300 mg Iodine/mL. The contrast agent was used to enhance visualization of blood vessels and lymph nodes.

The animal bed was then repositioned into the microPET gantry with a 7.5 cm axial field of view (MicroPET Focus 220, Siemens Preclinical Solutions, Knoxville, TN). Sixty minutes post [^{18}F] FBTre injection, the same tape marker used for the HR-CT was used as the PET reference point. A series of 3, 10 minute emission scans with 75 mm thick slices (windows) with a 20 mm overlap were collected were acquired caudal to cranial. Vital signs such as O_2 saturation, CO_2 saturation levels, respiratory rate, heart rate and temperature were continuously monitored throughout the procedure until the animal was recovered and ambulatory. Feed and water was provided in the scan room for at least 24 hours until the radiation levels returned to background, when the animal could be returned to its home cage.

Histograms of the emission list mode data were formed into 3D sinograms. All emission scans were corrected for random coincidences and normalized for detector efficiencies. The CT image was fused to the PET image using ASIPro VM (Siemens Medical Solutions USA, Inc., Knoxville, TN) and then

the attenuation correction data was obtained from a segmented μ-map of the CT image. Images were reconstructed using a Fourier rebinning (FORE) algorithm and an ordered subset-expectation maximization (OSEM2D) algorithm with attenuation correction and 6 iterations. Images were smoothed by using a Gaussian kernel with an FWHM of 2 mm which is 2.22 mm in the transverse and 2.51 mm in the axial directions. Reconstructed pixel size, in cm = (((X_crystal_pitch /918 2.0) * X_dim)/(image_size * zoom)) * (effective_radius / radius). The CT and PET scanners were calibrated monthly and daily prior to scanning as recommended by the manufacturers.

9.23.7 Pharmacokinetic Studies in Uninfected Animals

Blood was collected in heparinized tubes at 0, 0.25, 0.5, 0.75, 1, 2, 4, and 8 h post-dosing of uninfected mice. FITC-Tre was dosed at 3 mg/kg and Fluorbenzyl-trehalose was dosed at 4.7 mg/kg in phosphate buffered saline. Samples were then centrifuged for 10 min at 5000g, and plasma was stored at -80°C until analysis. To 5 mL of plasma, 1 mL of methanol and 44 mL of acetonitrile were added and the mixture was vortexed for 2 minutes, centrifuged at 10,000g for 10 minutes, and the supernatant was removed to a clean tube. Supernatants were transferred to 96-well plates and lyophilized. HPLC coupled to tandem mass spectrometry (LC/MS/MS) analysis was performed on a Sciex Applied Biosystems API4000 triple-quadrupole mass spectrometer coupled to Spark Symbiosis Pharma HPLC system. Gradient elution conditions with an Agilent Zorbax Phenyl 991 x 75mm 3.5μ column were used. The mobile phase A was 0.2% acetic acid (FITC-Tre) or 0.2% ammonium hydroxide (Fluorobenzyl-Tre) in deionized water and mobile phase B was 0.2% acetic acid or 0.2% ammonium hydroxide, respectively, in acetonitrile. Multiple reaction monitoring of parent/daughter transitions in electro-spray positive ionization mode were used to track the presence of analytes. Data processing was performed using Analyst software 1.4.2 (Applied Biosystems Sciex). Pharmacokinetic parameters were calculated using non-compartmental analysis with the WinNonLin software (Pharsight, Mountain View, CA).

9.24 Molecular dynamics

Protein structures were explicitly solvated with TIP3P water molecules and Na+ and Cl- counterions and all mutations were generated using the Visual Molecular Dynamics program (VMD, University of Illinois at Urbana-Champaign)[7]. Molecular dynamics simulations were performed under isobaric-isothermal conditions using the NAMD[8] program (v.2.7) on the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD (http://biowulf.nih.gov) and included periodic boundary conditions. Mutations were made in Electrostatic interactions were handled using a Particle-Mesh Ewald summation. The CHARMM27[9] forcefield was used with CHARMM atom types and charges. Prior to the start of the simulation, a system was minimized using a conjugate gradient method, followed by slow warming to 310°K in 10°K increments. Each increment ran for 5 psec which was

enough time for a system to equilibrate at a given temperature. Once the 310°K target temperature was reached, a system was further equilibrated for 50 psec. Production runs were conducted at 310°K and date was obtained for 20 nsec. For all simulations, a 1 fsec integration timestep was used along with a 12Å cutoff. Langevin dynamics were used to maintain temperature and a modified Nosé-Hoover Langevin piston was used to control pressure.

9.25 Phylogenetic Analysis

Sequences of all annotated Ag85 proteins were downloaded from the tuberculosis database (TBDB.org) and aligned by ClustalW in DNAstar laser gene MegAlign software.

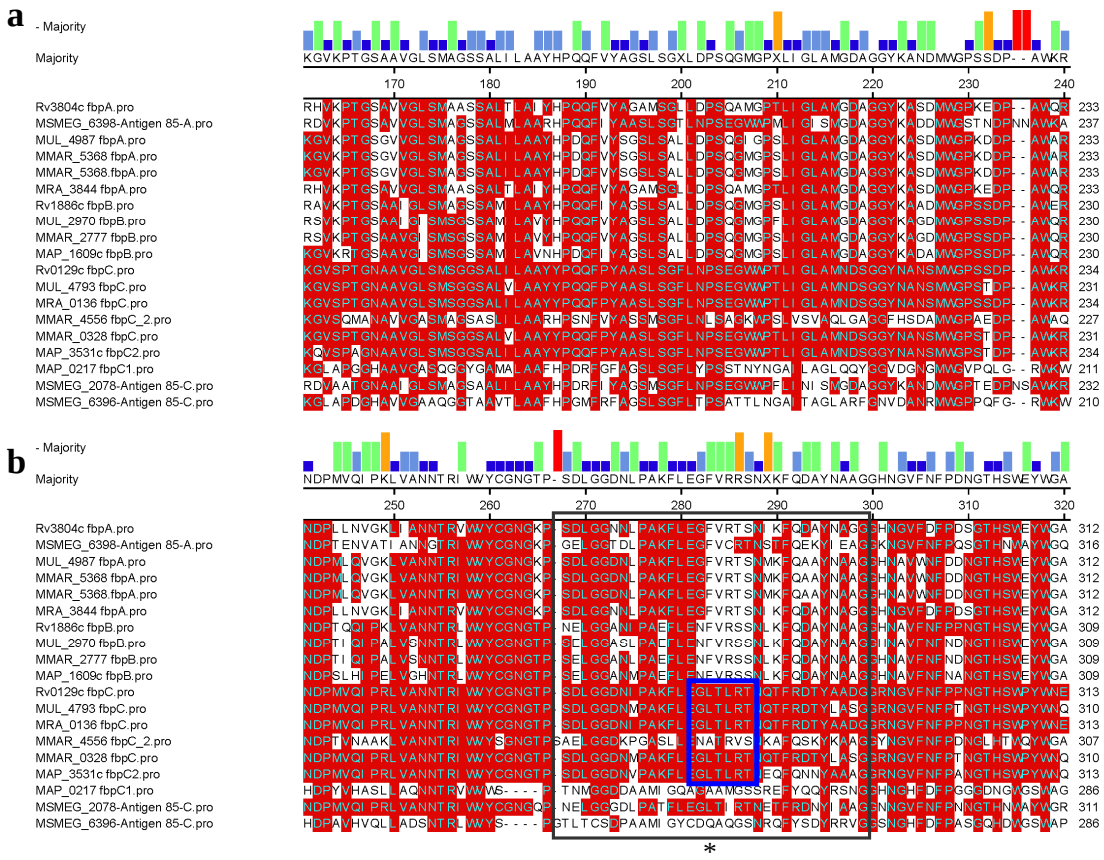


Figure 9.9. Clustal W alignment of Ag85 sequences for all annotated Ag85 genes in *M. smegmatis* (Msmeg), *M. avium M. paratuberulosis* (MAP), *M. marinum* (MMAR), *M. tuberculosis* H37Ra (MRA), *M. ulcerans* (MUL), and *M. tuberculosis* H37Rv (RV). Sequences were compared to Rv0129c (Ag85C). Red indicates homology. White indicates regions where sequences differed from Ag85C. Sequences are organized A (top), B(middle), C (bottom). (a) box indicates region of Ag85 dimer adjacent to pocket (b) helix 8, region that moves in "transition state." blue box indicates region that contacts secondary trehalose directly. Residues with * indicate those investigated by site directed mutagenesis.

9.26 Site Directed Mutagenesis

Plasmids were obtained from Colorado State tuberculosis research contract HHSN266200400091c as Rv3804/Ag85A Rv1886c/Ag85B Rv0129c/Ag85C in pET23b vector containing a His-tag and were

expressed as has been described above. Mutagenesis was conducted using a Stratagene Quikchange lightning kit according to the manufacturer's instructions. Primers were obtained from Invitrogen and were designed to have appropriate length and melting temperature. Reaction conditions were prepared as follows: 5 µl of 10 × reaction buffer, 1 µl (10 ng) of Ag85 plasmid. 1 µl (1 µg) of oligonucleotide primer #1 [XX-mer (1 µg /µl)] and 1 µl (1 µg) of oligonucleotide primer #2 [XX-mer (1 µg /µl)] (see **Table 9.1** for primers). 1 µl of dNTP mix, 3 µl of QuikSolution reagent 38 µl ddH2O (to bring the final reaction volume to 50 µl). Finally 1 µl of QuikChange® Lightning Enzyme was added directly before reaction was started. PCR cycle was run as follows: 95 °C for 2 minutes, followed by 20 X (95 °C for 20 seconds, 60 °C for 10 seconds, 68 °C for 5 minutes) then 68 °C for 5 minutes. Entire reaction mixture was digested with DPN enzyme (1.5 µL) 30 minutes. Electrocompetent cells (DH10β) were aliquoted (45 µL) and thawed on ice. 2 µL reaction mixture was added to each tube of competent cells. Competent cells were left with reaction mixture for 5 minutes on ice and then were transferred to 1 mm cuvette (Biorad) and electroporated; following which 0.5 mL SOC media was added. Cells were grown at 37 °C for 1 hour, and plated on LB plates containing appropriate antibiotic. Colonies were harvested after overnight incubation at 37 °C and plasmid was purified by Miniprep™ (Qiagen).

9.26.1 Mutagenesis of Ag85C

Table 9.1. Primers used for site directed mutagenesis of Ag85C in pET23b vector. Red indicates site of mutation. Mutations convert individual amino acids to those found in Ag85B or to alanine. Multiple rounds of mutagenesis with the corresponding primers were used to generate plasmids with triple mutations in distal sites.

Primer	Sequence 5' to 3'	Mutation
kb-16	CAACCCGTCCGAGGGC**GCGGCG**CCGACGCTGATCGGC	W157A and W158A
kb-17	GCCGATCAGCGTCGG**CGCCGC**GCCCTCGGACGGGTTG	W157A and W158A
kb-22	CAACCCGTCCGAGGGC**ATGGGC**CCGACGCTGATCGGCC	W158G and W157M
kb-23	GGCCGATCAGCGTCGG**GCCCAT**GCCCTCGGACGGGTTG	W158G and W157M
kb-44	CTCAACCCGTCCGAGGGC**GCG**TGGCCGACGCTGATCGGC	W157A
kb-45	GCCGATCAGCGTCGGCCA**CGC**GCCCTCGGACGGGTTGAG	W157A
kb-46	CAACCCGTCCGAGGGCTGG**GCG**CCGACGCTGATCGGCCTGGC	w158A
kb-47	GCCAGGCCGATCAGCGTCGG**CGC**CCAGCCCTCGGACGGGTTG	W158A
kb-52	CGAAGTTCCTGGAAGGC**GCC**ACCCTGCGCACCAACCAGAC	L230A
kb-53	GTCTGGTTGGTGCGCAGGGT**GGC**GCCTTCCAGGAACTTCG	L230A
kb-56	GGCGAAGTTCCTGGAAGGC**TTC**ACCCTGCGCACCAACCAGACC	l230F
kb-57	GGTCTGGTTGGTGCGCAGGGT**GAA**GCCTTCCAGGAACTTCGCC	L230F

9.26.2 Expression of Ag85C mutant proteins

LB Amp^{100} (30 mL) was inoculated with desired plasmid in BL21 DE3 strain and grown overnight at 37 ˚C with shaking to generate a starter culture, which was used to inoculate 1 L of autoclaved Amp^{100} LB broth. Culture was grown at 37 ˚C with shaking until it reaches an OD_{600} of 0.5-0.8 and removed from 37 ˚C and place in 18 ˚C incubator with shaking and allowed to cool to 18 ˚C. Expression was induced with isopropyl-β-D-thiogalactopyranoside (IPTG) (0.25 mM) for 16 hours. Cells were harvested and resuspended in buffer (50 mM Tris-HCl, 250 M NaCl, 10 mM imidazole pH 8) containing Complete® protease inhibitor cocktail (Roche) and lysozyme. Pellet was lysed by sonication and pelleted at 20,000 rpms for 15 min. Purification was carried out over nickel NTA agarose resin (invitrogen). 2 mL resin slurry was equilibrated with binding buffer (50mM Tris-HCl, 250mM NaCl, 10mM imidazole, pH 8). Lysate was applied to column slowly (1 mL/min) and then purified using a stepwise gradient of wash buffer (50 mM Tris-HCl, 25 mM Imidazole, 250 mM NaCl, pH 8) (20 column volumes followed by elution (50 mM Tris-HCl, 250 mM Imidazole, 250 mM NaCl, pH 8) (3 mL). Eluant was then switched into 20 mM TEA buffer pH 7.4 with PD10 desalting cartridge (Amersham). SDS-page was performed on all protein fractions. PD-10 desalting cartridges (Amersham) were utilized to remove endotoxins and to buffer exchange into 20 mM TEA pH 7.4 for mass spectrometry assay. Protein concentrations were calculated by standard BCA page assay as well as by absorbance on a nanodrop. Mass spectrometry and sequencing analysis confirmed the correct identity of the proteins.

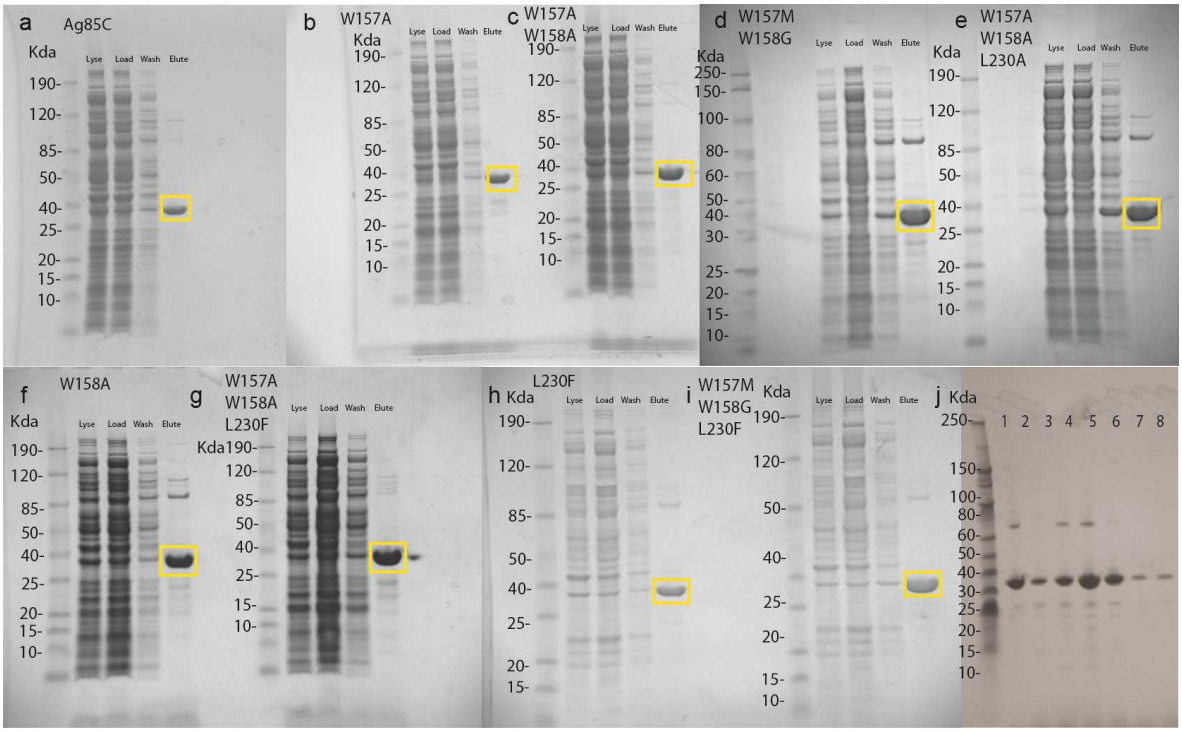


Figure 9.10. (a) Overexpression of Ag85C (b) and Ag85C containing the W157A mutation (c) W157A and W158A double mutant (d) W157M and W158G double mutant (e) W157A, W158A and L230A triple mutant (f) W158A mutant (g) W157A, W158A and L230A triple mutant (h) L230F mutant (i) W157M, W158G and L230F triple mutant with 'Ag85B-like' type (j) protein gel of selected mutants to compare size. 1-Ag85C, 2-W157A, 3- W158A, 4-W157A W158A, 5-W157M W158G, 6- W157M W158G L230F, 7-W157A W158A L230F 8- W157A W158A L230A. Boxes indicated purified protein. Ladder (New England BioLabs 10-250 Kda or Invitrogen Benchmark prestained protein ladder).

9.26.3 Circular Dichroism

Circular dichroism was measured in a Jasco J-715 spectropolarimeter, scanning four times from 260 - 190 nm at 50nm/min, time constant =1 s, bandwidth = 1nm, slit width = 500um. Protein solutions in 10mM phosphate buffer (protein concentrations approximately 100ug/mL) were held in 1 mm path length quartz cuvettes. Because of uncertainty about the protein concentrations, secondary structures were estimated as described in McPhie (2008, Analytical Biochemistry, vol 375, pp 379-381).

9.26.4 Mutagenesis of Ag85B

Table 9. 2. Primers used for site directed mutagenesis of Ag85B in pET23b vector. Red indicates site of mutation. Mutations convert individual amino acids to those found in Ag85C.

Primer	Sequence 5' to 3'	Mutation
kb-187	CCCGCCGAGTTCTTGGAGAACCTGGTTCGTAGCAGCAACCTGAAG	85B F231L
kb-188	CTTCAGGTTGCTGCTACGAACCAGGTTCTCCAAGAACTCGGCGGG	85B F231L
kb-189	CCCGCCGAGTTCTTGGAGGGCCTGGTTCGTAGCAGCAACCTGAAG	85B F231L N230G
kb-190	CTTCAGGTTGCTGCTACGAACCAGGCCCTCCAAGAACTCGGCGGG	85B F231L N230G
kb-191	GGCCGATCAGGCTAGGCCACCACCCCTGAGAGGGGTCC	85B M158W G159W
kb-192	GGACCCCTCTCAGGGGTGGTGGCCTAGCCTGATCGGCC	85B M158W G159W

9.26.5 Overexpression of Ag85B mutants

Overexpression and purification of Ag85B mutants was conducted as with Ag85C mutants (see previous section for methods).

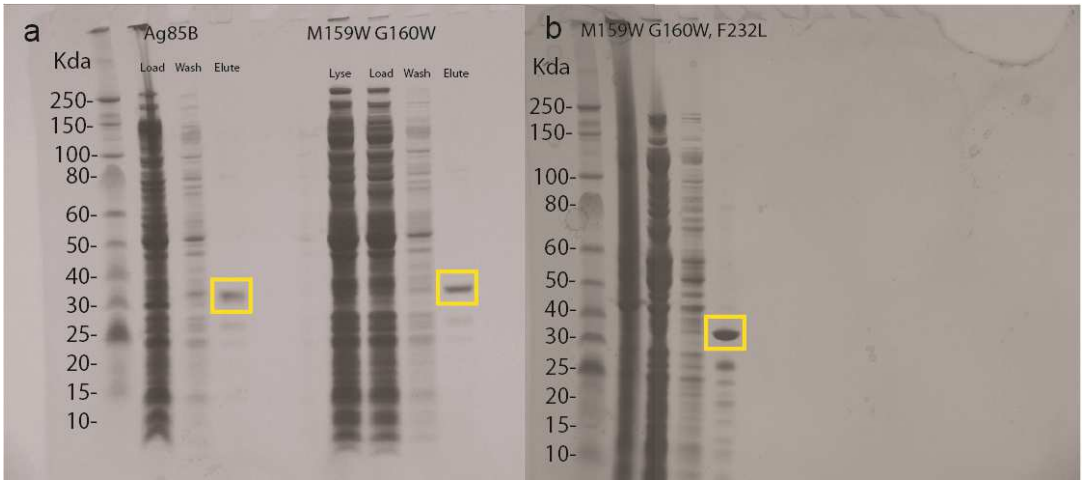


Figure 9.11. (a) Overexpression of Ag85B and Ag85B containing the M159W and G160W mutations (b) overexpression of Ag85B containing the M159W and G160W and F232L mutations. Boxes indicated purified protein. Ladder (New England BioLabs 10-250 Kda).

9.26.6 Native Page Gels

Dimeric nature of the Ag85 proteins was evaluated by native page gel (NativePAGE™ Novex® Bis-Tris 4-16% Gel Electrophoresis, Invitrogen) with NativeMark™ Unstained Protein Standard as markers for protein size, following the manufacturer's instructions. Samples were prepared in NativePAGE sample buffer (4x). 5-20µl sample was loaded on gel in wells prefilled with NativePAGE running buffer containing cathode buffer additive (cathode buffer). Inner chamber was filled with cathode buffer and outer chamber filled with NativePAGE running buffer 1x (anode buffer). Electrophoresis was performed for 1 hour at 150 volts and gel was destained in destaining buffer (40% MeOH, 10% acetic acid, water) with repeated microwave heating cycles and shaking.

9.27 General Cloning: PCR and ligation and transformation

Primers were designed in DNAStar lasergene SeqBuilder software and were obtained from Invitrogen. Genes were amplified by standard polymerase chain reaction (PCR), conducted per manufacturer's instructions for the following polymerases: High-fidelity Taq Polymerase (Invitrogen), Platinum Pfx Polymerase (Invitrogen) or Phusion High-Fidelity DNA polymerase (New England Biolabs). DMSO (3 %) was included in all reaction mixtures and the annealing temperature and elongation times were optimized for each reaction.

Cloning Vector

PCR products were ligated into cloning vectors with the use of either PCR Topo blunt, PCR 2.1 TA or PCR Topo 2.1 cloning kits (Invitrogen) following the manufacturers instructions. T4 ligase (Invitrogen) was used with the PCR 2.1 TA vector.

Transformations

Ligations were transformed into chemically competent or electrocompetent Ecoli cells. For cloning either Top10 (Invitrogen), XL-10 gold (stratagene), DH5α, DH10β were used. For protein overexpression: Bl21 DE3 cells with or without pLys were employed. Chemically competent cells were transformed by heat shock (45 sec, 42 °C) followed by 2 minutes on ice. Electrocompetent cells were electroporated in 1mm cuvette (Biorad) in Gene Pulser Xcell electroporation system (Biorad). The following settings were used for electroporation: Exponential decay of 25 µF charge, 200 Ω resistance, 1800 volts. Cells were recovered in SOC medium for 1 hour and plated on LB. X-Gal (30 µL 50 mg/ mL) was added to plates immediately before plating in order to facilitate blue/white selection on LB agar plates. Hygromycin 200 µg/mL, Kanamycin 40 µg/mL or Ampicillin 100 µg/mL were used for selection and colonies were cultured, plasmids purified by miniprep (Qiagen) and sent for sequencing (Macrogen).

9.27.1 Restriction Digests

All restriction enzymes were obtained from New England Biolabs and used according to the manufacturer's instructions. NdeI and HindIII digests and BamH1 and HindIII were conducted in buffer 2 (10x). SphI and EcorV digests or SmaI digests were conducted in buffer 4 (10x) with desired plasmid (10 µL), buffer (1.2 µL), enzymes (0.5 µL each). Plasmid was digested for 1.5 hours at 37℃, after which shrimp alkaline phosphatase (SAP) (1 µL) was added to target vector digest but not to insert digest reactions. Shrimp alkaline phosphatase (SAP) was allowed to react for 45 min, upon which time, enzymes were denatured (65℃ 20 min). Reactions were then analyzed by gel electrophoresis (1% agarose in TAE buffer containing Ethidium bromide) and desired DNA fragments were purified by gel extraction (Qiagen).

9.27.2 Ligation Reactions

Restriction digest products were ligated into appropriate vectors with the quick ligation kit (New England BioLabs) or T4 ligase kit (Invitrogen). Quick ligase was allowed to react at RT for 15 minutes. Standard ligase reactions were conducted at 11 °C overnight.

9.28 Cloning and expression of Rv0519C and Rv1288

Alignment of Genes

FASTA sequences for Rv1288 and Rv0519C were downloaded from the 'Tuberculist' http://genolist.pasteur.fr/TubercuList/ and were aligned with the sequences for Ag85A, B and C using Megalign with a Clustal W alignment.

Protein Modeling

Threading of Rv0519C and Rv1288 was conducted using LOMETS (Local MET-Threading-Server) and the resulting PDB files were obtained http://zhang.bioinformatics.ku.edu/LOMETS/output/S9065/ and http://zhang.bioinformatics.ku.edu/LOMETS/output/S9064/

Table 9.3. LOMETS prediction results for Rv0519C.

Prediction results by LOMETS consensus method

Rank	Template	Align_length	Coverage	Zscore	Seq_id	Confidence	Program	Target-template alignments	3-D models from threading alignments	Full-length models by MODELLER
1	3fcxa	237	0.79	25.649	0.156	High	SPARKS2	alignment_1	threading_1	model_1
2	3fcxa	237	0.79	24.052	0.156	High	SP3	alignment_2	threading_2	model_2
3	1pv1a	236	0.786	24.150	0.191	High	SPARKS2	alignment_3	threading_3	model_3
4	3fcxA	237	0.79	8.080	0.160	High	MUSTER	alignment_4	threading_4	model_4
5	1pv1a	236	0.786	22.736	0.186	High	SP3	alignment_5	threading_5	model_5
6	3c6b_A	230	0.766	44.525	0.183	High	HHsearch v1.5b	alignment_6	threading_6	model_6
7	1jjfa	237	0.79	21.675	0.165	High	SPARKS2	alignment_7	threading_7	model_7
8	3c6b_A	225	0.75	44.763	0.196	High	HHsearch v1.0	alignment_8	threading_8	model_8
9	1jjfa	237	0.79	16.524	0.194	High	PPA-I	alignment_9	threading_9	model_9
10	1sfr_A	226	0.753	25.620	0.168	High	HHsearch v1.5b	alignment_10	threading_10	model_10

Table 9.4. LOMETS prediction results for Rv1288.

Prediction results by LOMETS consensus method

Rank	Template	Align_length	Coverage	Zscore	Seq_id	Confidence	Program	Target-template alignments	3-D models from threading alignments	Full-length models by MODELLER
1	1sfrA	272	0.596	13.523	0.191	High	PPA-I	alignment_1	threading_1	model_1
2	1dqza	264	0.578	6.793	0.201	High	MUSTER	alignment_2	threading_2	model_2
3	1dqza	264	0.578	13.596	0.197	High	PPA-I	alignment_3	threading_3	model_3
4	1sfr_A	270	0.592	25.248	0.204	High	HHsearch v1.5a	alignment_4	threading_4	model_4
5	1dqz_A	266	0.583	25.286	0.203	High	HHsearch v1.5a	alignment_5	threading_5	model_5
6	1sfrA	272	0.596	6.716	0.199	High	MUSTER	alignment_6	threading_6	model_6
7	1dqza	266	0.583	21.805	0.199	High	SP3	alignment_7	threading_7	model_7
8	1r88A	259	0.567	13.654	0.205	High	PPA-I	alignment_8	threading_8	model_8
9	Esterase:1f0na	272	0.596	31.860	0.202	High	FUGUE	alignment_9	threading_9	model_9
10	1r88A	259	0.567	6.710	0.208	High	MUSTER	alignment_10	threading_10	model_10

9.28.1 Alignment of Rv1288 and Rv0519C with Ag85 proteins

Sequence alignment shows 20-30% sequence homology and conserved catalytic triad. ClustalW alignment in DNAStar LaserGene MegaAlign program.

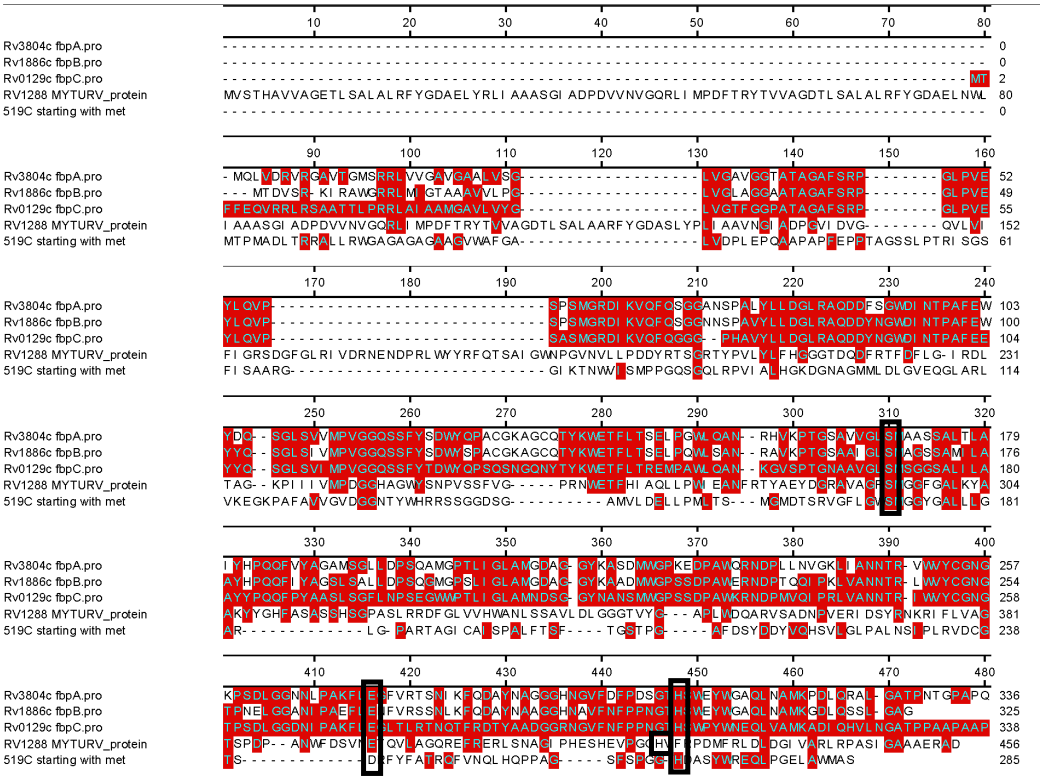


Figure 9.12. Alignment of Ag85 genes with Rv1288 and Rv0519C (generated by DNAstar lasergene MegAlign). Boxed residues indicate conserved catalytic triad.

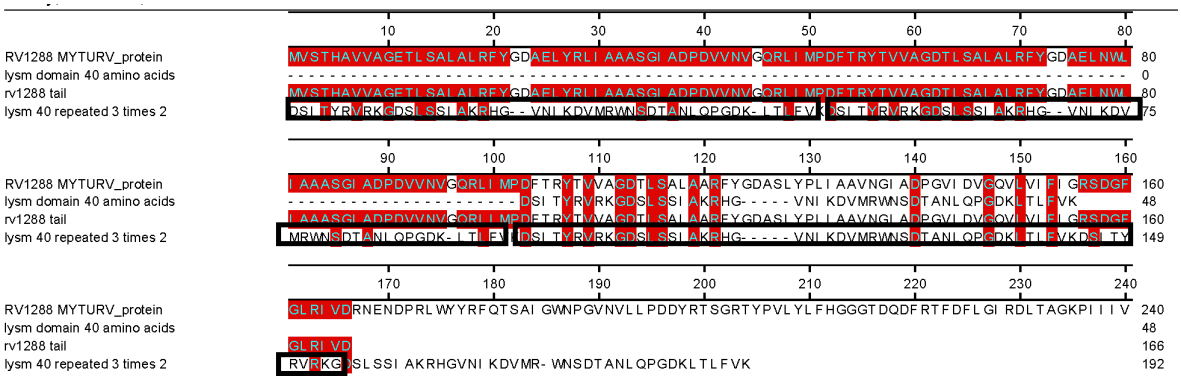


Figure 9.13. Alignment of Lysm domain with Rv1288 (generated by DNAstar lasergene MegAlign). Boxes indicate 3 repeated Lysm peptidoglycan binding domains.

9.28.2 Cloning of Rv0519C and Rv1288

The sequences for Rv0519C and Rv1288 were obtained in PUC vectors from Genescript, cloned with an N terminal BamHI site and C terminal HindIII site and truncated to remove the insoluble membrane binding domains. Genes were digested from parent PUC vectors with BamHI and HindIII restriction enzymes (**Figure 9.14a**) and ligated into pET23b vector. Vectors were transformed into Bl21 DE3 cells following the standard protocol and induced with IPTG. No protein induction was observed (**Figure 9.14b,c)** for Rv1288 and Rv0519C.

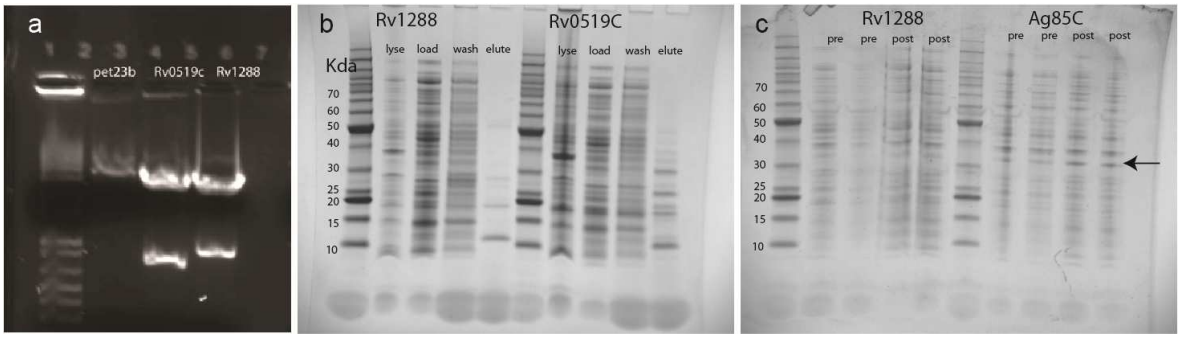


Figure 9.14. (a) Digest of Rv0519c, pET23b and Rv1288. 1Kb plus DNA ladder (Invitrogen) (b) Original attempted expression and purification of Rv1288 and Rv0519C. in pET23b shows no purified protein (c) expression profiling of Rv1288 and Ag85C as a control shows no induction of Rv1288. Arrow indicated Ag85C. Pre-pre induction post-post induction. Pagerule plus prestained protein ladder- Fermatas.

It was realized that the original cloning strategy had been faulty and a pET23a vector was required instead of pET23b. This is due to the fact that the PET vectors are designed to begin expressing in a specific frame after the T7 promoter. The frame is such that it allows the ATG present in the Nde restriction site CAT**ATG**, to be in frame. Therefore, in many cloning strategies, Nde restriction site is used as the upstream ligation site and either HindIII or Xho or BamHI are used downstream. The a, b, c notation for the individual pET vectors describes the reading frame relative to the BamH1 cloning site, GGATCC, which is downstream of the Nde site (**Figure 9.15**, pET23 map). Vectors denoted 'a' begin expressing in frame of the GGA triplet of the BamH1. Vectors denoted 'b' express from GAT because they have one base subtracted before the BanH1 site. Those denoted 'c' express from the ATC triplet.

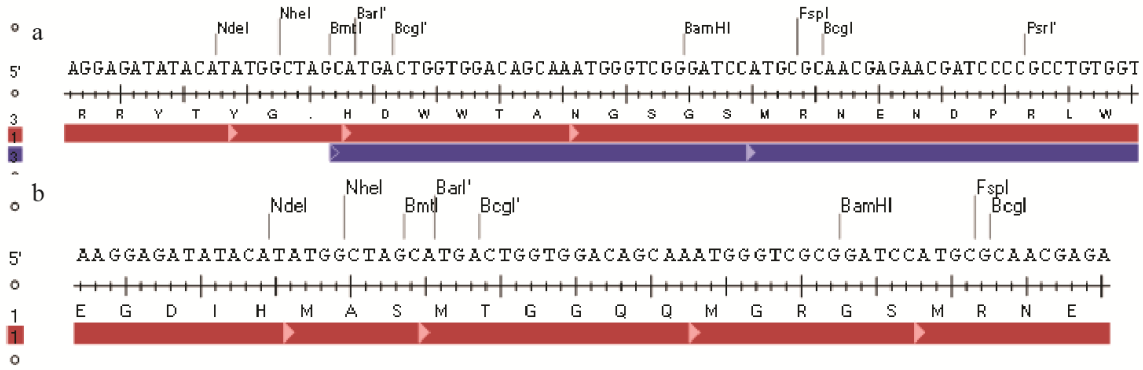


Figure 9.15. (a) Rv1288 in pET23b vector. Notice that 'in frame' expression begins at Nde ATG (Red ORF). However, the Rv1288 gene is in a different frame (blue ORF). (b) Rv1288 in pET23a vector. BamH1 cloning site is in frame with Nde. When this gene expresses, it will have some 'junk' amino acids upstream of the actual start site, due to the DNA between Nde and BamH1.

In the original cloning strategy the BamH1 site was utilized as the upstream ligation site, but cloned into a pET23b vector, thereby putting the entire gene out of frame (**Figure 9.15**). To remedy this pET23a was obtained and both Rv1288 and Rv0519c were ligated into this vector, transformed into BL21 DE3 cells and then cultured in LB medium and induced with IPTG. Robust induction for Rv1288 and Rv0519C was observed in the pET23a overexpression system. Purification over nickel NTA resin following standard purification protocol afforded pure Rv0519C-hexahis, but no Rv1288-hexahis. Expression of Rv1288 in Ecoli BL21 strain expressing additional chaperones afforded no soluble protein (**Figure 9.16**). Pellet of *E. coli* overexpressing Rv1288 were very white and it was determined that all protein was found to be in inclusion bodies.

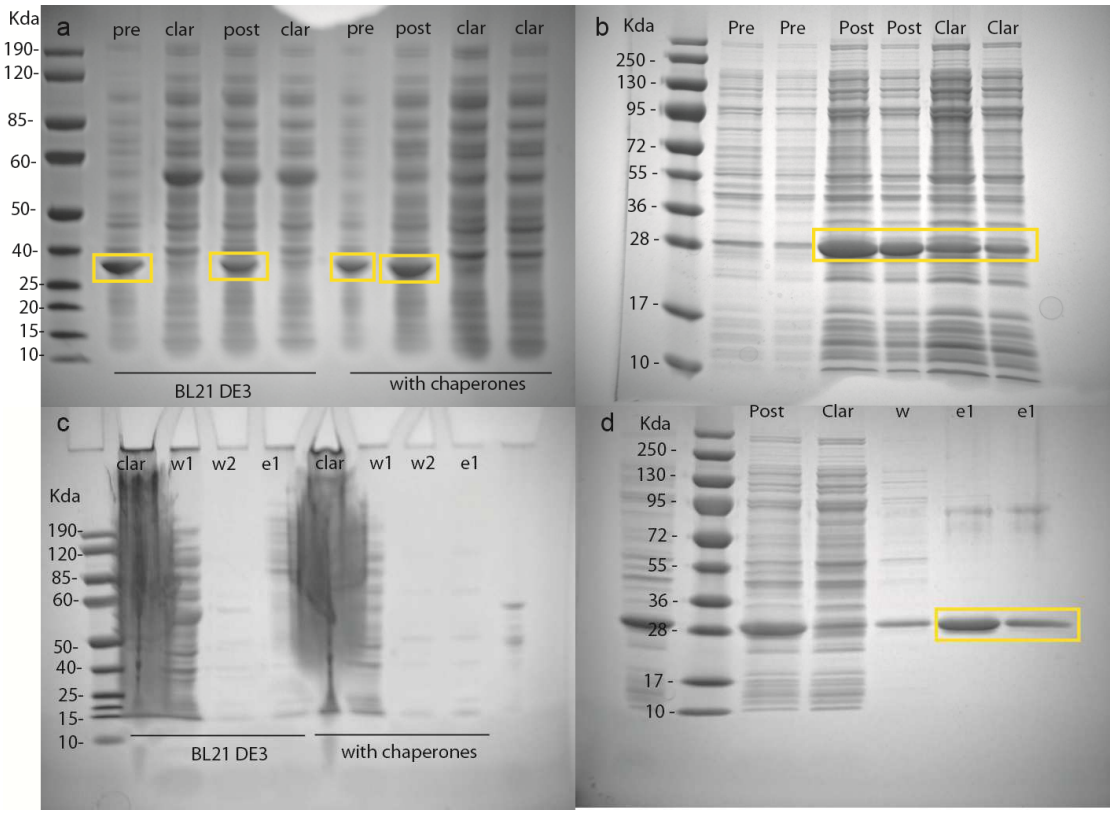


Figure 9.16. Induction of Rv1288 and Rv0519C constructs in pET23a shows robust overexpression for (a) Rv1288 and (b) Rv0519C. (c) attempted purification of Rv1288 affords no soluble protein (d) Purification of Rv0519C yields pure, His-tagged protein. Pre- pre induction, post- post induction, clar- clarified lysate. W- wash, e- elution.

We successfully obtained Rv1288 as a pET23a construct, through standard digest and ligation into pET23a vector (**Figure 9.16**). Under expression profiling, the gene expressed robustly. However, upon purification, we found that the protein was entirely insoluble, found in inclusion bodies. The *E. coli* pellet was extremely white, characteristic of inclusion bodies, and none of our desired protein was found in the clarified lysate. Attempts at expression both in standard BL21 DE3 cells and cells expressing chaperones both resulted in insoluble protein. Likewise, purification over nickel resin yielded no soluble protein.

The original cloning design, involving BamHI and HindIII that had caused the frame shift problem allowed for cloning into a vector (pMAL-C2X) encoding a C-terminal maltose binding protein tag (MBP). Rv1288, Rv0519C and pMAL-C2X were digested with BamHI and HindIII and were ligated the desired genes into pMAL-C2X to generate the MBP fusion proteins. Ligations proceeded smoothly and the desired constructs were obtained.

Genes exhibited no leaky pre-induction expression and upon induction with IPTG, robust overexpression was observed (**Figure 9.17**). Purification on amylase resin yielded the desired Rv0519C-MBP fusion protein in very high yield. Rv1288-MBP fusion appears to be partially insoluble so the production of the Rv1288 fusion was substantially lower and may be insoluble

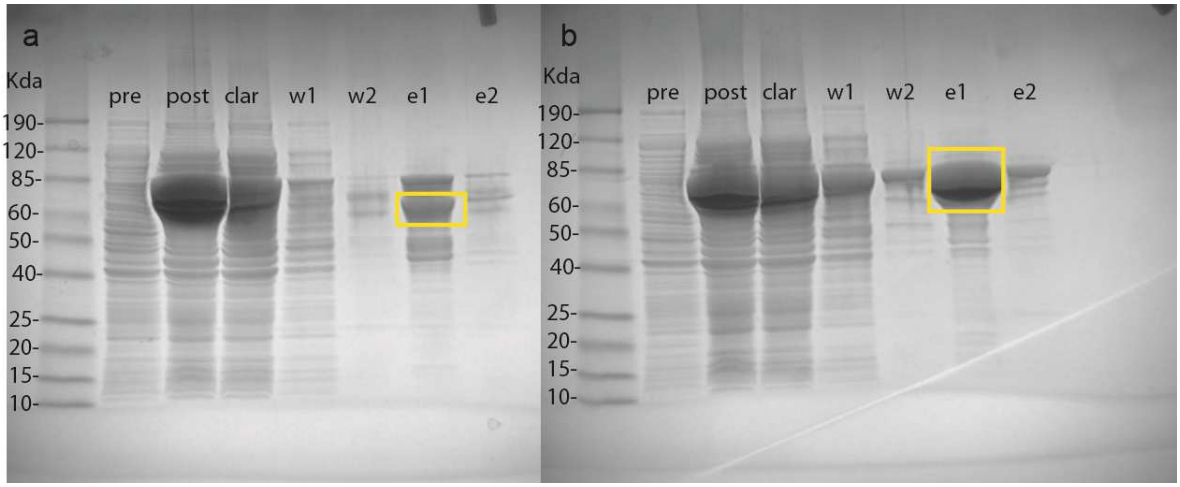


Figure 9.17. (a) expression and purification of Rv1288-MBP fusion protein over expressed in pMAL-C2X vector (b) expression and purification of Rv0519c-MBP fusion protein over expressed in pMAL-C2X vector. Pre- pre induction, post- post induction, clar- clarified lysate. W- wash, e- elution.

9.29 Microarray analysis

Microarray data was obtained from the supplementary information of Boshoff et al[10]. Additional array data was kindly shared to the author by Dr. Kriti Arora and Dr. Helena Boshoff.

M. tuberculosis H37Rv was grown to mid-logarithmic phase (OD_{650}=0.357). Treatment with isoniazid was carried out at 2µM (10-times minimum inhibitory concentration) or 0 µM (untreated) final concentration for 6 h following which the cells were harvested by spinning at 3,000 rpm for 5 min at 4 °C. The pellets were resuspended in 1ml Trizol reagent (Invitrogen) and bead beat with zirconia beads (BioSpec) (1/3rd volume of beads to sample) for 3 cycles of 30 sec each at 6,000 rpm with intermittent incubation on ice for 5 mins each. The samples were spun down at 13,000 rpm for 5mins at 4°C to remove lysed cell debris and the supernatant was then removed from the BSL-3. (Note: All work from this step on was conducted using surfaces and pipettes cleaned with RNA-Zap). One-tenth volume of bromo-chloro-propane (BCP; Molecular Research Center, Inc.) was added to the samples and the samples were inverted to mix contents. Phase separation was allowed to occur for 10mins on ice followed by centrifugation at 10,000 rpm for 10 min at 4 °C. The upper aqueous layer was removed and to it 1:1 volume of 70% ethanol was added. The samples were inverted to mix and immediately loaded on to RNeasy spin columns provided with the RNeasy Mini Kit (Qiagen) and centrifuged at 10,000 rpm for 1 min. The flow through was discarded and genomic DNA was

removed from the samples by an on-column DNA digestion using the RNase-free DNase set from (Qiagen). The RNA was further cleaned using RNeasy reagents provided in the kit according to the manufacturer's protocol. The concentration and purity of extracted RNA was determined using the BioAnalyzer (Agilent technologies). Samples with a RNA-integrity number (RIN) greater than 9.0 were sent to the NIAID Microarray Facility (Research Technologies Branch, National Institutes of Health) for cDNA labeling and hybridization. The raw mRNA expression data obtained were $\log_2$ transformed and reported as fold-change relative to untreated samples. Samples showing 2-fold or greater difference in expression (<-1.00 or >+1.00 in the $\log_2$-transformed data set) were considered significant.

9.30 Cloning of Ag85 into pVV16 His-tag vector

Vector pVV16 was obtained from the Colorado State TB Vaccine and Research Materials (TBVTRM) contract as a mycobacterial, replicating overexpression vector that encodes Kanamycin resistance and His-tag. Target genes were cloned at the NdeI and HindIII sites, upstream and in frame with the hexaHis-tag. Ag85C without signal peptide was amplified from the pET23b vector encoding Ag85C, also obtained from Colorado State. All other genes were amplified from genomic DNA. Taq polymerase (invitrogen) was found to be superior to Phusion polymerase (NEB). Genes were cloned into either PCR 2.1 TA or PCR Topo cloning vectors, following standard cloning protocols. Gene identity was confirmed by sequencing (Macrogen) and genes were then digested from cloning vector with NdeI and HindIII restrictions enzymes and ligated in pVV16 vector. BamHI restriction site was employed for Ag85A, which contained an internal HindIII site. pVV16 overexpression constructs were then electroporated into electrocompetent *M. smegmatis* (see the follow sections for electroporation and generation of competent cell protocols).

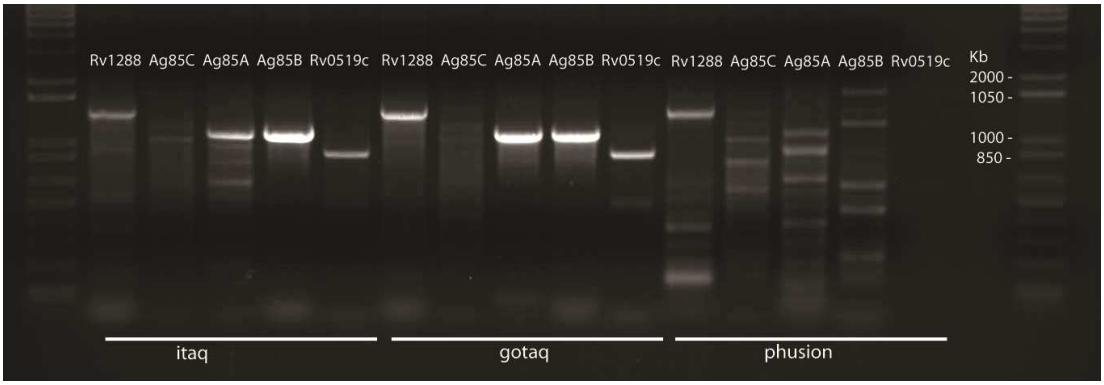


Figure 9.18. Comparison of different polymerases for PCR of desired genes, including signal peptides reveals that Taq efficiently amplifies most gene products with the appropriate restriction sites for cloning into pVV16 vector.

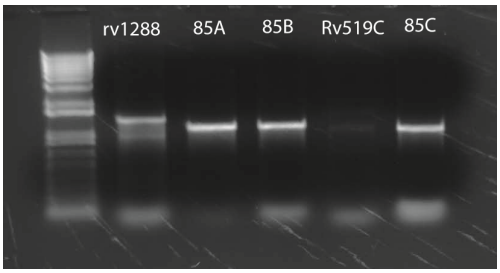


Figure 9.19. PCR of genes using high fidelity Platinum Taq®. High fidelity, Platinum Taq® was employed to reduce the number of errors during amplification. PCR with a hot-start proceeded smoothly for all genes, although Rv0519C amplified poorly.

Table 9.5. Table of primers employed to amplify genes for cloning into vector pVV16 for overexpression of His-tagged constructs.

Primer	Sequence 5' to 3'	Gene and restriction site amplified
kb-124	GGGCATATGCGCCCAAACATGCATGGATG	85A Nde forward
kb-108	GGGGGATCCTTGGCGCCCTGGG	85A BamH1 reverse
kb-74	GGGCATATGAGACGACTTTGCGCC	85B NdeI forward
kb-75	GGGAAGCTTGCCGGCGCCTAACG	85B HindIII end
kb-78	GGGCATATGACGTTCTTCGAACAGGTG	85C NdeI forward
kb-79	GGGAAGCTTGGCGGCCGGCGCAG	85C HindIII reverse
Kb-60	GGGCATATGTTCTCTAGGCCCG	85C NdeI no signal peptide forward
kb-61	GGGAAGCTTCTCGAGAGCGGC	85C HindIII no signal peptide reverse
kb-87	GGGCATATGACGCCCATGGCAGACCT	519C NdeI forward
kb-88	GGGAAGCTTGGACGCCATCCAAGCCA	519C HINDIII reverse
kb-62	GGGCATATGGTCAGCACACATG	1288 NdeI forward
kb-63	GGGAAGCTTATCGGCGCGTTC	1288 HIndIII reverse
kb-81	GGGCATATGGTGGACCGCAACGA	1288 soluble domain NdeI forward

9.31 Western blots

M smegmatis (50 mL in 7H9 medium), expressing pVV16 encoding target overexpression proteins was harvested by centrifugation (3000 rpm, 10 minutes). Supernatant was collected and concentrated in 10,000 Kda MW cutoff spin concentrators (Amicon ultracel 10,000k). Pellet was resuspended (2 mL PBS pH 7.2, NaCl 250 mM, Triton x100 0.05%, β-mercapto ethanol) and lysed by sonication (setting 50, 3 cycles of 1 min, 5 seconds on and 5 seconds off), using the microprobe. Lysate was centrifuged (13,000 rpm, 10 minutes) and clarified lysate was removed to a new tube. Lysate, supernatant and clarified lysate (1 µL each) were run by nupage Bis tris gel 4-12%.

Transfer buffer was made: A 10 X buffer consisted of 30.3 g tris base, 144.1 g glycine, 1 L ultrapure water, pH to 8.3. For 1X buffer, 100 mL of the 10X buffer was diluted with 200 mL methanol and

700 mL water. PVDF membrane was washed in methanol and equilibrated in 1X transfer buffer. Proteins were transferred to PVDF membrane, using standard transfer protocol in an XCell II Blot module system. Membrane and filter paper were cut to the same dimensions as the gel and a transfer stack was made (sponge, filter paper, gel, membrane, filterpaper, sponge) with care to prevent bubbles between the layers and keep the entire stack wet with transfer buffer. Stack was placed in the Transfer chamber, which was placed in the Blot module and filled with transfer buffer. The outerchamber was filled with ice water and leads were connected and transfer was performed (20 min at 30 volts). Membrane was then placed in a clean tray and blocked with 10 mL TBST (25 mM tris, 150 mM NaCl, pH 7.6, 0.05% Tween 20) with BSA (10 mg/mL) overnight at 4℃. Membrane was washed with TBST (2 X 15 mL, 10 min each) and incubated with HisProbe-HRP working solution for one hour (dilute HisProbe-HRP (Pierce 15165) 1:5000 in TBST). Blot was then washed with TBST (4 X 15 mL, 10 min each) and incubated with 10 mL of Supersignal West Pico substrate working solution for 5 minutes. Membrane was placed in membrane protector and exposed to Kodak film (Biomax MR) in a film cassette and then developed.

9.32 Mycolyl transferase genes into tetracycline inducible GFP fusions

The TET expression system was obtained as a kind gift from Dr. Sabine Ehrt, consisting of one plasmid, which contains no elements that mediate replication in mycobacteria but does contain both TetR and the tetO-containing promoter (pTC-mcs1), and two plasmids which mediate chromosomal integration (pTC-2M1L) or replication(pTC-2M1L) in mycobacteria. The system is designed so that the gene is first cloned into pTC-mcs1 and then the entire TetR-Pmyc1tetO expression cassette can be cloned into the myco/Ec shuttle vectors using the rare cutters PacI/NotI.

pTC-mcs1: integrative plasmid for mycobacteria, KanR, mcs with unique PacI and NotI.
pTE-mcs1: replicative plasmid for mycobacteria, HygR, mcs with unique PacI and NotI.
pTC-2M1L: contains PacI/NotI cassette from pTC-2M1L in pTE-mcs1.
pTE-2M1L: contains PacI/NotI cassette from pTX-2M1L in pTE-mcs1.

To reduce cloning steps, the cloning process could be done directly into the replicating mycobacterial vector pTE-2M1L as detailed in Chapter 6. GFP and RFP were cloned between the EcoRV and NotI sites and the Ag85 genes were cloned between the SpHI and EcoRV sites. SmaI was used as an alternate blunt cutter for Ag85B and C, as both contained internal EcoRV sites. Genes were amplified by PCR with High fidelity Taq polymerase and with primers encoding the ribosome binding site to replace the RBS, as it is removed during restriction digest with SphI.

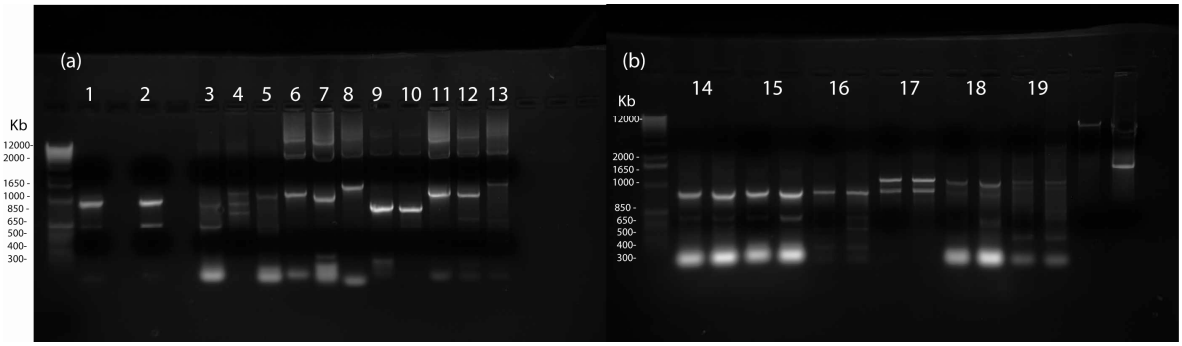


Figure 9.20. PCR of genes of interest with appropriate restriction sites indicated for cloning into pVV16 (NdeI and HindIII or BamHI), or pTE-2M1L (SphI, EcoRV, SmaI and NotI). Lanes: 1-Rv0519c NdeI and HindIII , 2- Rv0519c NdeI and BamHI, 3- Rv0519c SphI and EcoRV, 4- Ag85A NdeI and BamHI, 5- Ag85A SphI and EcoRV, 6-Ag85B SphI and SmaI, 7- Ag85C SphI and SmaI, 8- Rv1288 SphI and EcoRV, 9- GFP EcoRV and NotI, 10-RFP EcoRV and NotI, 11- Ag85B NdeI and HindIII, 12-Ag85C NdeI and HindIII, 13- Rv1288 NdeI and HindIII, 14- Rv0519C SphI and EcoRV, 15- Rv0519C SphI and EcoRV 16- Ag85A NdeI and BamHI, 17- Ag85A NdeI and BamHI, 18- Ag85A SphI and EcoRV, 19- Ag85A SphI and EcoRV. Different combinations of primers were employed to optimize amplification.

Table 9.6 Table of primers employed to amplify genes for cloning into tet inducible vector pTE-2M1L

Primer	Sequence 5' to 3'	Gene and restriction site amplified
kb-109	GGGGATATCATGGCCAAGACAATTGCGGA	GFP EcoRV forward
kb-111	GGGGCGGCCGCCTTGTACAGCTCGTC	GFP Not1 reverse
kb-113	GGGGATATCATGGCCAAGACAATTGCGGATC	RFP EcoRV forward
kb-116	GGGGGCGGCCGCTTAGGCGCCGGTG	RFP Not1 longer
kb-137	GGGGGCATGCGAAGGAGGTATACATATGCGCCCAAACATGCATG	85A SphI forward with RBS
kb-101	GGGGATATCGGCGCCCTGGGGC	85A EcoRV reverse
kb-141	GGGGGCATGCGAAGGAGGTATACATATGAGACGACTTTGCGCCCGA	85B SphI forward with RBS
kb-142	GGGCCCGGGGCCGGCGCCTAACGA	85B SmaI reverse
kb-136	GGGGGCATGCGAAGGAGGTATACATATGACGTTCTTCGAACAGGTGC	85C SphI forward with RBS
kb-106	GGGCCCGGGGGCGGCCGGCGCAG	85C SmaI reverse longer
kb-139	GGGGGCATGCGAAGGAGGTATACATATGACGCCCATGGCAGACCTG	519C SphI forward with RBS
kb-98	GGGGATATCACGCCCATGGCAGACCT	519c EcoRV reverse
kb-140	GGGGGCATGCGAAGGAGGTATACATATGGTCAGCACACATGCGG	1288 SphI forward with RBS
kb-96	GGGGATATCATCGGCGCGTTCTGC	1288 EcoRV reverse

Table 9.7 Table of primers employed to mutate catalytic serine residues to alanine (SA) in Ag85A,B and C. Red indicates site of mutation.

Primer	Sequence 5' to 3'	Mutant amplified
kb-193	GCGCCGTCGTCGGTCTT**GCG**ATGGCTGCTTCTTCGG	85A ser to ala
kb-194	CCGAAGAAGCAGCCAT**CGC**AAGACCGACGACGGCGC	85A ser to ala
kb-195	GCCGACGAGCCGGCCAT**CGC**CAAGCCGATTGCAGCGCTGC	85B ser to ala
kb-196	GCAGCGCTGCAATCGGCTTGGCGATGGCCGGCTCGTCGGC	85B ser to ala
kb-197	GCGCGGAACCGCCCGACAT**CGC**AAGACCCACCGCCGCGTTG	85C ser to ala
kb-198	CAACGCGGCGGTGGGTCTT**GCG**ATGTCGGGCGGTTCCGCGC	85C ser to ala

9.32.1 Preparation of electrocompetent *M. smegmatis* and *M. tuberculosis*

The protocol followed is from TBVTRM website http://www.cvmbs.colostate.edu/mip/tb/sop.htm

A fresh 50 ml culture of *M. smegmatis* MC^2155 or H37RV was grown in Middlebrook 7H9-ADC-Tween or to mid-log phase (OD^{600} = 0.5 to 1.0). *M. smegmatis* cells were incubated on ice for one hour and kept on ice for the entire protocol. *Mtb* cells were kept at RT for the entire process. Cells were then transferred to sterile 50 mL falcon tubes and harvested by centrifugation at 3000 rpm 4 °C (*M. smegmatis) or 25* °C (*Mtb)* for 10 minutes. Supernatant was decanted into Vesphene disinfectant. Pellet was resuspended in 30 mL sterile 10% glycerol (cold for *M. smeg)*. Centrifugation and resuspension steps were repeated twice more with a gradual reduction of the volume of 10% glycerol added (5 mL then 2 mL respectively). 400µl aliquots were distributed into 1.7 ml eppendof tubes and store at -70°C.

9.32.2 Electroporation into *M. smegmatis*

M. smegmatis was electroporated following the standard protocol[11]. 1 µL of plasmid in DEPC water was added to 100 µL of electrocompent mc^2155 *M. smegmatis* on ice. Cells were allowed to incubate with plasmid on ice for 5 minutes and were then transferred to 2mm electroporation cuvette (Biorad) with care to make sure that there were no bubbles present in the chamber. A control with no plasmid was treated under identical conditions. The following settings were used for electroporation: 25 µF, 25 KV, 200 Ω, 2 mm cuvette. Immediately after electroporation 1 mL 7H9 medium was added and cells were allowed to recover at 37 ℃ with shaking for 5 hours, upon which 100 µL aliquots were plated on 7H11 plates with Hyg^{50}. The plates were incubated at 37 ℃ for 48 hours at which point colonies were picked and cultured in 7H9 medium under Hyg selection.

9.32.3 Electroporation into Mtb

M. tuberculosis was electroporated following the standard protocol[11]. 4 µL of plasmid in DEPC water was added to 400 µL of electrocompent H37Rv *M. tuberculosis* at room temperature. Cells were allowed to incubate with plasmid for 5 minutes and were then transferred to 2mm electroporation cuvette (biorad) with care to make sure that there were no bubbles present in the chamber. A control with no plasmid was treated under identical conditions. The following settings were used for electroporation: 25 µF, 25 KV, 200 Ω, 2 mm cuvette. Immediately after electroporation 1 mL 7H9 medium was added and cells were allowed to recover at 37 ℃ with shaking for 24 hours, upon which 500 µL aliquots were plated on 7H11 plates with Hyg^{50}. Plates were incubated at 37 ℃ for 3 weeks at which point colonies were picked and cultured in 7H9 medium under Hyg selection.

9.32.4 Induction of strains overexpressing Tet inducible GFP fusion proteins

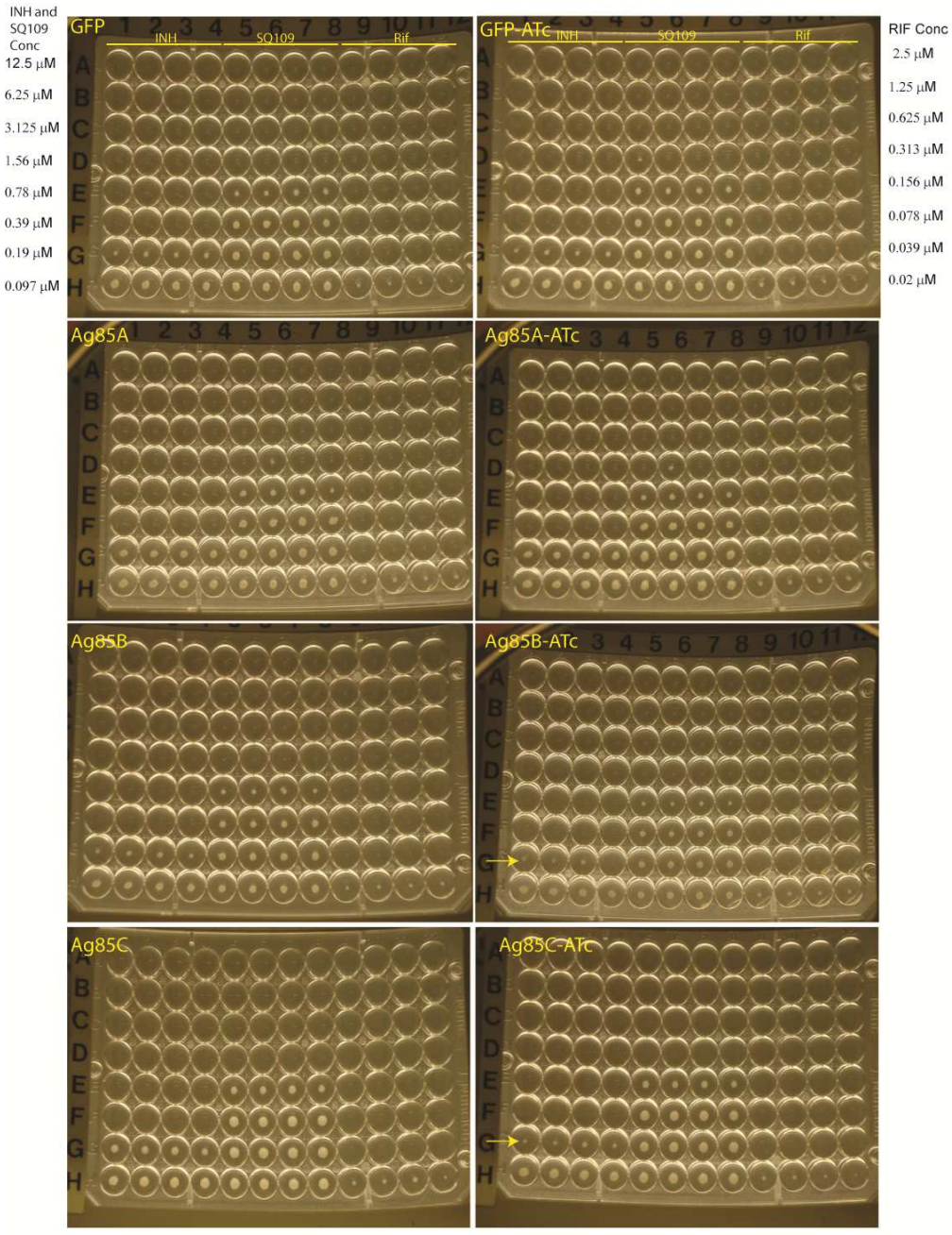
Cells were grown in 7H9 or GAS(T) medium (with or without Tween). Tween free medium was used for miscroscopy to encourage Ag85-GFP adherence to the mycobacteria. Overexpressing strains were grown in the presence of 100 ng/mL anhydrotetracyline (ATc) and 50 ng/mL hygromycin. For fluoresence studies, whole cell culture was evaluated for fluorescence in black plates (nunc) on a fluostar Optima fluorimeter. Supernatant fluorescence was measured by first harvesting the cells by centrifugation and then removal of the supernatant and followed by quantitation of fluorescence. For microscopy cells were grown in GAST medium, fixed with formalin (4%) and washed into PBS buffer. 50 µL PBS sample was applied to a round coverslip, allowed to dry and 1 drop mounting medium (Prolong Gold, Invitrogen) was placed on each slide. With forceps, coverslips were flipped over onto the droplet of mounting medium and allowed to dry prior to visualization by confocal microscopy. Macrophage infections were conducted as has been described earlier with the addition of 100 ng/mL anhydrotetracyline to the macrophage DMEM culture medium. Macrophages infected with tet inducible strains were allowed to grow for 72 hours under ATc induction prior to fixation and visualization by confocal microscopy.

9.32.5 Minimum Inhibitory Concentration (MIC) with Tet inducible strains

MIC analysis for drug sensitivity of tet inducible strains was conducted in the standard MIC format described above.

For *M. smegmatis* grown in the presence or absence of 200 ng/mL anhydrotetracyline the following initial drug concentrations were employed. INH was used at in 1mM. SQ109 was used at 62.5 µM. Rifampicin at 125 µM. MIC plates were read after 36, 48 and 72 h.

For *M. tuberculosis* grown in the presence or absence of 200 ng/mL anhydrotetracyline the following initial drug concentrations were employed. INH was used at in 12.5 µM. SQ109 was used at 12.5 µM. Rifampicin at 1.25 µM. MIC plates were read after 7 and 10 days.



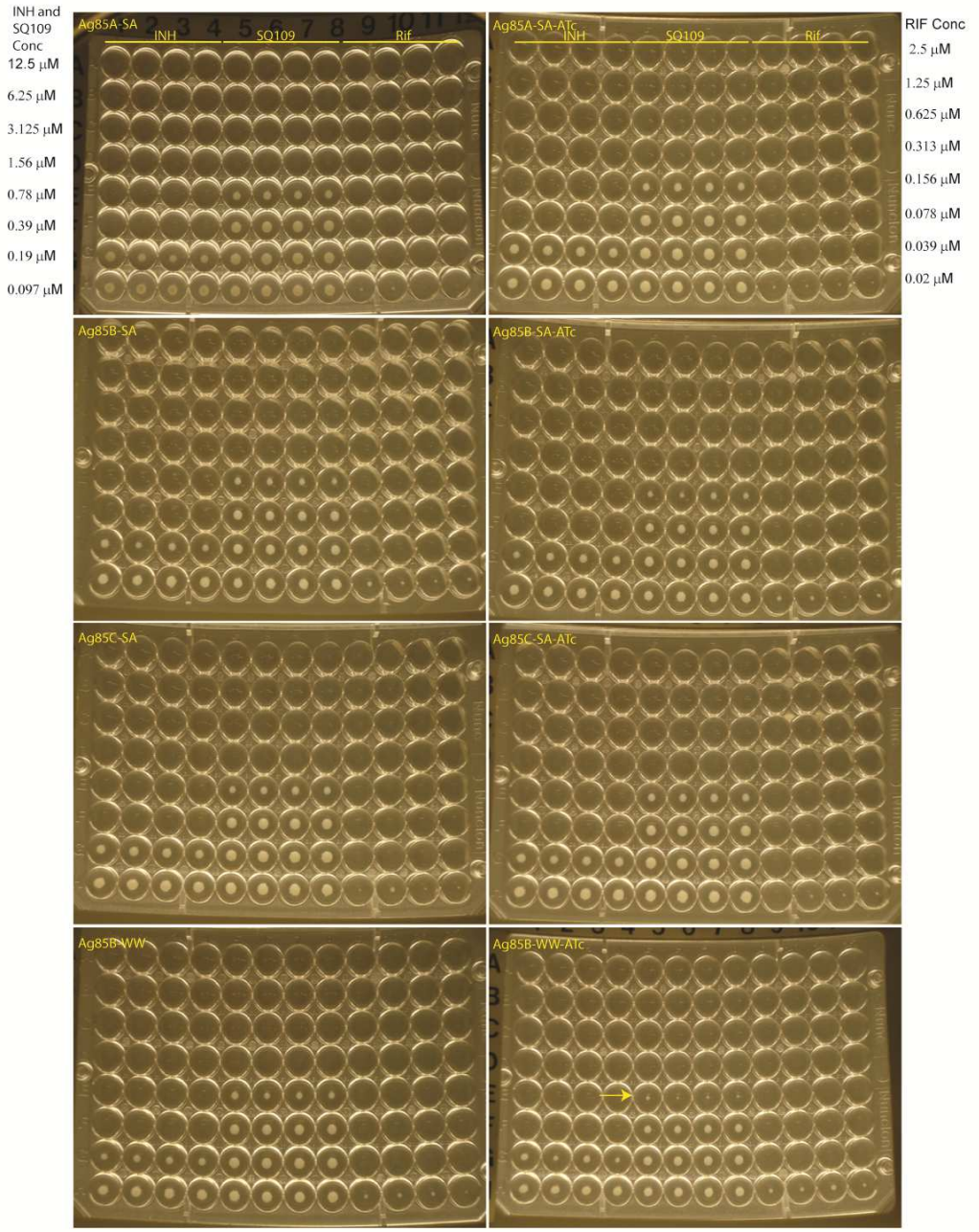


Figure 9.21. Minimum inhibitory concentrations for Isoniazid (INH), SQ109 and Rifampicin (RIF) in M. tuberculosis H37Rv, expressing tet inducible Ag85 fusion proteins. Induction of Ag85B and C protein expression with Anhydrotetracycline (ATc), may cause growth inhibition under INH treatment. Ag85C overexpression may also offer a modest growth inhibition under SQ109 treatment. Arrows indicate wells of interest. SA- catalytic serine-alanine mutant, WW- Ag85B encoding the TrpTrp motif of the Ag85C secondary site. Read after 10 days.

9.32.6 Lipid extractions of ATc inducible constructs

9.32.6.1 TDM and TMM extraction

Colabeling

Cells were grown to an OD_{650} of 0.1 and labeled with ^{14}C-acetate (1 uCi/mL culture) in the presence of anhydrotetracycline (200 ng/mL) for 4 hours (*M. smegmatis*) or 3 days (*Mtb)* and then harvested by centrifugation (3000 rpm, 5 min) and extracted into 2 mL (2:1 chloroform:methanol). The organic layer was removed, concentrated and the residue was resuspended in 200 µL (80:20 chloroform:methanol) and 20 µL was spotted onto silica TLC plates, which were developed for 1 hour (75:25:4 or 80:20:2 chloroform:methanol:water). Plates were scanned for fluorescence and exposed to a phosphor storage plate, which was scanned for radioactivity.

Prelabeling

Cells were grown to an OD_{650} of 0.1 and labeled with ^{14}C-acetate (1 uCi/mL culture) for 3 hours (*M. smegmatis*) or 3 days (*Mtb).* Cells were then harvested by centrifugation (3000 rpm, 5 min) and resuspended in fresh 7H9 medium containing anhydroxy tetracycline (200 ng/mL) and were grown for a further 3 hours (*M. smegmatis)*or 2 days (*Mtb).* Extraction and TLC analysis of lipids proceeded in an identical fashion to colabel.

9.32.6.2 Mycolate analysis

Mycolate analysis were conducted based on previous reports[5]. After TDM and TMM lipid extraction, the remaining aqueous cell fraction was saponified to remove the cell associated mycolates. 0.5 mL 20% TBAH was added to acqueous layer and heated to 100 °C for 3-4 hours. To this was added dichloromethane (0.5 mL) and methyl iodide (50 µL) and reaction was shaken vigorously at room temperature for 1-2 hours. Organic layer was removed to a new vial and concentrated under a stream of argon. The resulting white crystals were extracted with diethyl ether (2 × 250 µL) and the extract was placed in a fresh vial and concentrated under a stream of argon. The resulting clear oil was re-dissolved in 200 µL (2:1 toluene:acetonitrile) and 20 µL product was spotted on silica TLC plate and developed with three successive runs (3 × petroleum ether or hexanes:diethyl ether 85:15). Plate was then exposed to storage phosphor plate, which was scanned for radioactivity. Radioactivity was quantified in phosphorimager units and normalized to the FAMES.

9.33 Summary of all plasmids

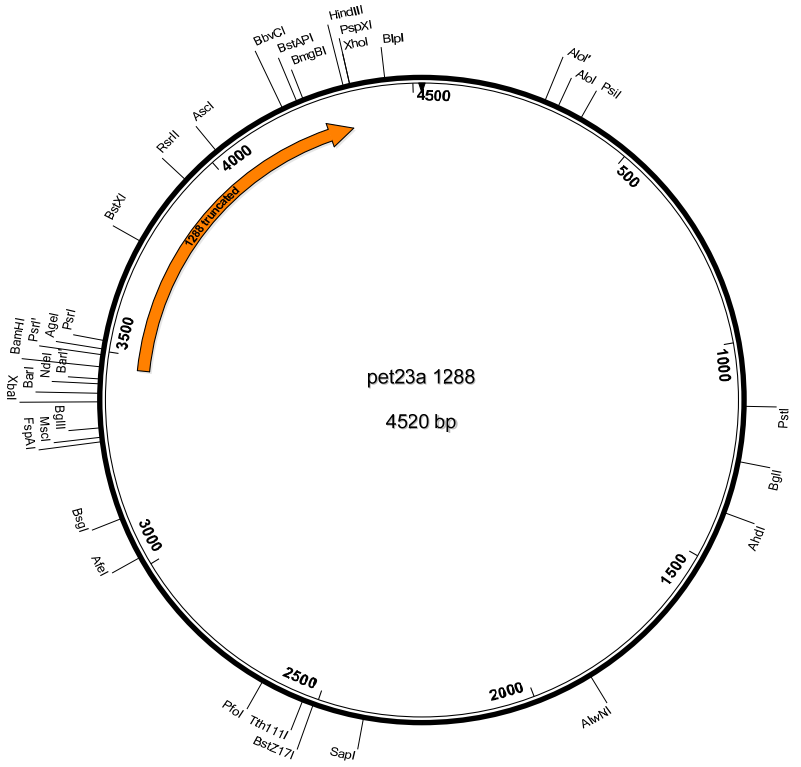
Table 9.8. Summary of all plasmids referenced in this thesis

Plasmid	Alternate name	Source and description
PCR 2.1 TA		Invitrogen, cloning vector, TA overhang
PCR Blunt TOPO		Invitrogen, cloning vector, TOPO activated, blunt
PCR TOPO TA		Invitrogen, cloning vector, TOPO, TA overhang
pVV16		TBVTRM, Colorado stateHexaHis TB overexpression
pTE-2M1L		Dr Sabine Ehrt, Tet inducible replicating
pTC-2M1L		Dr Sabine Ehrt, Tet inducible integrating
pMAL-C2X		Dr. Ujjin Manjunatha, Originally NEB, MBP overexpression
pET23b		EMD chemicals, Hexahis overexpression
pET23a		EMD chemicals, Hexahis overexpression
pMRLB41	Rv3804c, Ag85A	Colorado state TBVTRM, Ag85A overexpressing in pET23b
pMRLB47	Rv1886c, Ag85B	Colorado state TBVTRM, Ag85B overexpressing in pET23b
pMRLB16	Rv0129c, Ag85C	Colorado state TBVTRM, Ag85C overexpressing in pET23b
Rv0519C	Rv0519C PUC57	Gene synthesized by Genscript with BamHI and HindIII sites
Rv1288	Rv1288 PUC57	Gene synthesized by Genscript with BamHI and HindIII sites
Plasmids generated by KMB		
pKMB1	Ag85C W157A	Colorado state TBVTRM, Ag85C mutant overexpressing in pET23b
pKMB2	Ag85C W158A	Colorado state TBVTRM, Ag85C mutant overexpressing in pET23b
pKMB3	Ag85C L230F	Colorado state TBVTRM, Ag85C mutant overexpressing in pET23b
pKMB4	Ag85C W157A W158A (AA)	Colorado state TBVTRM, Ag85C mutant overexpressing in pET23b
pKMB5	Ag85C W157M W158G (MG)	Colorado state TBVTRM, Ag85C mutant overexpressing in pET23b
pKMB6	Ag85C W157A, W158A, L230A (AAA)	Colorado state TBVTRM, Ag85C mutant overexpressing in pET23b
pKMB7	Ag85C W157A, W158A, L230F (AAF)	Colorado state TBVTRM, Ag85C mutant overexpressing in pET23b
pKMB8	Ag85C W157M, W158G, L239F (MGF)	Colorado state TBVTRM, Ag85C mutant overexpressing in pET23b
pKMB9	Ag85B M159W G160W (WW)	Colorado state TBVTRM, Ag85B mutant overexpressing in pET23b
pKMB10	Ag85B M159W G160W F232L (WWL)	Colorado state TBVTRM, Ag85B mutant overexpressing in pET23b
pKMB11	Rv0519C pET23b	subcloned into pET23b vector (out of frame)
pKMB12	Rv1288 pET23b	subcloned into pET23b vector (out of frame)
pKMB13	Rv0519C pET23	subcloned into pET23a vector for overexpression
pKMB14	Rv1288 pET23	subcloned into pET23a vector for overexpression
pKMB15	Rv0519C Pmal	subcloned into pMAL-C2X as MBP fusion protein
pKMB16	Rv1288 Pmal	subcloned into pMAL-C2X as MBP fusion protein
pKMB17	Ag85A NdeI and BamHI	cloned PCR 2.1 with signal peptide and NdeI and BamHI
pKMB18	Ag85B NdeI and HindIII	cloned PCR 2.1 with signal peptide and NdeI and HindIII
pKMB19	Ag85C NdeI and HindIII	cloned PCR 2.1 with signal peptide and NdeI and HindIII
pKMB20	Ag85C short NdeI and HindIII	cloned PCR 2.1 with signal peptide and NdeI and HindIII
pKMB21	Rv0519C NdeI and HindIII	cloned PCR 2.1 with signal peptide and NdeI and HindIII
pKMB22	Rv1288 NdeI and HindIII	cloned PCR 2.1 with signal peptide and NdeI and HindIII
pKMB23	Rv1288 short NdeI and HindIII	cloned PCR 2.1 with signal peptide and NdeI and HindIII
pKMB24	pVV Ag85A	Subcloned into pVV16 vector for overexpression in Mtb with His-tag
pKMB25	pVV Ag85B	Subcloned into pVV16 vector for overexpression in Mtb with His-tag
pKMB26	pVV Ag85C	Subcloned into pVV16 vector for overexpression in Mtb with His-tag
pKMB27	pVV Ag85C short	Subcloned into pVV16 vector for overexpression in Mtb with His-tag
pKMB28	pVV Rv0519C	Subcloned into pVV16 vector for overexpression in Mtb with His-tag
pKMB29	pVV Rv1288	Subcloned into pVV16 vector for overexpression in Mtb with His-tag
pKMB30	pVV Rv1288 short	Subcloned into pVV16 vector for overexpression in Mtb with His-tag
pKMB31	GFP EcoRV NotI	cloned into PCR 2.1 with EcoRV and NotI sites
pKMB32	RFP EcoRV NotI	cloned into PCR 2.1 with EcoRV and NotI sites
pKMB33	Ag85ASphI, EcoRV	cloned into PCR 2.1 with SphI and EcoRV sites and no RBS
pKMB34	Ag85B SphI, SmaI	cloned into PCR 2.1 with SphI and SmaI sites and no RBS

Plasmid	Alternate name	Source and description
pKMB35	Ag85C SphI, SmaI	cloned into PCR 2.1 with SphI and SmaI sites and no RBS
pKMB36	pTE-GFP	subcloned into pTE-2M1L at EcoRV and NotI sites
pKMB37	pTE-RFP	subcloned into pTE-2M1L at EcoRV and NotI sites
pKMB38	pTE-Ag85C-GFP	subcloned into pTE-GFP between SphI and EcoRV sites (no RBS)
pKMB39	Ag85ASphI, EcoRV	cloned into PCR 2.1 with SphI and EcoRV sites and RBS
pKMB40	Ag85B SphI, SmaI	cloned into PCR 2.1 with SphI and SmaI sites and RBS
pKMB41	Ag85C SphI, SmaI	cloned into PCR 2.1 with SphI and SmaI sites and RBS
pKMB42	Ag85ASphI, EcoRV SA	cloned into PCR 2.1 with SphI and EcoRV sites and RBS and ser-ala mutation
pKMB43	Ag85B SphI, SmaI SA	cloned into PCR 2.1 with SphI and SmaI sites and RBS and ser-ala mutation
pKMB44	Ag85C SphI, SmaI SA	cloned into PCR 2.1 with SphI and SmaI sites and RBS and ser-ala mutation
pKMB45	Ag85B SphI, SmaI WW	cloned into PCR 2.1 with SphI and SmaI sites and RBS and M159W G160W
pKMB46	Ag85B SphI, SmaI WWL	cloned into PCR 2.1 with SphI and SmaI sites and RBS and M159W G160W F232L
pKMB47	Ag85B SphI, SmaI WWLG	cloned into PCR 2.1 with SphI and SmaI sites and RBS and M159W G160W N231G F232L
pKMB48	Rv0519C SphI, EcoRV	cloned into PCR 2.1 with SphI and EcoRV sites and RBS
pKMB49	Rv1288 SphI, EcoRV	cloned into PCR 2.1 with SphI and EcoRV sites and RBS
pKMB50	pTE-Ag85ASphI, EcoRV	subcloned into pTE-GFP between SphI and EcoRV sites (with RBS)
pKMB51	pTE-Ag85B SphI, SmaI	subcloned into pTE-GFP between SphI and EcoRV sites (with RBS)
pKMB52	pTE-Ag85C SphI, SmaI	subcloned into pTE-GFP between SphI and EcoRV sites (with RBS)
pKMB53	pTE-Ag85A SphI, EcoRV SA	subcloned into pTE-GFP between SphI and EcoRV sites (with RBS)
pKMB54	pTE-Ag85B SphI, SmaI SA	subcloned into pTE-GFP between SphI and EcoRV sites (with RBS)
pKMB55	pTE-Ag85B WW SphI, SmaI	subcloned into pTE-GFP between SphI and SmaI sites and RBS and M159W G160W
pKMB56	pTE-Ag85B WWL SphI, SmaI	subcloned into pTE-GFP between SphI and SmaI sites and RBS and M159W G160W F232L
pKMB57	pTE-Ag85B WWLG SphI, SmaI	subcloned into pTE-GFP between SphI and SmaI sites and RBS and M159W G160W N231G F232L
pKMB58	pTE-Rv1288 SphI, EcoRV	subcloned into pTE-GFP between SphI and EcoRV sites (with RBS)
pKMB59	pTE-Rv0519C SphI, EcoRV	subcloned into pTE-GFP between SphI and EcoRV sites (with RBS)

9.33.1 Selected plasmid maps

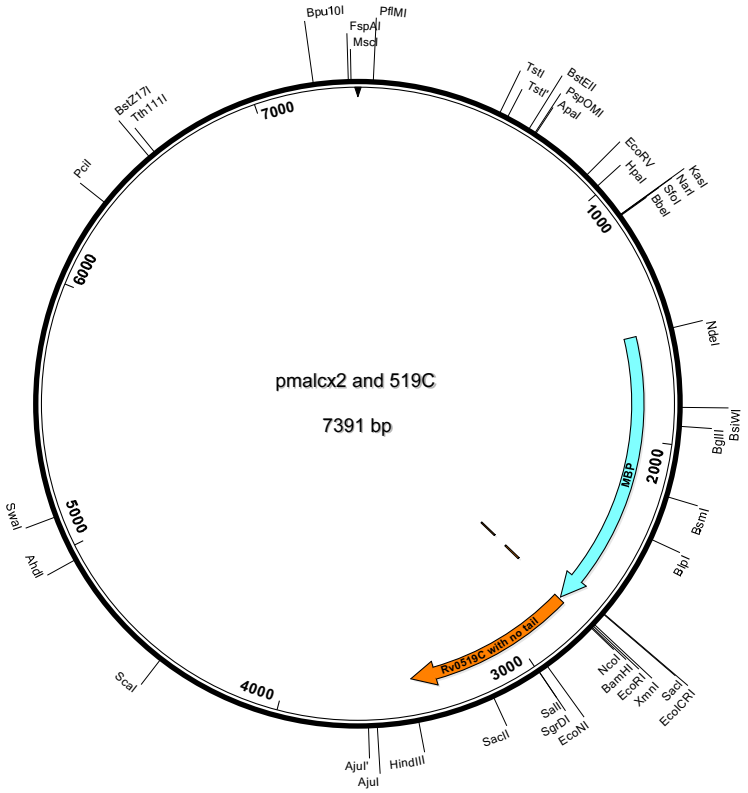
pKMB14, pET23a with truncated Rv1288



TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGC
TACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCC
CCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAA
CTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTC
CACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTT
ATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTT
AACAAAATATTAACGTTTACAATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATT
TTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAA
AGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTT
GCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAA
CTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTA
AAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTA
TTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA
TTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGA
AGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAA
TGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATT
AACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGA
CCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCG
CGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAG
GCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAG
ACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATC
CTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAA
GATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTAC
CAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCA
GATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACAT
ACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCA
AGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAG
CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGA
AAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAA

CGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGG
GGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTC
ACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTC
GCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTC
TCCTTACGCATCTGTGCGGTATTTCACACCGCATATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCAT
AGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCC
GCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT
GCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTC
GTGAAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCT
GGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGA
TTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAAC
ATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCA
CTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGAT
GCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAA
GACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTG
ATTCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCG
CACCCGTGGCCAGGACCCAACGCTGCCCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGAGAC
CACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCTAGCATGACTGGT
GGACAGCAAATGGGTCGCGGATCC**ATGCGCAACGAGAACGATCCCCGCCTGTGGTACTACCGGTTCCAGA**
CCTCCGCGATCGGCTGGAACCCCGGAGTCAACGTCCTGCTTCCCGATGACTACCGCACCAGCGGACGC
ACCTATCCCGTCCTCTACCTGTTCCACGGCGGCGGCACCGACCAGGATTTCCGCACGTTCGACTTTCTG
GGCATCCGCGACCTGACCGCCGGAAAGCCGATCATCATCGTGATGCCCGACGGCGGGCACGCGGGCTG
GTATTCCAACCCGGTCAGCTCGTTCGTCGGCCCACGGAACTGGGAGACATTCCACATCGCCCAGCTGCT
CCCCTGGATCGAGGCGAACTTCCGAACCTACGCCGAATACGACGGCCGCGCGGTCGCCGGGTTTTCGA
TGGGTGGCTTCGGCGCGCTGAAGTACGCAGCAAAGTACTACGGCCACTTCGCGTCGGCGAGCAGCCAC
TCCGGACCGGCAAGTCTGCGCCGCGACTTCGGCCTGGTAGTGCATTGGGCAAACCTGTCCTCGGCGGT
GCTGGATCTAGGCGGCGGCACGGTTTACGGCGCGCCGCTCTGGGACCAAGCTAGGGTCAGCGCCGACA
ACCCGGTCGAGCGTATCGACAGCTACCGCAACAAGCGGATCTTCCTGGTCGCCGGCACCAGTCCGGAC
CCGGCCAACTGGTTCGACAGCGTGAACGAGACCCAGGTGCTAGCCGGGCAGAGGGAGTTCCGCGAACG
CCTCAGCAACGCCGGCATCCCGCATGAATCGCACGAGGTGCCTGGCGGTCACGTCTTCCGGCCCGACA
TGTTCCGTCTCGACCTCGACGGCATCGTCGCCCGGCTGCGCCCCGCGAGCATCGGGGCGGCCGCAGAA
CGCGCCGATAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCC
CGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGG
GTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGAT

pKMB15 pMal-C2X with truncated Rv0519C

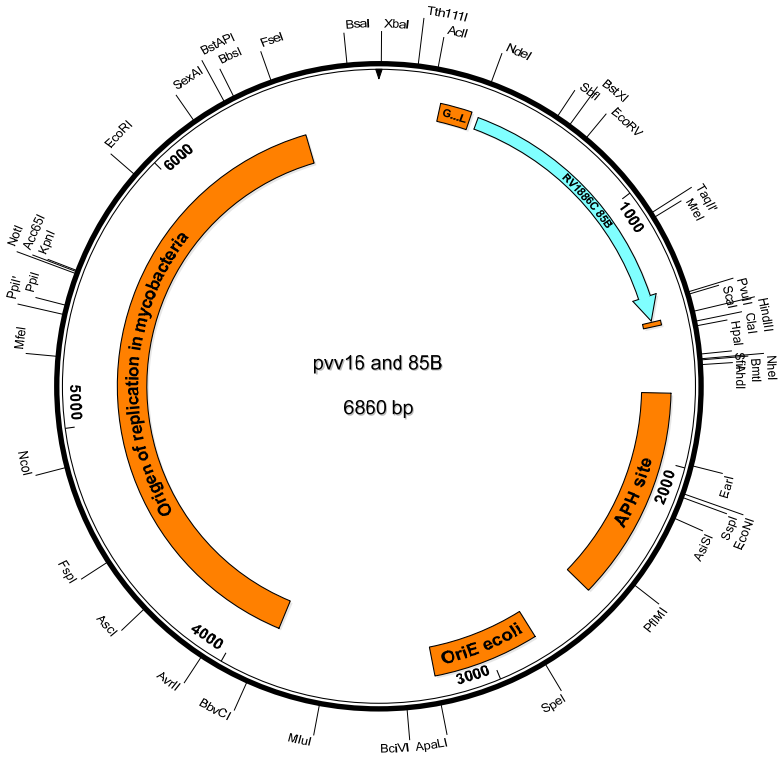


CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGG
TGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCG
CGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAA
TTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGT
CTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGG
TGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCG
TCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGT
TCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGC
GACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGT
CTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGG
GAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTG
CGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGG
TGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATC
AAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA
AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTC
TCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCG
CAACGCAATTAATGTAAGTTAGCTCACTCATTAGGCACAATTCTCATGTTTGACAGCTTATCATCGACTGCACG
GTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCA
TAATTCGTGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAA
ATATTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATT
TCACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGAGCACTTCACCAACAAGGACCATAGCATATGAAA
ATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAG
AAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAG
GTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTG
GCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACG
TTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGC
CGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCG
CTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGT
ATGAAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCC
TGGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTTTAATAA
AGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGG
TGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAAC

GCCGCCAGTCCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAA
GCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCA
CGTATTGCCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCT
GGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGC
GCAGACTAATTCGAGCTCGAACAACAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTC
AGAATTCGGATCC**ATGGAGCCCCCGACGGCGGGCAGCAGCTTGCCGACCAGGATCTCCGGCTCGTTCAT**
CTCGGCGGCTCGCGGAGGCATCAAGACCAACTGGGTGATCTCCATGCCGCCCGGCCAGAGCGGGCAGC
TGCGGCCAGTGATCGCACTGCATGGCAAGGACGGCAACGCCGGCATGATGCTGGATCTCGGCGTCGAG
CAGGGGTTGGCCCGCCTGGTCAAGGAGGGCAAGCCAGCGTTCGCGGTGGTCGGCGTCGACGGCGGCA
ACACCTACTGGCACCGTCGCTCTTCCGGCGGGGACTCGGGTGCGATGGTGCTCGACGAGCTGCTGCCG
ATGCTGACCTCGATGGGTATGGACACCTCCCGGGTGGGATTCCTGGGGTGGTCGATGGGGGGATACGG
CGCGCTGCTCCTGGGTGCCCGGTTGGGTCCGGCCAGAACCGCGGGGATCTGTGCCATCAGCCCGGCAT
TGTTCACGTCGTTCACCGGCAGCACACCCGGGGCGTTCGACAGCTACGACGACTATGTGCAACACAGC
GTGTTGGGCCTGCCCGCGCTGAACTCAATCCCGCTACGGGTGGATTGCGGCACCAGCGACCGCTTCTA
CTTCGCCACTCGCCAATTCGTCAACCAACTGCACCAACCGCCTGCCGGCAGCTTCTCGCCCGGAGGACA
CGACGCGTCGTACTGGCGCGAACAGTTGCCCGGCGAACTGGCTTGGATGGCGTCCAAGCTTGGCACTGG
CCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCG
AATGGCAGCTTGGCTGTTTTGGCGGATGAGATAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAG
CGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAA
GTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAAT
AAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTA
GGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGC
CATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCT
TTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAA
TATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTT
CCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTT
ACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTCCCAATGATGAG
CACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGC
ATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAG
TAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGG
AGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCG
GAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGC
AAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAG
TTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGT
GGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGG
GGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGT
AACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTACCCCGGTTGATAATCAGAAAAGCCCCAAA
AACAGGAAGATTGTATAAGCAAATATTTAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTG
TTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAG
ATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGC
GAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCAAATCAAGTTTTTTGGGGTCGAGGTG
CCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGT
GGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGC
GCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTAAAAGGATCTAGGTGAAGATCCTTT
TTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGAT
CAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAG
CGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGAT
ACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACC
TCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGA
CGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGA
ACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAG
GCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGC
CTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGG
GGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCAC
ATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGC
CGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTC
CTTACGCATCTGTGCGGTATTTCACACCGCATATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAG
TTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGC
TGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC
ATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGT
GCAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGG
CTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATT
TCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACAT
GCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACT
CAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGC
AGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGA

CCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGAT
TCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCA
CCCGTGGCCAGGACCCAACGCTGCCCGAAATT

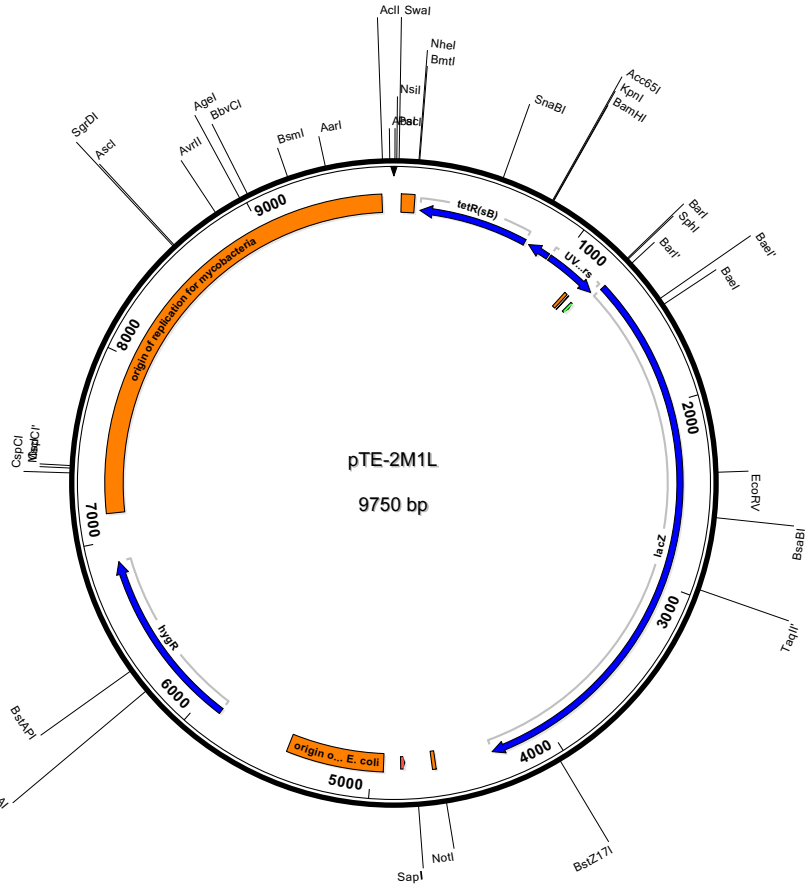
pKMB25, pVV16 and Ag85B



CAAATCTAGAGGTGACCACAACGACGCGCCCGCTTTGATCGGGGACGTCTGCGGCCGACCATTTACGGGTCTT
GTTGTCGTTGGCGGTCATGGGCCGAACATACTCACCCGGATCGGAGGGCCGAGGACAAGGTCGAACGAGGGG
CATGACCCGGTGCGGGGCTTCTTGCACTCGGCATAGGCGAGTGCTAAGAATAACGTTGGCACTCGCGACCGGT
GAGTCGTAGGTCGGGACGGTGAGGCCAGGCCCGTCGTCGCAGCGAGTGGCAGCGAGGACAACTTGAGCCGTC
CGTCGCGGGCACTGCGCCCGGCCAGCGTAAGTAGCGGGGTTGCCGTCACCCGGTGACCCCCGGTTTCATCCCC
GATCCGGAGGAATCACTTCCAT**ATGAGACGACTTTGCGCCCGAATCGACATTTGGCCTCCACACACGGTA**
TGTTCTGGCCCGAGCACACGACGACATACAGGACAAAGGGGCACAAGTATGACAGACGTGAGCCGAAA
GATTCGAGCTTGGGGACGCCGATTGATGATCGGCACGGCAGCGGCTGTAGTCCTTCCGGGCCTGGTGG
GGCTTGCCGGCGGAGCGGCAACCGCGGGCGCGTTCTCCCGGCCGGGGCTGCCGGTCGAGTACCTGCA
GGTGCCGTCGCCGTCGATGGGCCGCGACATCAAGGTTCAGTTCCAGAGCGGTGGGAACAACTCACCTG
CGGTTTATCTGCTCGACGGCCTGCGCGCCCAAGACGACTACAACGGCTGGGATATCAACACCCCGGCG
TTCGAGTGGTACTACCAGTCGGGACTGTCGATAGTCATGCCGGTCGGCGGGCAGTCCAGCTTCTACAG
CGACTGGTACAGCCCGGCCTGCGGTAAGGCTGGCTGCCAGACTTACAAGTGGGAAACCTTCCTGACCA
GCGAGCTGCCGCAATGGTTGTCCGCCAACAGGGCCGTGAAGCCCACCGGCAGCGCTGCAATCGGCTTG
TCGATGGCCGGCTCGTCGGCAATGATCTTGGCCGCCTACCACCCCCAGCAGTTCATCTACGCCGGCTCG
CTGTCGGCCCTGCTGGACCCCTCTCAGGGGATGGGGCCTAGCCTGATCGGCCTCGCGATGGGTGACGC
CGGCGGTTACAAGGCCGCAGACATGTGGGGTCCCTCGAGTGACCCGGCATGGGAGCGCAACGACCCTA
CGCAGCAGATCCCCAAGCTGGTCGCAAACAACACCCGGCTATGGGTTTATTGCGGGAACGGCACCCCG
AACGAGTTGGGCGGTGCCAACATACCCGCCGAGTTCTTGGAGAACTTCGTTCGTAGCAGCAACCTGAA
GTTCCAGGATGCGTACAACGCCGCGGGCGGGCACAACGCCGTGTTCAACTTCCCGCCCAACGGCACGC
ACAGCTGGGAGTACTGGGGCGCTCAGCTCAACGCCATGAAGGGTGACCTGCAGAGTTCGTTAGGCGCC
GGCAAGCTTCACCACCACCACCACCACTGACAGCTTATCGATGTCGACGTAGTTAACTAGCGTACGATCGACT
GCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCCGGCC
ATCATGGCCGCGGTGATCAGCTAGCCAACAAAGCGACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAA
GATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCC
ATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAA
ATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGCTTGTATGGGAAGCCCCATGCGCCAGAG
TTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGA
CGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCG

ATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGG
CAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCG
CTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCC
TGTTGAACAAGTCTGGAAAGAAATGCATAATCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATT
TCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCA
GACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTT
TCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAAT
CAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGGCTTTGTTGAATA
AATCGAACTTTTGCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGC
AAAAGTTCAAAATCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCT
GGATGATGGGGCGATTCAGGCCTGGTATGAGTCAGCAACACCTTCTTCACGAGGCAGACCTCACTAGTTCCAC
TGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTT
GCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAG
GTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCA
AGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAA
GTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGG
TTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTGAGA
AAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGC
GCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGA
GCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGG
TTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATT
ACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAACGCGTGAGCCCACCAGCTCC
GTAAGTTCGGGTGCTGTGTGGCTCGTACCCGCGCATTCAGGCGGCAGGGGGTCTAACGGGTCTAAGGCGGCGT
GTACGGCCGCCACAGCGGCTCTTAGCGGCCCGGAAACGTCCTCGAAACGACGCATGTGTTCCTCCTGGTTGGT
ACAGGTGGTTGGGGGTGCTCGGCTGTCGCTGGTGTTTCATCATCAGGGCTCGACGGGAGAGCGGGGGAGTGTG
CAGTTGTGGGGTGGCCCCTCAGCGAAATATCTGACTTGGAGCTCGTGTCGGACCATACACCGGTGATTAATCG
TGGTTTATTATCAAGCGTGAGCCACGTCGCCGACGAATTTGAGCAGCTCTGGCTGCCGTACTGGTCCCTGGCA
AGCGACGATCTGCTCGAGGGGATCTACCGCCAAAGCCGCGCGTCGGCCCTAGGCCGCCGGTACATCGAGGCG
AACCCAACAGCGCTGGCAAACCTGCTGGTCGTGGACGTAGACCATCCAGACGCAGCGCTCCGAGCGCTCAGC
GCCCGGGGGTCCCATCCGCTGCCCAACGCGATCGTGGGCAATCGCGCCAACGGCCACGCACACGCAGTGTGG
GCACTCAACGCCCCTGTTCCACGCACCGAATACGCGCGGCGTAAGCCGCTCGCATACATGGCGGCGTGCGCCG
AAGGCCTTCGGCGCGCCGTCGATGGCGACCGCAGTTACTCAGGCCTCATGACCAAAAACCCCGGCCACATCGC
CTGGGAAACGGAATGGCTCCACTCAGATCTCTACACACTCAGCCACATCGAGGCCGAGCTCGGCGCGAACAT
GCCACCGCCGCGCTGGCGTCAGCAGACCACGTACAAAGCGGCTCCGACGCCGCTAGGGCGGAATTGCGCACT
GTTCGATTCCGTCAGGTTGTGGGCCTATCTTCCCGCCCTCATGCGGATCTACCTGCCGACCCGGAACGTGGACG
GACTCGGCCGCGCGATCTATGCCGAGTGCCACGCGCGAAACGCCGAATTTCCGTGCAACGACGTGTGTCCCGG
ACCGCTACCGGACAGCGAGGTCCGCGCCATCGCCAACAGCATTTGGCGTTGGATCACAACCAAGTCGCGCATT
TGGGCGGACGGGATCGTGGTCTACGAGGCCACACTCAGTGCGCGCCATGCGGCCATCTCGCGGAAGGGCGCA
GCAGCGCGCACGGCGGCGAGCACAGTTGCGCGGCGCGCAAAGTCCGCGTCAGCCATGGAGGCATTGCTATGA
GCGACGGCTACAGCGACGGCTACAGCGACGGCTACAACTGGCAGCCGACTGTCCGCAAAAAGCGGCGCGTGA
CCGCCGCCGAAGGCGCTCGAATCACCGGACTATCCGAACGCCACGTCGTCCGGCTCGTGGCGCAGGAACGCA
GCGAGTGGTTCGCCGAGCAGGCTGCACGCCGCGAACGCATCCGCGCCTATCACGACGACGAGGGCCACTCTT
GGCCGCAAACGGCCAAACATTTCGGGCTGCATCTGGACACCGTTAAGCGACTCGGCTATCGGGCGAGGAAAG
AGCGTGCGGCAGAACAGGAAGCGGCTCAAAAGGCCCACAACGAAGCCGACAATCCACCGCTGTTCTAACGCA
ATTGGGGAGCGGGTGTCGCGGGGGTTCCGTGGGGGGTTCCGTTGCAACGGGTCGGACAGGTAAAAGTCCTGG
TAGACGCTAGTTTTCTGGTTTGGGCCATGCCTGTCTCGTTGCGTGTTTCGTTGCGTCCGTTTTGAATACCAGCCA
GACGAGACGGGGTTCTACGAATCTTGGTCGATACCAAGCCATTTCCGCTGAATATCGTGGAGCTCACCGCCAG
AATCGGTGGTTGTGGTGATGTACGTGGCGAACTCCGTTGTAGTGCTTGTGGTGGCATCCGTGGCGCGGCCGCG
GTACCAGATCTTTGATCCGGCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAA
GTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCT
ATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGACGTGGCCGACCAGCCCGTCATCGT
CAACGCCTGACCGCGGTGCGGACAGGCCGTGTCGCGACCGGCCGTGCGGAATTAAGCCGGCCCGTACCCTGT
GAATAGAGGTCCGCTGTGACACAAGAATCCCTGTTACTTCTCGACCGTATTGATTCGGATGATTCCTACGCGA
GCCTGCGGAACGACCAGGAATTCTGGGAGCCGCTGGCCCGCCGAGCCCTGGAGGAGCTCGGGCTGCCGGTGC
CGCCGGTGCTGCGGGTGCCCGGCGAGAGCACCAACCCCGTACTGGTCGGCGAGCCCGACCCGGTGATCAAGC
TGTTCGGCGAGCACTGGTGCGGTCCGGAGAGCCTCGCGTCGGAGTCGGAGGCGTACGCGGTCCTGGCGGACG
CCCCGGTGCCGGTGCCCCGCCTCCTCGGCCGCGGCGAGCTGCGGCCCGGCACCGGAGCCTGGCCGTGGCCCTA
CCTGGTGATGAGCCGGATGACCGGCACCACCTGGCGGTCCGCGATGGACGGCACGACCGACCGGAACGCGCT
GCTCGCCCTGGCCCGCGAACTCGGCCGGGTGCTCGGCCGGCTGCACAGGGTGCCGCTGACCGGGAACACCGT
GCTCACCCCCCATTCCGAGGTCTTCCCGGAACTGCTGCGGGAACGCCGCGCGGCGACCGTCGAGGACCACCGC
GGGTGGGGCTACCTCTCGCCCCGGCTGCTGGACCGCCTGGAGGACTGGCTGCCGGACGTGGACACGCTGCTGG
CCGGCCGCGAACCCCGGTTCGTCCACGGCGACCTGCACGGGACCAACATCTTCGTGGACCTGGCCGCGACCGA
GGTCACCGGGATCGTCGACTTCACCGACGTCTATGCGGGAGACTCCCGCTACAGCCTGGTGCAACTGCATCTC
AACGCCTTCCGGGGCGACCGCGAGATCCTGGCCGCGCTGCTCGACGGGGCGCAGTGGAAGCGGACCGAGGAC
TTCGCCCGCGAACTGCTCGCCTTCACCTTCCTGCACGACTTCGAGGTGTTCGAGGAGACCCCGCTGGATCTCTC
CGGCTTCACCGATCCGGAGGAACTGGCGCAGTTCCTCTGGGGGCCGCCGGACACCGCCCCCGGCGCCTGACGC
CCCCCGGATC

pTE-2M1L

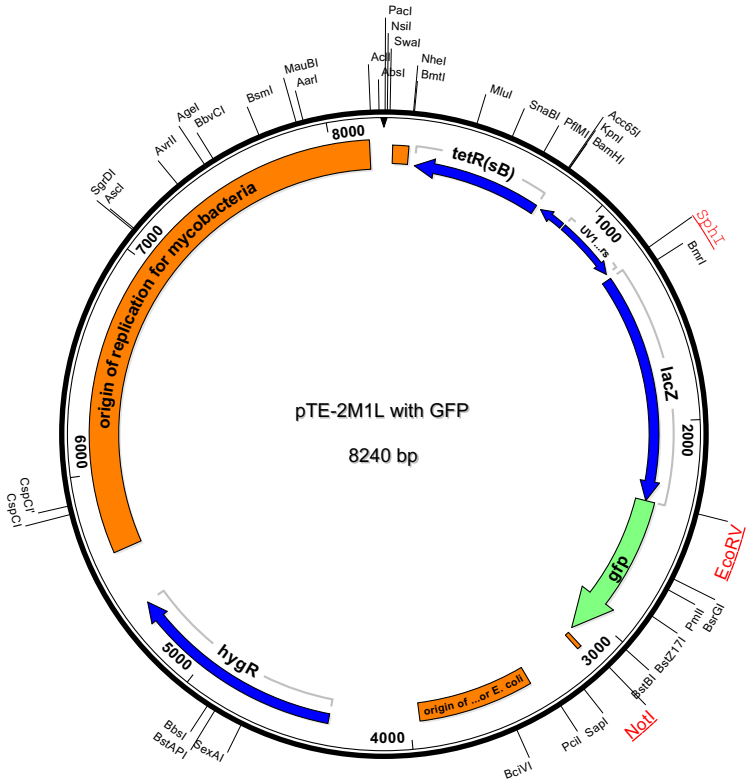


TTAATTAATATGCATCAATTGATTTAAATACTAGTGTTAACTATTTAATTGGGGACCCTAGAGGTCCCCTTTTTT
ATTTTAAAAATTTTTTCNAAGGGGACCTCTAGGGTCCCCAGCTGGCTAGCGAGTCATGAGGTNACCATGGTCA
TTAAGACCCACTTTCACATTTAAGTTGTTTTTCTAATCCGCATATGATCAATTCAAGGCCGAATAAGAAGGCTG
GCTCTGCACCTTGGTGATCAAATAATTCGATAGCTTGTCGTAATAATGGCGGCATACTATCAGTAGTAGGTGTT
TCCCTTTCTTCTTTAGCGACTTGATGCTCTTGATCTTCCAATACGCAACCTAAGGTAAAATGCCCCACAGCGCT
GAGTGCATATAACGCGTTCTCTAGTGAAAAACCTTGTTGGCATAAAAAGGCTAATTGATTTTCGAGAGTTTCA
TACTGTTTTTCTGTAGGCCGTGTACCTAAATGTACTTTTGCTCCATCGCGATGACTTAGTAAAGCACATCTAAA
ACTTTTAGCGTTATTACGTAAAAAATCTTGCCAGCTTTCCCCTTCTAAAGGGCAAAAGTGAGTATGGTGCCTAT
CTAACATCTCAATGGCTAAGGCGTCGAGCAGGGCCCGCTTGTTCTTGACGTGCCAGTACAGGGTCGGCTGCTC
GACGCCCAGCTTCTGGGCCAGCTTGCGGGTGGTCAGGCCCTCGATGCCGACCTCGTTCAGCAGCTCCAGGGCG
GAGTTGATGACCTTGGACTTGTCCAGGCGGGACATTGCGAAGTGATTCCTCCTGATGGTACCGGATCCGTGGC
AGGAGGGTCGCCTGCGTCGGCAACTCGCCCAGGCTACACGAGAGCCCTCGACTGCGAAAATGCCGACCAGGC
ACCACCGCACTCCAGGCTTAACCCGGTGCGACGATCCAATATTACGAGATCTACTAGAAATATTGGATCGTCG
GCACCGTCACGGCCGTGGGAGGCGGCACGATCCGCGACGTGATGATCGGCCGCATCCCCACGGTGCTGCGCA
GTGAGCTCTACGCCATCCCGGCGTTGATCTGTGCGTTCGCACGCACAGGCCCGGTGTGAGAAGGGTCTCTGAC
GAGCGGGAGAACTCCCTATCAGTGATAGAGTTTGTCCTCCCTATCAGTGATAGATAGGCTCTGGGAGTACCCG
TGTGTACGACCAGCACGGCATACATCATTTCGACGCCGAGAGATTCGCCGCCCGAAATGAGCACGATCCGCTC
GAGGCATGCGAAGGAGATATACATATGGAGATCGATCCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTG
GCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCAC
CGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGCG
GTGCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGATACTGTCGTCGTCCCCTCAAACTGGCAGATGC
ACGGTTACGATGCGCCCATCTACACCAACGTAACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAG
AATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTA
TTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAACGGGCGCTGGGTCGGTTACGGCCAGGACAGTCGT
TTGCCGTCTGAATTTGACCTGAGCGCATTTTTACGCGCCGGAGAAAACCGCCTCGCGGTGATGGTGCTGCGTT

GGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGCATTTTCCGTGACGTCTCGTTGCT
GCATAAACCGACTACACAAATCAGCGATTTCCATGTTGCCACTCGCTTTAATGATGATTTCAGCCGCGCTGTAC
TGGAGGCTGAAGTTCAGATGTGCGGCGAGTTGCGTGACTACCTACGGGTAACAGTTTCTTTATGGCAGGGTGA
AACGCAGGTCGCCAGCGGCACCGCGCCTTTCGGCGGTGAAATTATCGATGAGCGTGGTGGTTATGCCGATCGC
GTCACACTACGTCTGAACGTCGAAAACCCGAAACTGTGGAGCGCCGAAATCCCGAATCTCTATCGTGCGGTGG
TTGAACTGCACACCGCCGACGGCACGCTGATTGAAGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTGCGGAT
TGAAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATTCGAGGCGTTAACCGTCACGAGCATCATCCT
CTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGATATCCTGCTGATGAAGCAGAACAACTTTAAC
GCCGTGCGCTGTTCGCATTATCCGAACCATCCGCTGTGGTACACGCTGTGCGACCGCTACGGCCTGTATGTGGT
GGATGAAGCCAATATTGAAACCCACGGCATGGTGCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCG
GCGATGAGCGAACGCGTAACGCGAATGGTGCAGCGCGATCGTAATCACCCGAGTGTGATCATCTGGTCGCTG
GGGAATGAATCAGGCCACGGCGCTAATCACGACGCGCTGTATCGCTGGATCAAATCTGTCGATCCTTCCCGCC
CGGTGCAGTATGAAGGCGGCGGAGCCGACACCACGGCCACCGATATTATTTGCCCGATGTACGCGCGCGTGG
ATGAAGACCAGCCCTTCCCGGCTGTGCCGAAATGGTCCATCAAAAAATGGCTTTCGCTACCTGGAGAGACGCG
CCCGCTGATCCTTTGCGAATACGCCCACGCGATGGGTAACAGTCTTGGCGGTTTCGCTAAATACTGGCAGGCG
TTTCGTCAGTATCCCCGTTTACAGGGCGGCTTCGTCTGGGACTGGGTGGATCAGTCGCTGATTAAATATGATGA
AAACGGCAACCCGTGGTCGGCTTACGGCGGTGATTTTGGCGATACGCCGAACGATCGCCAGTTCTGTATGAAC
GGTCTGGTCTTTGCCGACCGCACGCCGCATCCAGCGCTGACGGAAGCAAAACACCAGCAGCAGTTTTTCCAGT
TCCGTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGATAACGAGCTCCTGCA
CTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGTCGCTCCACAAGGTAA
ACAGTTGATTGAACTGCCTGAACTACCGCAGCCGGAGAGCGCCGGGCAACTCTGGCTCACAGTACGCGTAGTG
CAACCGAACGCGACCGCATGGTCAGAAGCCGGGCACATCAGCGCCTGGCAGCAGTGGCGTCTGGCGGAAAAC
CTCAGTGTGACGCTCCCCGCCGCGTCCCACGCCATCCCGCATCTGACCACCAGCGAAATGGATTTTTGCATCG
AGCTGGGTAATAAGCGTTGGCAATTTAACCGCCAGTCAGGCTTTCTTTCACAGATGTGGATTGGCGATAAAAA
ACAACTGCTGACGCCGCTGCGCGATCAGTTCACCCGTGCACCGCTGGATAACGACATTGGCGTAAGTGAAGCG
ACCCGCATTGACCCTAACGCCTGGGTCGAACGCTGGAAGGCGGCGGGCCATTACCAGGCCGAAGCAGCGTTG
TTGCAGTGCACGGCAGATACACTTGCTGATGCGGTGCTGATTACGACCGCTCACGCGTGGCAGCATCAGGGGA
AAACCTTATTTATCAGCCGGAAAACCTACCGGATTGATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGA
AGTGGCGAGCGATACACCGCATCCGGCGCGGATTGGCCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGT
AAACTGGCTCGGATTAGGGCCGCAAGAAAACTATCCCGACCGCCTTACTGCCGCCTGTTTTGACCGCTGGGAT
CTGCCATTGTCAGACATGTATACCCCGTACGTCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGACGCGCGAAT
TGAATTATGGCCCACACCAGTGGCGCGGCGACTTCCAGTTCAACATCAGCCGCTACAGTCAACAGCAACTGAT
GGAAACCAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCACATGGCTGAATATCGACGGTTTCCATATGGG
GATTGGTGGCGACGACTCCTGGAGCCCGTCAGTATCGGCGGAATTCCAGCTGAGCGCCGGTCGCTACCATTAC
CAGTTGGTCTGGTGTCAAAAATAATAATAACCGGGCAGGCCATGTCTGCCCGTATTTCGCGTAAGGAAATCCA
TTATGTACTATTTAAAAAACACAAACTTTTGGATGTTCGGTTTATTCTTTTTCTTTTACTTTTTTATCATGGGAG
CCTACTTCCCGTTTTTCCCGATTTGGCTACATGACATCAACCATATCAGCAAAAGTGATACGGGTATTATTTTT
GCCGCTATTTCTCTGTTCTCGCTATTATTCCAACCGCTGTTTGGTCTGCTTTCTGACAAACTCGGAACTTGTTTA
TTGCAGCTTAAAACGCGGCCGCAAACAGTATTAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAG
CTAATCAACTAGTTCTAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT
CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAG
GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAA
GGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAA
AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGC
TCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG
GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCA
CGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACAC
GACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG
TTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG
TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTT
TGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACG
CTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCT
TTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCT
TAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGGACGTGGC
CGACCAGCCCGTCATCGTCAACGCCTGACCGCGGTGCGGACAGGCCGTGTCGCGACCGGCCGTGCGGAATTA
AGCCGGCCCGTACCCTGTGAATAGAGGTCCGCTGTGACACAAGAATCCCTGTTACTTCTCGACCGTATTGATTC
GGATGATTCCTACGCGAGCCTGCGGAACGACCAGGAATTCTGGGAGCCGCTGGCCCGCCGAGCCCTGGAGGA
GCTCGGGCTGCCGGTGCCGCCGGTGCTGCGGGTGCCCGGCGAGAGCACCAACCCCGTACTGGTCGGCGAGCC
CGACCCGGTGATCAAGCTGTTCGGCGAGCACTGGTGCGGTCCGGAGAGCCTCGCGTCGGAGTCGGAGGCGTA
CGCGGTCCTGGCGGACGCCCCGGTGCCGGTGCCCCGCCTCCTCGGCCGCGGCGAGCTGCGGCCCGGCACCGGA
GCCTGGCCGTGGCCCTACCTGGTGATGAGCCGGATGACCGGCACCACCTGGCGGTCCGCGATGGACGGCACG
ACCGACCGGAACGCGCTGCTCGCCCTGGCCCGCGAACTCGGCCGGGTGCTCGGCCGGCTGCACAGGGTGCCG
CTGACCGGGAACACCGTGCTCACCCCCCATTCCGAGGTCTTCCCGGAACTGCTGCGGGAACGCCGCGCGGCGA
CCGTCGAGGACCACCGCGGGTGGGGCTACCTCTCGCCCCGGCTGCTGGACCGCCTGGAGGACTGGCTGCCGGA
CGTGGACACGCTGCTGGCCGGCCGCGAACCCCGGTTCGTCCACGGCGACCTGCACGGGACCAACATCTTCGTG
GACCTGGCCGCGACCGAGGTCACCGGGATCGTCGACTTCACCGACGTCTATGCGGGAGACTCCCGCTACAGCC
TGGTGCAACTGCATCTCAACGCCTTCCGGGGCGACCGCGAGATCCTGGCCGCGCTGCTCGACGGGGCGCAGTG

GAAGCGGACCGAGGACTTCGCCCGCGAACTGCTCGCCTTCACCTTCCTGCACGACTTCGAGGTGTTCGAGGAG
ACCCCGCTGGATCTCTCCGGCTTCACCGATCCGGAGGAACTGGCGCAGTTCCTCTGGGGGCCGCCGGACACCG
CCCCCGGCGCCTGACGCCCCGGGCCGCCCGGCGCCGCCCCCGGCCCCCGGCGGCCGGCCGCCCGGAGCCCCG
CCCGCGCTCGGGAGCCCCGGGCCCGCGCCGAAGCCCGCTGCTGCGAGCCCGGAGCGGGCCGGCCGACGGCGG
TGCGGGCCCGGCGGCGGACGCTCAGCAGCGGCGGGCGTGAAAGGCCCTGGCATCCTCGATCATCTCCTCCAG
GGTGGTCGGCCGAGCTTCCATCCCAGCTCGGCAAGGATCAACGAGGACAGTCGCACGACGAAGTTCTTCTGGA
TCGCGCCCGTGCTGGAAGCACTCAACCTCGAAGCGTGTGGTTGCGGAGCCATCTAGCAACCACACGAAACATG
CGCAACGAACCGCGCAACGAACAACGCCTAGAACTGGCCCTAGATGAGCTGACTCGTATCGTTGGTAAACCT
AGTTTGACCAGCATGTTTTAACTACGTTCGGTGAGCTGTCAACGGGGCCTGTAACGGCACAACGAACCGTGCA
ACGAGAGTGGCCACGGATGCCACCACAGGCACTACAACGGAGTTCGCCACGTACATCACCACAACCACCGAT
TCTGGCGGTGAGCTCCCCGATATTCAGCGGAAATGGCTTGGTATCGACCAAGATTCGTAGAACCCCGTCTCGT
CTGGCTGGTATTCAAAACGGGCGCAACGAAACACGCAACGAGACAGGCATGGCCCAAACCAGAAAACTAGCG
TCTACCAGGACTTTTACCTGTCCGACCCGTTGCAACGGAACCCCCCACGGAACCCCCGCGACACCCGCTCCCC
AATTGCGTTAGAACAGCGGTGGATTGTCGGCTTCGTTGTGGGCCTTTTGAGCCGCTTCCTGTTCTGCCGCACGC
TCTTTCCTCGCCCGATAGCCGAGTCGCTTAACGGTGTCCAGATGCAGCCCGAAATGTTTGGCCGTTTGCGGCCA
AGAGTGGCCCTCGTCGTCGTGATAGGCGCGGATGCGTTCGCGGCGTGCAGCCTGCTCGGCGAGCCACTCGCTG
CGTTCCTGCGCCACGAGCCGGACGACGTGGCGTTCGGATAGTCCGGTGATTCGAGCGCCTTCGGCGGCGGTCA
CGCGCCGCTTTTTGCGGACAGTCGGCTGCCGGTTGTAGCCGTCGCTGTAGCCGTCGCTGTAGCCGTCGCTCATA
GCAATGCCTCCATGGCTGACGCGGACTTTGCGCGCCGCGCAACTGTGCTCGCCGCCGTGCGCGCTGCTGCGCC
CTTCCGCGAGATGGCCGACTGGCGCGCACTGAGTGTGGCCTCGTAGACCACGATCCCGTCCGCCCAAATGCGC
GACTTGGTTGTGATCCAACGCCAAATGCTGTTGGCGATGGCGCGGACCTCGCTGTCCGGTAGCGGTCCGGGAC
ACACGTCGTTGCACGGGAATTCGGCGTTTCGCGCGTGGCACTCGGCATAGATCGCGCGGCCGAGTCCGTCCAC
GTTCCGGGTCGGCAGGTAGATCCGCATGAGGGCGGGACGATAGGCCCACAACCTGACGGAATCGAACAGTGC
GCAATTCCGCCCTAGCGGCGTCGGAGCCGCTTTGTACGTGGTCTGCTGACGCCAGCGCGGCGGTGGCATGTTC
GCGCCGAGCTCGGCCTCGATGTGGCTGAGTGTGTAGAGATCTGAGTGGAGCCATTCCGTTTCCCAGGCGATGT
GGCCGGGGTTTTTGGTCATGAGGCCTGAGTAACTGCGGTCGCCGTCGACGGCGCGCCGAAGGCCTTCGGCGCA
CGCCGCCATGTATGCGAGCGGCTTACGCCGCGCGTATTCGGTGCGTGGAACAGGGGCGTTGAGTGCCCACACT
GCGTGTGCGTGGCCGTTGGCGCGATTGCCCACGATCGCGTTGGGCAGCGGATGGGACCCCCGGGCGCTGAGC
GCTCGGAGCGCTGCGTCTGGATGGTCTACGTCCACGACCAGCAGGTTTGCCAGCGCTGTTGGGTTCGCCTCGA
TGTACCGGCGGCCTAGGGCCGACGCGCGGCTTTGGCGGTAGATCCCCTCGAGCAGATCGTCGCTTGCCAGCGG
CCAGTACGGCAGCCAGAGCTGCTCAAATTCGTCGGCGACGTGGCTCACGCTTGGTAGTAGACCACGATTAATC
ACCGGTGTATGGTCCGACACGAGCTCCAAGTCAGATATTTCGCTGAGGGGCCACCCCACAACTGCACACTCCC
CCGCTCTCCCGTCGAGCCCTGGTGGTGGAACACCAGCGACAGCCGAGCACCCCCAACCACCTGTACCAACCAG
GAGGAACACATGCGTCGTTTCGAGGACGTTTCCGGGCCGCTGAGAGCCGCTGTGGCGGCCGTACACGCCGCCT
TAGACCCGTTAGACCCCCTGCCGCCTGAATGCGCGGGTACGAGCCACACAGCGCCCGAACTTACGGAGCTGGT
GGGCTCACCTGGCTTTATGGCGTACGAATCGGCTGTGTGCGACCTGTTGGGCGAGGTGAGGTACGCGCTACTC
ACGCTGGCAAGGGCGACACAGCCGCCCCACCGAGCCCGCACGGCCGCGCGCGGTGTCAACAACCGGGTGAGT
CGTGCACACCAGCAGGTGTTCGAGGCTTGGCTCGAAGTGCAGGACATCGTGGCGAACGCCGCCCGATGAGCC
GCGCCTTACGCTGGCTGCCAGCCGTTCGCGGGCTGGTTGGTGCAGCGCGTCGAGCGGTTAGAGGCCCTGCGGT
GTTCCACCACCGCAGGCCTCGCCCTTTTTAAGGCTGAATTTGCTTGTCTCCGAATCCAACTGGCTTGTCCAAGG
GTGTATCTACGCTTAATCCAAAGTTCAAACGAGGGGATTACACATGACCAACTTCGATAACGTTCTCGGCTCG
ATGATCCCCGGCGGGCCCCCCCTCGAGGTCGACGGTATCGAT

pKMB36, pTE-2M1L with GFP

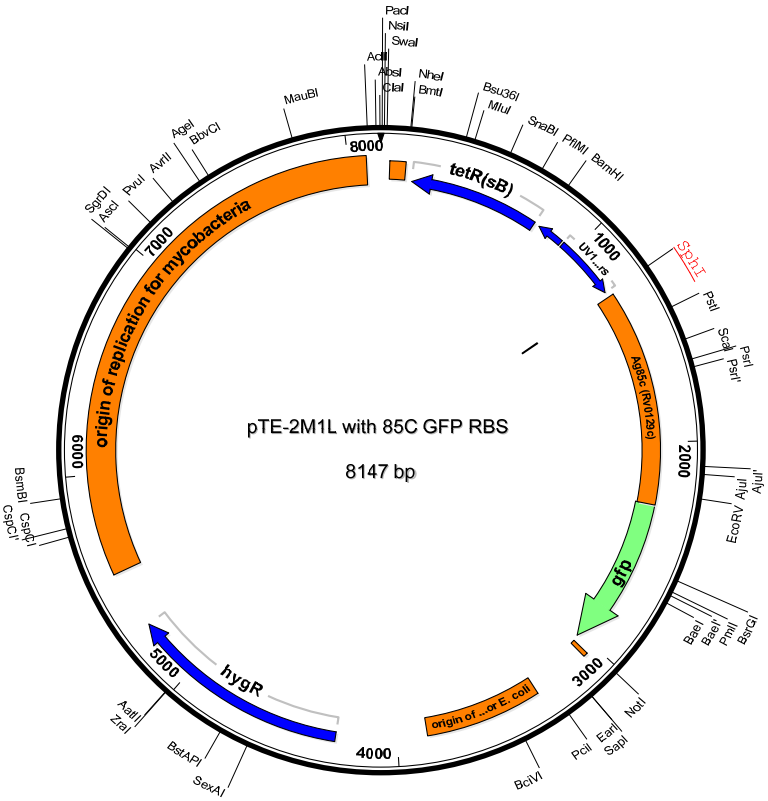


TTAATTAATATGCATCAATTGATTTAAATACTAGTGTTAACTATTTAATTGGGGACCCTAGAGGTCCCCTTTTTT
ATTTTAAAAATTTTTTCNAAGGGGACCTCTAGGGTCCCCAGCTGGCTAGCGAGTCATGAGGTNACCATGGTCA
TTAAGACCCACTTTCACATTTAAGTTGTTTTTCTAATCCGCATATGATCAATTCAAGGCCGAATAAGAAGGCTG
GCTCTGCACCTTGGTGATCAAATAATTCGATAGCTTGTCGTAATAATGGCGGCATACTATCAGTAGTAGGTGTT
TCCCTTTCTTCTTTAGCGACTTGATGCTCTTGATCTTCCAATACGCAACCTAAGGTAAAATGCCCCACAGCGCT
GAGTGCATATAACGCGTTCTCTAGTGAAAAACCTTGTTGGCATAAAAAGGCTAATTGATTTTCGAGAGTTTCA
TACTGTTTTTCTGTAGGCCGTGTACCTAAATGTACTTTTGCTCCATCGCGATGACTTAGTAAAGCACATCTAAA
ACTTTTAGCGTTATTACGTAAAAAATCTTGCCAGCTTTCCCCTTCTAAAGGGCAAAAGTGAGTATGGTGCCTAT
CTAACATCTCAATGGCTAAGGCGTCGAGCAGGGCCCGCTTGTTCTTGACGTGCCAGTACAGGGTCGGCTGCTC
GACGCCCAGCTTCTGGGCCAGCTTGCGGGTGGTCAGGCCCTCGATGCCGACCTCGTTCAGCAGCTCCAGGGCG
GAGTTGATGACCTTGGACTTGTCCAGGCGGGACATTGCGAAGTGATTCCTCCTGATGGTACCGGATCCGTGGC
AGGAGGGTCGCCTGCGTCGGCAACTCGCCCAGGCTACACGAGAGCCCTCGACTGCGAAAATGCCGACCAGGC
ACCACCGCACTCCAGGCTTAACCCGGTGCGACGATCCAATATTACGAGATCTACTAGAAATATTGGATCGTCG
GCACCGTCACGGCCGTGGGAGGCGGCACGATCCGCGACGTGATGATCGGCCGCATCCCCACGGTGCTGCGCA
GTGAGCTCTACGCCATCCCGGCGTTGATCTGTGCGTTCGCACGCACAGGCCCGGTGTGAGAAGGGTCTCTGAC
GAGCGGGAGAACTCCCTATCAGTGATAGAGTTTGTCCTCCCTATCAGTGATAGATAGGCTCTGGGAGTACCCG
TGTGTACGACCAGCACGGCATACATCATTTCGACGCCGAGAGATTCGCCGCCCGAAATGAGCACGATCCGCTC
GAGGCATGCGAAGGAGATATACATATGGAGATCGATCCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTG
GCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCAC
CGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGCG
GTGCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGATACTGTCGTCGTCCCCTCAAACTGGCAGATGC
ACGGTTACGATGCGCCCATCTACACCAACGTAACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAG
AATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTA
TTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAACGGGCGCTGGGTCGGTTACGGCCAGGACAGTCGT
TTGCCGTCTGAATTTGACCTGAGCGCATTTTTACGCGCCGGAGAAAACCGCCTCGCGGTGATGGTGCTGCGTT
GGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGCATTTTCCGTGACGTCTCGTTGCT
GCATAAACCGACTACACAAATCAGCGATTTCCATGTTGCCACTCGCTTTAATGATGATTTCAGCCGCGCTGTAC
TGGAGGCTGAAGTTCAGATGTGCGGCGAGTTGCGTGACTACCTACGGGTAACAGTTTCTTTATGGCAGGGTGA
AACGCAGGTCGCCAGCGGCACCGCGCCTTTCGGCGGTGAAATTATCGATGAGCGTGGTGGTTATGCCGATCGC
GTCACACTACGTCTGAACGTCGAAAACCCGAAACTGTGGAGCGCCGAAATCCCGAATCTCTATCGTGCGGTGG
TTGAACTGCACACCGCCGACGGCACGCTGATTGAAGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTGCGGAT
TGAAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATTCGAGGCGTTAACCGTCACGAGCATCATCCT
CTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGAT**ATCCATATGAGTAAAGGAGAAGAACTTTT**

CACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGG
AGAGGGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACC
TGTTCCATGGCCAACACTTGTCACTACTTTCGGTTATGGTGTTCAATGCTTTGCGAGATACCCAGATCAT
ATGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTC
AAAGATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATC
GAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAAC
TCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACAC
AACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCT
GTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGA
GACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAA
TAAAGCGGCCGCAAACAGTATTAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAGCTAATCAACT
AGTTCTAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTC
CTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATA
CGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA
CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGC
TCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC
GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTC
ATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCC
GTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC
ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG
GTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA
AAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC
AGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAA
CGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAA
AAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTG
AGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGGACGTGGCCGACCAGCC
CGTCATCGTCAACGCCTGACCGCGGTGCGGACAGGCCGTGTCGCGACCGGCCGTGCGGAATTAAGCCGGCCC
GTACCCTGTGAATAGAGGTCCGCTGTGACACAAGAATCCCTGTTACTTCTCGACCGTATTGATTCGGATGATTC
CTACGCGAGCCTGCGGAACGACCAGGAATTCTGGGAGCCGCTGGCCCGCCGAGCCCTGGAGGAGCTCGGGCT
GCCGGTGCCGCCGGTGCTGCGGGTGCCCGGCGAGAGCACCAACCCCGTACTGGTCGGCGAGCCCGACCCGGT
GATCAAGCTGTTCGGCGAGCACTGGTGCGGTCCGGAGAGCCTCGCGTCGGAGTCGGAGGCGTACGCGGTCCT
GGCGGACGCCCCGGTGCCGGTGCCCCGCCTCCTCGGCCGCGGCGAGCTGCGGCCCGGCACCGGAGCCTGGCC
GTGGCCCTACCTGGTGATGAGCCGGATGACCGGCACCACCTGGCGGTCCGCGATGGACGGCACGACCGACCG
GAACGCGCTGCTCGCCCTGGCCCGCGAACTCGGCCGGGTGCTCGGCCGGCTGCACAGGGTGCCGCTGACCGG
GAACACCGTGCTCACCCCCCATTCCGAGGTCTTCCCGGAACTGCTGCGGGAACGCCGCGCGGCGACCGTCGAG
GACCACCGCGGGTGGGGCTACCTCTCGCCCCGGCTGCTGGACCGCCTGGAGGACTGGCTGCCGGACGTGGAC
ACGCTGCTGGCCGGCCGCGAACCCCGGTTCGTCCACGGCGACCTGCACGGGACCAACATCTTCGTGGACCTGG
CCGCGACCGAGGTCACCGGGATCGTCGACTTCACCGACGTCTATGCGGGAGACTCCCGCTACAGCCTGGTGCA
ACTGCATCTCAACGCCTTCCGGGGCGACCGCGAGATCCTGGCCGCGCTGCTCGACGGGGCGCAGTGGAAGCG
GACCGAGGACTTCGCCCGCGAACTGCTCGCCTTCACCTTCCTGCACGACTTCGAGGTGTTCGAGGAGACCCCG
CTGGATCTCTCCGGCTTCACCGATCCGGAGGAACTGGCGCAGTTCCTCTGGGGGCCGCCGGACACCGCCCCCG
GCGCCTGACGCCCCGGGCCGCCCGGCGCCGCCCCCGGCCCCCGGCGGCCGGCCGCCCGGAGCCCCGCCCGCG
CTCGGGAGCCCCGGGCCCGCGCCGAAGCCCGCTGCTGCGAGCCCGGAGCGGGCCGGCCGACGGCGGTGCGGG
CCCGGCGGCGGACGCTCAGCAGCGGCGGGCGTGAAAGGCCCTGGCATCCTCGATCATCTCCTCCAGGGTGGTC
GGCCGAGCTTCCATCCCAGCTCGGCAAGGATCAACGAGGACAGTCGCACGACGAAGTTCTTCTGGATCGCGCC
CGTGCTGGAAGCACTCAACCTCGAAGCGTGTGGTTGCGGAGCCATCTAGCAACCACACGAAACATGCGCAAC
GAACCGCGCAACGAACAACGCCTAGAACTGGCCCTAGATGAGCTGACTCGTATCGTTGGTAAACCTAGTTTGA
CCAGCATGTTTTAACTACGTTCGGTGAGCTGTCAACGGGGCCTGTAACGGCACAACGAACCGTGCAACGAGAG
TGGCCACGGATGCCACCACAGGCACTACAACGGAGTTCGCCACGTACATCACCACAACCACCGATTCTGGCGG
TGAGCTCCCCGATATTCAGCGGAAATGGCTTGGTATCGACCAAGATTCGTAGAACCCCGTCTCGTCTGGCTGG
TATTCAAAACGGGCGCAACGAAACACGCAACGAGACAGGCATGGCCCAAACCAGAAAACTAGCGTCTACCAG
GACTTTTACCTGTCCGACCCGTTGCAACGGAACCCCCCACGGAACCCCCGCGACACCCGCTCCCCAATTGCGT
TAGAACAGCGGTGGATTGTCGGCTTCGTTGTGGGCCTTTTGAGCCGCTTCCTGTTCTGCCGCACGCTCTTTCCT
CGCCCGATAGCCGAGTCGCTTAACGGTGTCCAGATGCAGCCCGAAATGTTTGGCCGTTTGCGGCCAAGAGTGG
CCCTCGTCGTCGTGATAGGCGCGGATGCGTTCGCGGCGTGCAGCCTGCTCGGCGAGCCACTCGCTGCGTTCCT
GCGCCACGAGCCGGACGACGTGGCGTTCGGATAGTCCGGTGATTCGAGCGCCTTCGGCGGCGGTCACGCGCC
GCTTTTTGCGGACAGTCGGCTGCCGGTTGTAGCCGTCGCTGTAGCCGTCGCTGTAGCCGTCGCTCATAGCAATG
CCTCCATGGCTGACGCGGACTTTGCGCGCCGCGCAACTGTGCTCGCCGCCGTGCGCGCTGCTGCGCCCTTCCGC
GAGATGGCCGACTGGCGCGCACTGAGTGTGGCCTCGTAGACCACGATCCCGTCCGCCCAAATGCGCGACTTGG
TTGTGATCCAACGCCAAATGCTGTTGGCGATGGCGCGGACCTCGCTGTCCGGTAGCGGTCCGGGACACACGTC
GTTGCACGGGAATTCGGCGTTTCGCGCGTGGCACTCGGCATAGATCGCGCGGCCGAGTCCGTCCACGTTCCGG
GTCGGCAGGTAGATCCGCATGAGGGCGGGACGATAGGCCCACAACCTGACGGAATCGAACAGTGCGCAATTC
CGCCCTAGCGGCGTCGGAGCCGCTTTGTACGTGGTCTGCTGACGCCAGCGCGGCGGTGGCATGTTCGCGCCGA
GCTCGGCCTCGATGTGGCTGAGTGTGTAGAGATCTGAGTGGAGCCATTCCGTTTCCCAGGCGATGTGGCCGGG
GTTTTTGGTCATGAGGCCTGAGTAACTGCGGTCGCCGTCGACGGCGCGCCGAAGGCCTTCGGCGCACGCCGCC
ATGTATGCGAGCGGCTTACGCCGCGCGTATTCGGTGCGTGGAACAGGGGCGTTGAGTGCCCACACTGCGTGTG
CGTGGCCGTTGGCGCGATTGCCCACGATCGCGTTGGGCAGCGGATGGGACCCCCGGGCGCTGAGCGCTCGGA

GCGCTGCGTCTGGATGGTCTACGTCCACGACCAGCAGGTTTGCCAGCGCTGTTGGGTTCGCCTCGATGTACCG
GCGGCCTAGGGCCGACGCGCGGCTTTGGCGGTAGATCCCCTCGAGCAGATCGTCGCTTGCCAGCGGCCAGTAC
GGCAGCCAGAGCTGCTCAAATTCGTCGGCGACGTGGCTCACGCTTGGTAGTAGACCACGATTAATCACCGGTG
TATGGTCCGACACGAGCTCCAAGTCAGATATTTCGCTGAGGGGCCACCCCACAACTGCACACTCCCCCGCTCT
CCCGTCGAGCCCTGGTGGTGGAACACCAGCGACAGCCGAGCACCCCCAACCACCTGTACCAACCAGGAGGAA
CACATGCGTCGTTTCGAGGACGTTTCCGGGCCGCTGAGAGCCGCTGTGGCGGCCGTACACGCCGCCTTAGACC
CGTTAGACCCCCTGCCGCCTGAATGCGCGGGTACGAGCCACACAGCGCCCGAACTTACGGAGCTGGTGGGCTC
ACCTGGCTTTATGGCGTACGAATCGGCTGTGTGCGACCTGTTGGGCGAGGTGAGGTACGCGCTACTCACGCTG
GCAAGGGCGACACAGCCGCCCCACCGAGCCCGCACGGCCGCGCGCGGTGTCAACAACCGGGTGAGTCGTGCA
CACCAGCAGGTGTTCGAGGCTTGGCTCGAAGTGCAGGACATCGTGGCGAACGCCGCCCGATGAGCCGCGCCTT
ACGCTGGCTGCCAGCCGTTCGCGGGCTGGTTGGTGCAGCGCGTCGAGCGGTTAGAGGCCCTGCGGTGTTCCAC
CACCGCAGGCCTCGCCCTTTTTAAGGCTGAATTTGCTTGTCTCCGAATCCAACTGGCTTGTCCAAGGGTGTATC
TACGCTTAATCCAAAGTTCAAACGAGGGGATTACACATGACCAACTTCGATAACGTTCTCGGCTCGATGATCC
CCGGCGGGCCCCCCCTCGAGGTCGACGGTATCGAT

pKMB55, pTE-2MIL with GFP and Ag85C



TTAATTAATATGCATCAATTGATTTAAATACTAGTGTTAACTATTTAATTGGGGACCCTAGAGGTCCCCTTTTTT
ATTTTAAAAATTTTTTCNAAGGGGACCTCTAGGGTCCCCAGCTGGCTAGCGAGTCATGAGGTNACCATGGTCA
TTAAGACCCACTTTCACATTTAAGTTGTTTTTCTAATCCGCATATGATCAATTCAAGGCCGAATAAGAAGGCTG
GCTCTGCACCTTGGTGATCAAATAATTCGATAGCTTGTCGTAATAATGGCGGCATACTATCAGTAGTAGGTGTT
TCCCTTTCTTCTTTAGCGACTTGATGCTCTTGATCTTCCAATACGCAACCTAAGGTAAAATGCCCCACAGCGCT
GAGTGCATATAACGCGTTCTCTAGTGAAAAACCTTGTTGGCATAAAAAGGCTAATTGATTTTCGAGAGTTTCA
TACTGTTTTTCTGTAGGCCGTGTACCTAAATGTACTTTTGCTCCATCGCGATGACTTAGTAAAGCACATCTAAA
ACTTTTAGCGTTATTACGTAAAAAATCTTGCCAGCTTTCCCCTTCTAAAGGGCAAAAGTGAGTATGGTGCCTAT
CTAACATCTCAATGGCTAAGGCGTCGAGCAGGGCCCGCTTGTTCTTGACGTGCCAGTACAGGGTCGGCTGCTC
GACGCCCAGCTTCTGGGCCAGCTTGCGGGTGGTCAGGCCCTCGATGCCGACCTCGTTCAGCAGCTCCAGGGCG
GAGTTGATGACCTTGGACTTGTCCAGGCGGGACATTGCGAAGTGATTCCTCCTGATGGTACCGGATCCGTGGC
AGGAGGGTCGCCTGCGTCGGCAACTCGCCCAGGCTACACGAGAGCCCTCGACTGCGAAAATGCCGACCAGGC
ACCACCGCACTCCAGGCTTAACCCGGTGCGACGATCCAATATTACGAGATCTACTAGAAATATTGGATCGTCG

GCACCGTCACGGCCGTGGGAGGCGGCACGATCCGCGACGTGATGATCGGCCGCATCCCCACGGTGCTGCGCA
GTGAGCTCTACGCCATCCCGGCGTTGATCTGTGCGTTCGCACGCACAGGCCCGGTGTGAGAAGGGTCTCTGAC
GAGCGGGAGAACTCCCTATCAGTGATAGAGTTTGTCCTCCCTATCAGTGATAGATAGGCTCTGGGAGTACCCG
TGTGTACGACCAGCACGGCATACATCATTTCGACGCCGAGAGATTCGCCGCCCGAAATGAGCACGATCCGCTC
GAGGCATGCGAAGGAGGTATACAT**ATGACGTTCTTCGAACAGGTGCGAAGGTTGCGGAGCGCAGCGACA**
ACCCTGCCGCGCCGGCTGGCTATCGCGGCTATGGGGGCTGTCCTGGTTTACGGTCTGGTCGGTACCTT
CGGCGGGCCGGCCACCGCGGGCGCATTCTCTAGGCCCGGTCTTCCAGTGGAATATCTGCAGGTGCCAT
CCGCGTCGATGGGCCGCGACATCAAGGTCCAGTTCCAGGGCGGCGGACCGCACGCGGTCTACCTGCTC
GACGGTCTGCGGGCCCAGGATGACTACAACGGCTGGGACATCAACACCCCGGCCTTCGAGGAGTACTA
CCAGTCAGGGTTGTCGGTGATCATGCCCGTGGGCGGCCAATCCAGTTTCTACACCGACTGGTATCAGCC
CTCGCAGAGCAACGGCCAGAACTACACCTACAAGTGGGAGACCTTCCTTACCAGAGAGATGCCCGCCT
GGCTACAGGCCAACAAGGGCGTGTCCCCGACAGGCAACGCGGCGGTGGGTCTTTCGATGTCGGGCGGT
TCCGCGCTGATCCTGGCCGCGTACTACCCGCAGCAGTTCCCGTACGCCGCGTCGTTGTCGGGCTTCCTC
AACCCGTCCGAGGGCTGGTGGCCGACGCTGATCGGCCTGGCGATGAACGACTCGGGCGGTTACAACGC
CAACAGCATGTGGGGTCCGTCCAGCGACCCGGCCTGGAAGCGCAACGACCCAATGGTTCAGATTCCCC
GCCTGGTCGCCAACAACACCCGGATCTGGGTGTACTGCGGTAACGGCACACCCAGCGACCTCGGCGGC
GACAACATACCGGCGAAGTTCCTGGAAGGCCTCACCCTGCGCACCAACCAGACCTTCCGGGACACCTA
CGCGGCCGACGGTGGACGCAACGGGGTGTTTAACTTCCCGCCCAACGGAACACACTCGTGGCCCTACT
GGAACGAGCAGCTGGTCGCCATGAAGGCCGATATCCAGCATGTGCTCAACGGCGCGACACCCCCGGCC
GCCCCTGCTGCGCCGGCCGCCCCCATCCATATGAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCA
ATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGAT
GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACAC
TTGTCACTACTTTCGGTTATGGTGTTCAATGCTTTGCGAGATACCCAGATCATATGAAACAGCATGACTT
TTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTA
CAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGA
TTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTATACATC
ATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGC
GTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAAC
CATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTT
GAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAATAAAGCGGCCGCAAACAGT
ATTAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAGCTAATCAACTAGTTCTAGCTGCATTAATG
AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTG
CGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAG
GGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTG
CTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA
ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTG
CCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTA
TCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG
TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC
ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTT
GATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAA
AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGG
ATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT
CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATC
TGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGGACGTGGCCGACCAGCCCGTCATCGTCAACGCCTGAC
CGCGGTGCGGACAGGCCGTGTCGCGACCGGCCGTGCGGAATTAAGCCGGCCCGTACCCTGTGAATAGAGGTC
CGCTGTGACACAAGAATCCCTGTTACTTCTCGACCGTATTGATTCGGATGATTCCTACGCGAGCCTGCGGAAC
GACCAGGAATTCTGGGAGCCGCTGGCCCGCCGAGCCCTGGAGGAGCTCGGGCTGCCGGTGCCGCCGGTGCTG
CGGGTGCCCGGCGAGAGCACCAACCCCGTACTGGTCGGCGAGCCCGACCCGGTGATCAAGCTGTTCGGCGAG
CACTGGTGCGGTCCGGAGAGCCTCGCGTCGGAGTCGGAGGCGTACGCGGTCCTGGCGGACGCCCCGGTGCCG
GTGCCCCGCCTCCTCGGCCGCGGCGAGCTGCGGCCCGGCACCGGAGCCTGGCCGTGGCCCTACCTGGTGATGA
GCCGGATGACCGGCACCACCTGGCGGTCCGCGATGGACGGCACGACCGACCGGAACGCGCTGCTCGCCCTGG
CCCGCGAACTCGGCCGGGTGCTCGGCCGGCTGCACAGGGTGCCGCTGACCGGGAACACCGTGCTCACCCCCCA
TTCCGAGGTCTTCCCGGAACTGCTGCGGGAACGCCGCGCGGCGACCGTCGAGGACCACCGCGGGTGGGGCTA
CCTCTCGCCCCGGCTGCTGGACCGCCTGGAGGACTGGCTGCCGGACGTGGACACGCTGCTGGCCGGCCGCGAA
CCCCGGTTCGTCCACGGCGACCTGCACGGGACCAACATCTTCGTGGACCTGGCCGCGACCGAGGTCACCGGGA
TCGTCGACTTCACCGACGTCTATGCGGGAGACTCCCGCTACAGCCTGGTGCAACTGCATCTCAACGCCTTCCG
GGGCGACCGCGAGATCCTGGCCGCGCTGCTCGACGGGGCGCAGTGGAAGCGGACCGAGGACTTCGCCCGCGA
ACTGCTCGCCTTCACCTTCCTGCACGACTTCGAGGTGTTCGAGGAGACCCCGCTGGATCTCTCCGGCTTCACCG
ATCCGGAGGAACTGGCGCAGTTCCTCTGGGGGCCGCCGGACACCGCCCCCGGCGCCTGACGCCCCGGGCCGC
CCGGCGCCGCCCCCGGCCCCCGGCGGCCGGCCGCCCGGAGCCCCGCCCGCGCTCGGGAGCCCCGGGCCCGCG
CCGAAGCCCGCTGCTGCGAGCCCGGAGCGGGCCGGCCGACGGCGGTGCGGGCCCGGCGGCGGACGCTCAGCA
GCGGCGGGCGTGAAAGGCCCTGGCATCCTCGATCATCTCCTCCAGGGTGGTCGGCCGAGCTTCCATCCCAGCT
CGGCAAGGATCAACGAGGACAGTCGCACGACGAAGTTCTTCTGGATCGCGCCCGTGCTGGAAGCACTCAACC
TCGAAGCGTGTGGTTGCGGAGCCATCTAGCAACCACACGAAACATGCGCAACGAACCGCGCAACGAACAACG
CCTAGAACTGGCCCTAGATGAGCTGACTCGTATCGTTGGTAAACCTAGTTTGACCAGCATGTTTTAACTACGTT

CGGTGAGCTGTCAACGGGGCCTGTAACGGCACAACGAACCGTGCAACGAGAGTGGCCACGGATGCCACCACA
GGCACTACAACGGAGTTCGCCACGTACATCACCACAACCACCGATTCTGGCGGTGAGCTCCCCGATATTCAGC
GGAAATGGCTTGGTATCGACCAAGATTCGTAGAACCCCGTCTCGTCTGGCTGGTATTCAAAACGGGCGCAACG
AAACACGCAACGAGACAGGCATGGCCCAAACCAGAAAACTAGCGTCTACCAGGACTTTTACCTGTCCGACCC
GTTGCAACGGAACCCCCCACGGAACCCCCGCGACACCCGCTCCCCAATTGCGTTAGAACAGCGGTGGATTGTC
GGCTTCGTTGTGGGCCTTTTGAGCCGCTTCCTGTTCTGCCGCACGCTCTTTCCTCGCCCGATAGCCGAGTCGCTT
AACGGTGTCCAGATGCAGCCCGAAATGTTTGGCCGTTTGCGGCCAAGAGTGGCCCTCGTCGTCGTGATAGGCG
CGGATGCGTTCGCGGCGTGCAGCCTGCTCGGCGAGCCACTCGCTGCGTTCCTGCGCCACGAGCCGGACGACGT
GGCGTTCGGATAGTCCGGTGATTCGAGCGCCTTCGGCGGCGGTCACGCGCCGCTTTTTGCGGACAGTCGGCTG
CCGGTTGTAGCCGTCGCTGTAGCCGTCGCTGTAGCCGTCGCTCATAGCAATGCCTCCATGGCTGACGCGGACTT
TGCGCGCCCGCGCAACTGTGCTCGCCGCCGTGCGCGCTGCTGCGCCCTTCCGCGAGATGGCCGACTGGCGCGCA
CTGAGTGTGGCCTCGTAGACCACGATCCCGTCCGCCCAAATGCGCGACTTGGTTGTGATCCAACGCCAAATGC
TGTTGGCGATGGCGCGGACCTCGCTGTCCGGTAGCGGTCCGGGACACACGTCGTTGCACGGGAATTCGGCGTT
TCGCGCGTGGCACTCGGCATAGATCGCGCGGCCGAGTCCGTCCACGTTCCGGGTCGGCAGGTAGATCCGCATG
AGGGCGGGACGATAGGCCCACAACCTGACGGAATCGAACAGTGCGCAATTCCGCCCTAGCGGCGTCGGAGCC
GCTTTGTACGTGGTCTGCTGACGCCAGCGCGGCGGTGGCATGTTCGCGCCGAGCTCGGCCTCGATGTGGCTGA
GTGTGTAGAGATCTGAGTGGAGCCATTCCGTTTCCCAGGCGATGTGGCCGGGGTTTTTGGTCATGAGGCCTGA
GTAACTGCGGTCGCCGTCGACGGCGCGCCGAAGGCCTTCGGCGCACGCCGCCATGTATGCGAGCGGCTTACGC
CGCGCGTATTCGGTGCGTGGAACAGGGGCGTTGAGTGCCCACACTGCGTGTGCGTGGCCGTTGGCGCGATTGC
CCACGATCGCGTTGGGCAGCGGATGGGACCCCCGGGCGCTGAGCGCTCGGAGCGCTGCGTCTGGATGGTCTAC
GTCCACGACCAGCAGGTTTGCCAGCGCTGTTGGGTTCGCCTCGATGTACCGGCGGCCTAGGGCCGACGCGCGG
CTTTGGCGGTAGATCCCCTCGAGCAGATCGTCGCTTGCCAGCGGCCAGTACGGCAGCCAGAGCTGCTCAAATT
CGTCGGCGACGTGGCTCACGCTTGGTAGTAGACCACGATTAATCACCGGTGTATGGTCCGACACGAGCTCCAA
GTCAGATATTTCGCTGAGGGGCCACCCCACAACTGCACACTCCCCCGCTCTCCCGTCGAGCCCTGGTGGTGGA
ACACCAGCGACAGCCGAGCACCCCCAACCACCTGTACCAACCAGGAGGAACACATGCGTCGTTTCGAGGACG
TTTCCGGGCCGCTGAGAGCCGCTGTGGCGGCCGTACACGCCGCCTTAGACCCGTTAGACCCCCTGCCGCCTGA
ATGCGCGGGTACGAGCCACACAGCGCCCGAACTTACGGAGCTGGTGGGCTCACCTGGCTTTATGGCGTACGAA
TCGGCTGTGTGCGACCTGTTGGGCGAGGTGAGGTACGCGCTACTCACGCTGGCAAGGGCGACACAGCCGCCCC
ACCGAGCCCGCACGGCCGCGCGCGGTGTCAACAACCGGGTGAGTCGTGCACACCAGCAGGTGTTCGAGGCTT
GGCTCGAAGTGCAGGACATCGTGGCGAACGCCGCCCGATGAGCCGCGCCTTACGCTGGCTGCCAGCCGTTCGC
GGGCTGGTTGGTGCAGCGCGTCGAGCGGTTAGAGGCCCTGCGGTGTTCCACCACCGCAGGCCTCGCCCTTTTT
AAGGCTGAATTTGCTTGTCTCCGAATCCAACTGGCTTGTCCAAGGGTGTATCTACGCTTAATCCAAAGTTCAAA
CGAGGGGATTACACATGACCAACTTCGATAACGTTCTCGGCTCGATGATCCCCGGCGGGCCCCCCCTCGAGGT
CGACGGTATCGAT

9.34 References

1. Boucau, J., Sanki, A.K., Voss, B.J., Sucheck, S.J. & Ronning, D.R. A coupled assay measuring *Mycobacterium tuberculosis* antigen 85C enzymatic activity. *Anal. Biochem.* **385**, 120-127 (2009).
2. Geoghegan, K.F. et al. Spontaneous α-N-6-Phosphogluconoylation of a "His Tag" inEscherichia coli:The Cause of Extra Mass of 258 or 178 Da in Fusion Proteins. *Anal. Biochem.* **267**, 169-184 (1999).
3. Barry, C.S., Backus, K.M., Barry, C.E. & Davis, B.G. ESI-MS Assay of *M. tuberculosis* Cell Wall Antigen 85 Enzymes Permits Substrate Profiling and Design of a Mechanism-Based Inhibitor. *J. Am. Chem. Soc.*, ASAP (2011).
4. Domenech, P., Reed, M.B. & Barry, C.E., 3rd. Contribution of the *Mycobacterium tuberculosis* MmpL protein family to virulence and drug resistance. *Infect. Immun.* **73**, 3492-501 (2005).
5. Slayden, R.A. & Barry, C.E. Analysis of the Lipids of *Mycobacterium tuberculosis*. Vol. 54 229-245 (2001).
6. Celada, A., Gray, P.W., Rinderknecht, E. & Schreiber, R.D. Evidence for a gamma-interferon receptor that regulates macrophage tumoricidal activity. *J. Exp. Med.* **160**, 55-74 (1984).
7. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33-8, 27-8 (1996).
8. Phillips, J.C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781-802 (2005).
9. Brooks, B.R. et al. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**, 1545-614 (2009).
10. Boshoff, H.I.M. et al. The Transcriptional Responses of *Mycobacterium tuberculosis* to Inhibitors of Metabolism. *J. Biol. Chem.* **279**, 40174-40184 (2004).
11. Parish, T. & Stoker, N.G. Electroporation of Mycobacteria. in *Mycobacteria Protocols*, Vol. 101 129-144 (Humana Press, 1998).