

**Next generation approaches to polysaccharide preparation
for *Burkholderia pseudomallei* vaccine development.**

Submitted by Victoria Mae Baldwin, to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Biological Sciences, September 2016.

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

(Signature)

Burkholderia pseudomallei is the aetiological agent of melioidosis and a potential bioterror threat. Infections are difficult to treat due to extensive antibiotic resistance and there is no prophylactic vaccine available. Studies have shown that the capsular polysaccharide (CPS) of *B. pseudomallei* is a virulence factor, immunogen and candidate antigen for a glycoconjugate vaccine. However, polysaccharides are complex to synthesise. One approach is to genetically engineer *Escherichia coli* to express the CPS; however, previous attempts at cloning the CPS coding locus from *B. pseudomallei* into *E. coli* were unsuccessful. This project proposes to clone only the essential genes from *B. pseudomallei* and to use native *E. coli* mechanisms to complete CPS synthesis. This would contribute to development of a new platform for the expression of any bespoke polysaccharide in *E. coli*.

Six biosynthetic genes for the nucleotide sugar precursor were successfully expressed in *E. coli*. The structure of the precursor was verified by mass spectrometry. Precursor synthesis was also performed in an *in vitro* microfluidics system. This minimised the quantity of substrates and enzymes required, in preparation for the characterisation of glycosyltransferases required for CPS assembly. A novel assay for characterising glycosyltransferase activity was also developed, as current available options are prohibitively expensive and require significant quantities of glycosyltransferase which are difficult to purify. Finally, plasmids for the expression of additional glycosyltransferases to link the nascent *B. pseudomallei* CPS to truncated polysaccharides in *E. coli* were constructed.

The aim of this project was to contribute to the development of a platform for the expression of bespoke polysaccharides in *E. coli*. The CPS of *B. pseudomallei* was chosen as the model polysaccharide as it has a simple structure and its manufacture is desirable for use in a vaccine against melioidosis.

Table of contents

Abstract	1
Table of contents	2
List of figures	10
List of tables	15
List of presentations and publications	17
Author's declaration	18
Acknowledgements	19
List of abbreviations	20
Chapter 1	23
Chapter 2	90
Chapter 3	156
Chapter 4	180
Chapter 5	246
Chapter 6	279
References	287
Appendices	321

Chapter 1 – Introduction 23

<u>1.1 <i>Burkholderia pseudomallei</i> and melioidosis</u>	24
1.1.1 Phylogeny and genome	24
1.1.2 Epidemiology of <i>B. pseudomallei</i>	26
1.1.2.1 Distribution in the environment	26
1.1.2.2 Global distribution	27
1.1.3 <i>B. pseudomallei</i> infection in humans	29
1.1.3.1 Route of transmission	29
1.1.3.2 Risk factors for infection in humans	31
1.1.4 Melioidosis	33
1.1.4.1 Clinical symptoms	33
1.1.4.2 Diagnosis	34
1.1.4.3 Treatment	35
1.1.4.4 Mortality rates	36

1.1.5 Virulence factors	36
1.1.5.1 Intracellular survival	36
1.1.5.2 Flagellar motility	39
1.1.5.3 Host cell adhesion	39
1.1.5.4 Secretion systems	41
1.1.5.6 Polysaccharides	43
1.1.5.7 Antibiotic resistance	43
1.1.5.8 Latent, persistent and chronic infections	44
1.1.5.9 Other virulence factors	45
<u>1.2 Polysaccharides in <i>B. pseudomallei</i></u>	46
1.2.1 Capsular polysaccharide	47
1.2.1.1 Structure	47
1.2.1.2 Function	50
1.2.1.3 Coding locus	52
1.2.1.4 GDP-6d _B Hep biosynthetic pathway	53
1.2.1.5 Glycosyltransferase-mediated assembly of CPS	55
1.2.1.6 Capsule assembly	56
1.2.2 Lipopolysaccharide	57
1.2.3 Minor exopolysaccharides	60
<u>1.3 Vaccine development</u>	62
1.3.1 Animal models of infection	63
1.3.2 Routes for vaccine administration	64
1.3.3 Host immune response	65
1.3.3.1 Innate immunity	65
1.3.3.2 Adaptive immune system	67
1.3.4 Immunisation strategies	72
1.3.4.1 Whole cell vaccination	72
1.3.4.2 Subunit vaccines	74
<u>1.4 Engineering <i>B. pseudomallei</i> CPS in <i>E. coli</i></u>	84
1.4.1 Synthetic biology for polysaccharides	84
1.4.2 Operon assembly for biosynthesis of <i>B. pseudomallei</i> CPS precursor .	86
1.4.1.2 BioBricks™ 3A assembly method	88
<u>1.5 Project aims</u>	89

<u>Chapter 2 - Purification of biosynthetic GDP-6-deoxy-α-D-manno-heptose from genetically engineered <i>E. coli</i></u>	90
<u>2.1 Introduction</u>	91
2.1.1 Aims	92
<u>2.2 Materials and methods</u>	93
2.2.1 Bacterial strains and reagents	93
2.2.2 Growth, storage and competent cell preparation of <i>E. coli</i>	93
2.2.2.1 Growth and storage conditions	93
2.2.2.2 Preparation of chemically competent cells	93
2.2.3 Plasmid construction for GDP-6ddHep biosynthesis	94
2.2.3.1 Gene synthesis	94
2.2.3.2 Plasmid transformation	95
2.2.3.3 Plasmid purification	95
2.2.3.4 Polymerase chain reaction	95
2.2.3.5 Restriction digests	96
2.2.3.6 Ligations	96
2.2.3.7 Nucleic acid visualisation	97
2.2.3.8 Subcloning into a BioBricks™ vector	97
2.2.3.9 BioBricks™ 3A assembly	97
2.2.4 Verification of plasmid expression	98
2.2.4.1 <i>E. coli</i> growth curve	98
2.2.4.2 Plasmid sequencing	99
2.2.4.3 RNA extraction	99
2.2.4.4 Reverse transcriptase PCR	99
2.2.4.5 Small volume protein extraction	100
2.2.4.6 Purification of c-Myc-tagged proteins	100
2.2.4.7 Sodium dodecyl sulphate polyacrylamide gel electrophoresis	100
2.2.4.8 Blue native-PAGE	101
2.2.4.9 Second dimension-PAGE	102
2.2.4.10 Western blot	103
2.2.5 Total metabolite extraction from <i>E. coli</i>	104
2.2.6 GDP-6ddHep identification by mass spectrometry	105
2.2.6.1 Liquid chromatography for mass spectrometry	105
2.2.6.2 Untargeted LC-MS metabolite profiling (Q-TOF)	105

2.2.7 Purification of GDP-6dHep	106
2.2.7.1 Extraction of GDP-6dHep	106
2.2.7.2 Purification of GDP-6dHep by HPLC	107
2.2.8 Verification of GDP-6dHep structure	107
2.2.8.1 Sample purity analysis (Q-TOF)	107
2.2.8.2 Targeted LC-MS metabolite screening (QQQ)	107
2.2.8.3 Gas chromatography-MS structural analysis	108
2.2.8.4 Nuclear magnetic resonance	110
2.3 Results	111
2.3.1 Plasmid construction for GDP-6dHep biosynthesis	111
2.3.1.1 Synthetic genes for plasmid construction	111
2.3.1.2 Assembly of GDP-6dHep biosynthesis plasmid	113
2.3.1.3 Terminator and promoter sequences	117
2.3.2 Verification of plasmid expression	119
2.3.2.1 Plasmid DNA sequence	119
2.3.2.2 RNA expression (transcription)	122
2.3.2.3 Protein expression (translation)	125
2.3.2.4 <i>E. coli</i> growth curve	127
2.3.3 Total metabolite extraction	128
2.3.4 GDP-6dHep purification	130
2.3.5 GDP-6dHep structure verification	131
2.3.5.1 LC-MS Q-TOF	131
2.3.5.2 LC-MS QQQ	133
2.3.5.3 GC-MS	137
2.3.5.4 NMR	142
2.4 Discussion	150
2.4.1 Construction of GDP-6d _b Hep biosynthetic operon	150
2.4.1.1 The problem of WcbM (part 1)	152
2.4.1.2 Operon optimisation to improve GDP-6d _b Hep yield	153
2.4.1.3 Alternative reasons for low GDP-6d _b Hep yield	154
2.4.2 GDP-6d _b Hep purification	155
Chapter 3 - Synthesis of GDP-6-deoxy-α-D-manno-heptose <i>in vitro</i>	156
3.1 Introduction	157
3.1.1 Aims	158

3.2 Materials and methods	159
3.2.1 Large scale protein purification	159
3.2.1.1 Isopropyl β -D-1-thiogalactopyranoside induction	159
3.2.1.2 ZYM-5052 auto-induction	159
3.2.1.3 Harvesting cells	159
3.2.1.4 Cell lysis	160
3.2.2 Purification of proteins for <i>B. pseudomallei</i> CPS synthesis	160
3.2.2.1 Nickel affinity and size exclusion chromatography	160
3.2.2.2 Gravity flow nickel affinity chromatography	161
3.2.2.3 Quality control, concentration and storage of proteins	161
3.2.3 Analysis of <i>B. pseudomallei</i> CPS biosynthetic pathway <i>in vitro</i>	162
3.2.3.1 Pi ColorLock™ Gold colourimetric kit	162
3.2.3.2 Adapted Pi ColorLock™ Gold colourimetric kit	163
3.2.3.3 Kinetic NADPH oxidation assay	163
3.2.3.4 DTNB assay for WcbI activity	164
3.2.4 Synthesis of GDP-6-deoxy- α -D-manno-heptose in a microfluidics system	165
3.2.4.1 Microfluidics assay	165
3.2.4.2 Preparation of samples for targeted LC-MS (QQQ) analysis	166
3.3 Results	167
3.3.1 Purification of the biosynthetic enzymes	167
3.3.2 GDP-6dHep biosynthetic pathway <i>in vitro</i>	168
3.3.2.1 GmhA, WcbL and WcbN activity	168
3.3.2.2 Addition of WcbM	169
3.3.2.3 Addition of WcbK and WcbJ	172
3.3.3 GDP-6dHep biosynthetic pathway in a microfluidics system	173
3.4 Discussion	176
3.4.1. Microfluidics as a novel system for conducting enzyme assays <i>in vitro</i>	176
3.4.1.1 Droplet formation	176
3.4.1.2 Droplet structure and stability	177
3.4.1.3 The problem of WcbM (part 2)	178

<u>Chapter 4 - Development of a novel assay for characterisation of glycosyltransferases</u>	180
<u>4.1 Introduction</u>	181
4.1.1 Discontinuous kinase assay	182
4.1.1.1 Mesophilic Cmk and Gmk from <i>E. coli</i>	183
4.1.1.2 Thermophilic Cmk, Gmk, Tmk and Umpk from <i>T. thermophilus</i>	184
4.1.2 Discontinuous phosphatase assay using an apyrase	186
4.1.3 β -1,4-Galactosyltransferase and LgtC	187
4.1.1 Aims	188
<u>4.2 Materials and methods</u>	189
4.2.1 Preparation of kinases for assay development	189
4.2.1.1 Cloning genes for mesophilic kinases	189
4.2.1.2 Gene synthesis for thermostable kinases	189
4.2.1.3 Purification of mesophilic kinases	190
4.2.1.4 Purification of thermophilic kinases	190
4.2.2 Kinase activity assay	191
4.2.3 Characterisation of β -galactosyltransferase and LgtC	191
4.2.3.1 Coupled discontinuous kinase assay	191
4.2.3.2 Coupled discontinuous phosphatase assay	192
4.2.3.3 LgtC characterisation	192
4.2.5 Data Analysis	193
4.2.5.1 Michaelis-Menten and substrate inhibition analyses	193
4.2.6 Fluorescent antibody assay	194
4.2.7 Apyrase purification	195
4.2.7.1 Extraction from <i>E. coli</i>	195
4.2.7.2 Denaturation	195
4.2.7.3 Dialysis	195
4.2.7.4 TCA precipitation	197
4.2.7.5 Nickel affinity purification	197
<u>4.3 Results</u>	199
4.3.1 Kinase purification	199
4.3.1.1 Mesophilic kinases	199
4.3.1.2 Thermophilic kinases	203

4.3.2 Kinase assay optimisation	211
4.3.2.1 ADP concentration	211
4.3.2.2 Kinase concentration	212
4.3.2.3 Assay range (NDP concentration)	214
4.3.3 Kinase activity by ATP production	216
4.3.4 Kinase standard curves for coupled assay development	219
4.3.4.1 Mesophilic kinases	219
4.3.4.2 Thermophilic kinases	220
4.3.4.3 Comparison of Cmk	221
4.3.4.4 Comparing mesophilic and thermophilic kinases at 37 °C	222
4.3.5 Characterisation of β 14GalT	224
4.3.5.1 β 14GalT concentration	224
4.3.5.2 Donor substrate (UDP-Gal) concentration	225
4.3.5.3 Acceptor substrate (GlcNAc) concentration	226
4.3.6 Determining kinetic parameters for β 14GalT	228
4.3.6.1 Novel coupled kinase assay	228
4.3.6.2 Commercial phosphatase assay	229
4.3.7 Determining kinetic parameters for LgtC	230
4.3.7.1 Novel coupled kinase assay	230
4.3.7.2 Commercial phosphatase assay	232
4.3.8 Commercial fluorescence assay	233
4.3.9 Apyrase purification	233
4.4 Discussion	237
4.4.1 Methods for glycosyltransferase characterisation	237
4.4.2 Novel coupled-kinase assay	238
4.4.3 Purification of kinases	239
4.4.4 Kinase activity	241
4.4.5 Characterisation of glycosyltransferases	243
4.4.6 Apyrase purification	244
<u>Chapter 5 - Polysaccharide synthesis: glycosyltransferase purification and plasmid construction</u>	246
<u>5.1 Introduction</u>	247
5.1.1 Aims	251
<u>5.2 Materials and methods</u>	254

5.2.1 Purification of glycosyltransferases	254
5.2.1.1 Gene synthesis	254
5.2.1.2 Bacterial cultures and detergents	254
5.2.1.3 Mini nickel column	255
5.2.1.4 Äkta and gravity flow nickel affinity chromatography	256
5.2.1.5 Differential Scanning Fluorimetry	256
5.2.2 Additional biosynthetic constructs	256
5.2.2.1 Gene synthesis for co-expression studies	256
5.2.2.2 BioBricks™ assembly of additional plasmids	257
5.2.2.3 Co-expression in <i>E. coli</i>	257
5.2.2.4 Enzyme-linked immunosorbent assay (ELISA)	258
5.2.2.5 Immunoblot assay	260
<u>5.3 Results</u>	262
5.3.1 Purification of glycosyltransferases	262
5.3.1.1. High-throughput screening for purification conditions of WcbH	262
5.3.1.2 Gravity flow column purification	264
5.3.1.3 Purification by size exclusion chromatography	266
5.3.1.4 Protein stability analysed by differential scanning fluorimetry	267
5.3.2 Additional plasmid construction	268
5.3.2.1 <i>B. pseudomallei</i> glycosyltransferases	268
5.3.2.2 Additional glycosyltransferases	269
5.3.3 Immunoblot and ELISA	270
<u>5.4 Discussion</u>	272
5.4.1 Glycosyltransferase purification	272
5.4.2 Glycosyltransferase function from bioinformatics	273
5.4.3 Assembly of CPS on truncated <i>E. coli</i> polysaccharides	275
5.4.4 Alternative methods for CPS biosynthesis for vaccine production	276
<u>Chapter 6 – Final discussion and conclusions</u>	279
6.1 Project background	280
6.2 Biosynthesis of GDP-6d _B Hep	281
6.3 Characterising GDP-6d _B Hep biosynthesis <i>in vitro</i>	282
6.4 A novel assay for glycosyltransferase activity	283
6.5 Exploiting native <i>E. coli</i> polysaccharide biosynthesis	284
6.6 Final conclusion	286

List of figures

Chapter 1

Figure 1	<i>Burkholderia</i> phylogenetic tree	25
Figure 2	Global distribution of <i>Burkholderia pseudomallei</i> and incidence of melioidosis	28
Figure 3	<i>B. pseudomallei</i> colony morphologies on Ashdown selective media	34
Figure 4	Intracellular lifecycle of <i>B. pseudomallei</i>	37
Figure 5	Persister cell tolerance to antibiotics	45
Figure 6	Structure of the repeated <i>manno</i> -heptose in <i>B. pseudomallei</i> CPS	48
Figure 7	Structure of <i>B. pseudomallei</i> CPS	49
Figure 8	Immunogold electron microscopy of CPS expression in <i>B. pseudomallei</i> 1026b	50
Figure 9	CPS coding locus in <i>B. pseudomallei</i>	53
Figure 10	Biosynthetic pathway for GDP-6 _D Hep	54
Figure 11	Proposed model for the assembly of CPS on the surface of <i>B. pseudomallei</i>	57
Figure 12	Generic structure of LPS	58
Figure 13	O-antigen of <i>B. pseudomallei</i> LPS	59
Figure 14	Basic structure of the five major antibody classes	69
Figure 15	Avici model for the mechanism of T-cell stimulation by glycoconjugate vaccines	79
Figure 16	Outer membrane vesicle formation	82
Figure 17	Enzymatic (A) and chemical (B) synthesis of oligosaccharides from <i>Neisseria meningitidis</i> .	85

Chapter 2

Figure 18	Synthetic gene constructs for BioBricks™ 3A assembly	94
Figure 19	BioBricks™ 3A assembly protocol	98
Figure 20	pEX plasmids containing the six CPS biosynthetic genes	112
Figure 21	Addition of <i>wcbL</i> upstream of <i>gmhA-wcbN</i> in pSB1A3	113
Figure 22	Confirmation of assembly of <i>wcbM-wcbL-gmhA-wcbN</i> in pSB1C3	114
Figure 23	PCR of <i>wcbM-wcbL-gmhA-wcbN_pSB1C3</i>	115
Figure 24	Progressive construction of the plasmid for GDP-6 _D Hep biosynthetic operon	116
Figure 25	Insertion of BioBricks™ promoter BBa_J23118 into <i>wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_pSB1C3</i>	117
Figure 26	PCR of completed GDP-6 _D Hep biosynthetic operon in BioBricks™ vector pSB1C3	118
Figure 27	Final plasmid construct for GDP-6 _D Hep biosynthesis	119
Figure 28	GDP-6 _D Hep biosynthesis plasmid with primer position and direction for sequencing indicated	120

Figure 29	Sequencing reads for GDP-6d _D Hep biosynthesis plasmid	121
Figure 30	Sequencing data showing an 88 bp deletion and 54 bp insertion in the GDP-6d _D Hep biosynthesis plasmid	122
Figure 31	Reverse-transcriptase PCR of the six genes on the GDP-6d _D Hep biosynthesis plasmid	124
Figure 32	SDS-PAGE and Western blot of c-Myc-tagged CPS biosynthetic proteins	125
Figure 33	2D-PAGE of CPS biosynthetic proteins	126
Figure 34	Anti-c-Myc Western blot of 2D-PAGE of CPS biosynthetic proteins	127
Figure 35	Effect of GDP-6d _D Hep biosynthesis plasmid expression on growth of <i>E. coli</i> .	128
Figure 36	Anion exchange chromatography for the purification of GDP-6d _D Hep.	130
Figure 37	HPLC of GDP-6d _D Hep	131
Figure 38	Accuracy score for GDP-6d _D Hep from Q-TOF scan in negative mode	132
Figure 39	Q-TOF scan of GDP-6d _D Hep sample to assess purity	133
Figure 40	MRM on QQQ LC-MS of purified GDP-6d _D Hep	134
Figure 41	MRM on QQQ LC-MS of purified extract from <i>E. coli</i> expressing empty pSB1C3 vector.	135
Figure 42	Overlay of MRM spectra for purified GDP-6d _D Hep and pSB1C3 empty vector	136
Figure 43	Overlay of MRM for GDP-mannose in pSB1C3 empty vector extract	137
Figure 44	Spectra of 219.1230 fragment scan in GC-MS samples and standards	139
Figure 45	Spectra of 333.1738 fragment scan in GC-MS samples and standards	140
Figure 46	Extracted ion chromatograms for GC-MS of GDP-6d _D Hep	141
Figure 47	Overlays of fragments scanned for in GC-MS of GDP-6d _D Hep	142
Figure 48	structures of GDP-6d _D Hep and five standards analysed by ¹ H and ³¹ P NMR	143
Figure 49	Spectra for ³¹ P NMR of GDP-mannose standard and GDP-6d _D Hep	144
Figure 50	Spectra for ¹ H NMR of the GDP-mannose standard and GDP-6d _D Hep	145
Figure 51	Spectra for ¹ H NMR of TEAA in D ₂ O and GDP-6d _D Hep	146
Figure 52	Spectra for ¹ H NMR of _D -manno-heptose standard and GDP-6d _D Hep	147
Figure 53	Spectra for ¹ H NMR of methyl-6-deoxy- _D -manno-heptoside standard and GDP-6d _D Hep	148
Figure 54	Spectra for ¹ H NMR of _D -manno-heptose standard and GDP-6d _D Hep	149

Chapter 3

Figure 55	Detection of GmhA, WcbL and WcbN activity	162
Figure 56	Detection of GmhA, WcbL, WcbN and WcbM activity	163
Figure 57	Detection of activity from the <i>B. pseudomallei</i> CPS biosynthetic pathway	164
Figure 58	SDS-PAGE of <i>B. pseudomallei</i> CPS biosynthetic proteins purified from <i>E. coli</i> DH5 α	167
Figure 59	GDP-6d _B Hep biosynthetic pathway from GmhA through to WcbJ monitoring oxidation of NADH or NADPH	172

Chapter 4

Figure 60	Discontinuous coupled kinase assay for glycosyltransferase activity	182
Figure 61	BacTiter Glo™ reagent activity	183
Figure 62	Glycosyltransferase assay coupled with mesophilic or thermophilic kinase	185
Figure 63	Discontinuous coupled phosphatase assay for glycosyltransferase activity	186
Figure 64	Dialysis of apyrase from <i>Cimex lectularius</i>	196
Figure 65	Elution from nickel affinity purification of mesophilic kinase, Cmk, expressed in <i>E. coli</i> Rosetta (DE3) cells.	199
Figure 66	SDS-PAGE showing denaturation of mesophilic Cmk	200
Figure 67	Elution from size exclusion purification of mesophilic kinase, Cmk, expressed in <i>E. coli</i> Rosetta (DE3) cells	201
Figure 68	SDS-PAGE showing denaturation of mesophilic Cmk	202
Figure 69	SDS-PAGE showing fractions eluted from size exclusion chromatography containing purified mesophilic Cmk	203
Figure 70	Elution from size exclusion purification of thermophilic kinase, Umpk, expressed in <i>E. coli</i> BL21 (DE3) cells	204
Figure 71	Elution from anion exchange of thermophilic kinase, Umpk expressed in <i>E. coli</i> BL21 (DE3) cells	205
Figure 72	Elution from anion exchange of thermophilic kinase, Umpk, expressed in <i>E. coli</i> BL21 (DE3) cells	206
Figure 73	Elution from nickel affinity purification of thermophilic kinase, Umpk expressed in <i>E. coli</i> BL21 (DE3) cells	207
Figure 74	Elution from size exclusion purification of thermophilic kinase, Umpk, expressed in <i>E. coli</i> BL21 (DE3) cells	208
Figure 75	SDS-PAGE of thermophilic kinase Umpk extracted from <i>E. coli</i> BL21 (DE3)	209
Figure 76	SDS-PAGE of purified mesophilic and thermostable kinases from <i>E. coli</i> Rosetta (DE3) and BL21 (DE3) cells	210
Figure 77	Standard curve of 0 - 1 mM ATP	211
Figure 78	Optimising Cmk concentration, with 0.1 mM ADP and 0-0.6 mM UDP	213
Figure 79	Michaelis-Menten and substrate inhibition of 0.2 mg mL ⁻¹ Cmk, 0.1 mM ADP and 0 – 5 mM UDP	215

Figure 80	Standard curve of 0 - 0.6 mM ATP	217
Figure 81	Michaelis-Menten analyses of ATP (μM) produced by kinases	218
Figure 82	Michaelis-Menten curves comparing mesophilic Cmk with UDP and Gmk with GDP.	220
Figure 83	Michaelis-Menten curves comparing thermophilic Cmk and Umpk with UDP, Gmk with GDP and Tmk with TDP	221
Figure 84	Michaelis-Menten curves comparing mesophilic Cmk and thermophilic Cmk with UDP and thermophilic Cmk with CDP	222
Figure 85	Michaelis-Menten curves comparing mesophilic Cmk and thermophilic Cmk with UDP and thermophilic Gmk with GDP	223
Figure 86	Enzyme titration of $\beta 14\text{GalT}$ with 5 mM UDP-Gal and 5 mM GlcNAc	225
Figure 87	Michaelis-Menten analysis to optimise UDP-Gal concentration.	226
Figure 88	Michaelis-Menten analysis to optimise GlcNAc concentration	227
Figure 89	Michaelis-Menten curve of 0.25 mU mL^{-1} $\beta 14\text{GalT}$ with 15 mM GlcNAc and 0 – 6 mM UDP-Gal	228
Figure 90	Michaelis-Menten curve of 0.25 mU mL^{-1} $\beta 14\text{GalT}$ with 15 mM GlcNAc and 0 – 6 mM UDP-Gal	230
Figure 91	Michaelis-Menten curve of $7 \mu\text{g mL}^{-1}$ LgtC with 2 mM lactose and 0 – 0.5 mM UDP-Gal	231
Figure 92	Michaelis-Menten curve of $7 \mu\text{g mL}^{-1}$ LgtC with 2 mM lactose and 0 – 0.5 mM UDP-Gal	232
Figure 93	SDS-PAGE of apyrase after dialysis for 48 h.	234
Figure 94	SDS-PAGE of apyrase after dialysis for 48 h and TCA precipitation	235
Figure 95	SDS-PAGE of apyrase after dialysis for 48 h and nickel affinity chromatography	236

Chapter 5

Figure 96	Hybrid <i>B. pseudomallei</i> CPS attached to truncated <i>E. coli</i> polysaccharides	248
Figure 97	Biosynthesis of hybrid polysaccharides in <i>E. coli</i>	250
Figure 98	Co-expression of three plasmids in <i>E. coli</i> for the synthesis of hybrid CPS.	251
Figure 99	Enzyme-linked immunosorbent assay for detecting <i>B. pseudomallei</i> CPS	258
Figure 100	Immunoblot for detecting <i>B. pseudomallei</i> CPS	261
Figure 101	Testing detergents for purification of WcbH	262
Figure 102	Mini nickel column high-throughput screening for IPTG-induction of WcbH expression	264
Figure 103	Gravity-flow nickel column purification of WcbB, WcbE and WcbH	265
Figure 104	SDS-PAGE of size exclusion chromatography of WcbH and WcbB	266
Figure 105	Differential scanning fluorimetry curve for WcbB	267
Figure 106	Plasmid map for <i>wcbI</i> , <i>wcbB</i> , <i>wcbE</i> and <i>wcbH</i> operon in BioBricks™ vector pSB4A5	269

Figure 107	Plasmid constructs for expression of additional glycosyltransferases in <i>E. coli</i>	270
Figure 108	Immunoblot of IgG positive sera from <i>B. pseudomallei</i> -challenged mice against extracts from <i>E. coli</i>	271
Figure 109	PglB assembly of glycoconjugates in a bacterial cell	277

List of tables

Chapter 2

Table 1	In-house SDS- and BN-PAGE gel composition	101
Table 2	Marker proteins for BN-PAGE	102
Table 3	Antibodies for Western blotting against His- and c-Myc-tagged proteins	104
Table 4	Samples and standards for GC-MS analysis	108
Table 5	Length of genes and fragments used for CPS biosynthetic operon assembly	116
Table 6	Retention time and relative abundance of GDP-6d ₅ Hep (618.0855 m/z) determined by Q-TOF LC-MS	129
Table 7	Retention time of samples and standards analysed by GC-MS	138

Chapter 3

Table 8	Enzymes and substrates for GmhA, WcbL, WcbN activity assay	168
Table 9	Enzymes and substrates for GmhA, WcbL, WcbN, WcbM activity assay	169
Table 10	Enzymes and substrates for separate GmhA, WcbL and WcbN then WcbM reactions	170
Table 11	Enzymes and substrates for testing the effect of different enzyme concentrations and separation of reactions	171
Table 12	MS-MS QQQ analysis of droplets recovered from microfluidics samples	174

Chapter 4

Table 13	Mesophilic and thermophilic kinase parameters	184
Table 14	V_{max} and K_M for ATP standard curve with BacTiter-Glo™ reagent	212
Table 15	V_{max} and K_M for various concentrations of Cmk with 0.1 mM ADP and 0 – 0.6 mM UDP	214
Table 16	Michaelis-Menten and substrate inhibition analysis of 0 – 5 mM UDP with 0.2 mg mL ⁻¹ mesophilic Cmk and 0.1 mM ADP	216
Table 17	V_{max} and K_M parameters for mesophilic and thermostable kinase with 0 – 0.6 mM NDP, 0.1 mM ADP and 0.2 mg mL ⁻¹ kinase.	217
Table 18	V_{max} and K_M parameters for mesophilic and thermostable kinase with 0 – 0.6 mM NDP, 0.1 mM ADP and 0.2 mg mL ⁻¹ kinase	219
Table 19	V_{max} and K_M parameters for mesophilic Cmk and thermostable Cmk and Gmk with 0 – 0.6 mM NDP, 0.1 mM ADP and 0.2 mg mL ⁻¹ kinase	224

Table 20	V_{max} and K_M parameters of 0.25 mU mL ⁻¹ β 14GalT with either 0 – 2 mM UDP-Gal (or 0 – 30 mM GlcNAc	227
Table 21	V_{max} and K_M for β 14GalT with UDP-Gal determined using a novel kinase assay and commercial phosphatase assay	229
Table 22	V_{max} and K_M for LgtC with UDP-Gal determined using a novel kinase assay and commercial phosphatase assay.	232
Table 23	Additional glycosyltransferases for assembly of hybrid <i>B. pseudomallei</i> CPS in <i>E. coli</i>	249
Table 24	Vector properties for co-expression of three plasmids in <i>E. coli</i>	252
Table 25	Samples for immunoblot assay for <i>B. pseudomallei</i> CPS expression in <i>E. coli</i> DH5 α	260

List of presentations and publications

This work was presented internally at the University of Exeter and Dstl as well as the following scientific meetings:

Poster

- South West Structural Biology Conference (Bristol, July 2013)
- Synthetic Biology 6.0 Conference (London, July 2013)
- World Melioidosis Congress (Bangkok, September 2013)
- Society for Glycobiology and Japanese Society of Carbohydrate Research Joint Annual Meeting (Honolulu, November 2014)
- European Melioidosis Congress (Cambridge, March 2015).

Presentation

- Society for Experimental Biology Synthetic Biology Conference (London, January 2014)
- South West Synthetic Biology GW4 group (Bristol, August 2014).

Publications

“A novel coupled kinase assay for the characterisation of glycosyltransferases.”

- Manuscript currently being prepared for submission to Carbohydrate Research journal

Author's declaration

Unless otherwise stated, the work presented in this thesis is solely the work of Victoria Mae Baldwin.

^1H and ^{31}P NMR was performed and analysed by Dr. Martin Rejzek at the John Innes Centre, Norwich. Prof Rob Field's group also provided a sample of 6-deoxy-*manno*-heptose for GC-MS analysis.

GC-MS was run by Dr. Debbie Salmon and analysed by Prof Nick Smirnoff at the University of Exeter.

Dr. Hannah Florance and Dr. Venura Peera ran the statistical analyses on the Q-TOF total metabolite extract data at the University of Exeter.

Experiments for the enzyme assays in a microfluidics system were conducted alongside Divesh Baxani, as part of a collaboration with Dr. Oliver Castell at the University of Cardiff.

Acknowledgements

My eternal love and gratitude for my wonderful Mam. Without her support and encouragement this thesis would not exist.

I would like to thank my supervisors Dr. Nic Harmer and Assoc. Prof. Jo Prior for their guidance and enthusiasm throughout this project. This work was made possible by the funding and facilities provided by the EPSRC CASE Award, the University of Exeter and Dstl.

I would also like to thank Dr. Debbie Salmon, Prof Nick Smirnoff, Dr. Hannah Florence and Maciej Trznadel for their patience and humour when teaching me about mass spectrometry.

Huge thank you to my parents and brother for all their love and encouragement throughout. Also to my human and equine pals at various stables for keeping me active and getting me outdoors, even on my worst days.

Finally, I would like to thank Mirella, David, Sam and Charlie for being the funniest, kindest, smartest and fiercest friends a person could wish for.

List of abbreviations

List of abbreviations

General Terms

2D-PAGE	second dimension-PAGE
BN-PAGE	blue native-PAGE
cDNA	complementary DNA
CFU	colony forming units
CPS	capsular polysaccharide
DNA	deoxyribonucleic acid
GC-MS	gas chromatography mass spectrometry
gDNA	genomic DNA
GI	genomic island
HILIC	hydrophilic interaction chromatography
HPLC	high performance liquid chromatography
IFN- γ	gamma interferon
Ig	Immunoglobulin
IL	Interleukin
LD ₅₀	50 % lethal dose (dose required for 50 % mortality)
LC-MS	liquid chromatography mass spectrometry
LIC	ligation-independent cloning
MHC	major histocompatibility complex
MLD	median lethal dose
MLST	multilocus sequence typing
MNGC	multinucleated giant cell
MRM	multiple reaction monitoring
MS/MS	tandem mass spectrometry
MTTD	mean time to death
NHS	normal human serum
NMR	nuclear magnetic resonance
OD	optical density
OMP	outer membrane protein
PAGE	polyacrylamide gel electrophoresis
PAMP	pathogen associated molecular pattern
PCR	polymerase chain reaction
PFGE	pulsed-field gel electrophoresis
QQQ	triple quadrupole
Q-TOF	quadrupole time of flight
RBS	ribosome binding site
RD	restriction digest
RNA	ribonucleic acid
RNase	Ribonuclease
RT-PCR	reverse transcriptase PCR
SDS-PAGE	sodium dodecyl sulphate-PAGE
TLR	toll-like receptor
TNF- α	tumour necrosis factor α
UV	Ultraviolet
v/v	volume to volume ratio
w/v	weight to volume ratio

Compounds / Proteins

β 14GalT	β -1,4-galactosyltransferase
Acetyl-CoA	Acetyl coenzyme A
ACN	Acetonitrile
ADP	adenosine diphosphate
AMP	adenosine monophosphate
Amp	Ampicillin
ATP	adenosine triphosphate
bis-Tris	bis(2-hydroxyethyl)aminotris(hydroxymethyl)methane
BSA	bovine serum albumin
CaCl ₂	calcium chloride
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
Chl	Chloramphenicol
D ₂ O	heavy water
DAB	3,3'-diaminobenzidine
DEAE	diethylaminoethanol
dH ₂ O	deionised water
dNTP	deoxynucleotide triphosphate
DTNB	5,5'-dithio-bis-[2-nitrobenzoic acid]
DTT	Dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EtOH	Ethanol
GC	guanosine and cytosine
GDP	guanosine diphosphate
GDP-6dbHep	GDP-6-deoxy- α -D- <i>manno</i> -heptose
GDP-4k6Hep	GDP-4-keto-6-deoxy-D- <i>manno</i> -heptose
GlcNAc	N-acetylglucosamine
GTP	guanosine triphosphate
Gu-HCl	guanidinium chloride
H ₂ O ₂	hydrogen peroxide
HCl	hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	horseradish peroxidase
IPTG	isopropyl β -D-1-thiogalactopyranoside
Kan	Kanamycin
KCl	potassium chloride
LB	Luria-Bertani
M1P	D- <i>glycero</i> - α -D- <i>manno</i> -heptopyranose-1-phosphate
M7P	D- <i>glycero</i> - α -D- <i>manno</i> -heptopyranose-7-phosphate
mAb	monoclonal antibody
MeOH	Methanol
MgCl ₂	magnesium chloride
MnCl ₂	manganese chloride
MOPS	3-(N-morpholino)propanesulphonic acid
MSTFA	N-methyl-N-(trimethylsilyl)trifluoroacetamide
NaCl	sodium chloride
NAD ⁺ /NADH	nicotinamide adenine dinucleotide

NAP ⁺ /NADPH	nicotinamide adenine dinucleotide phosphate
Octyl glucoside	octyl β-D-glucopyranoside
PBS	phosphate buffered saline
PMSF	phenylmethylsulfonyl fluoride
RbCl	rubidium chloride
S7P	sedoheptulose-7-phosphate
Sarkosyl	sodium lauroyl sarcosinate
TEAA	triethylammonium acetate
TEMED	N,N,N',N'-tetramethylethane-1,2-diamine
TFA	trifluoroacetic acid
UDP	uridine diphosphate
UDP-Gal	uridine diphosphate galactose

Units

%	Percentage
A	Ampere
bp	base pair
Da	Dalton
eV	Electronvolt
G	Gram
H	Hour
K	Kilo
L	Litre
M	milli or metre
M	Molar
min	Minute
N	Nano
°C	degrees Centigrade
P	Pico
pH	potential of hydrogen (acidity)
psig	pounds per square inch gauge
rA	relative abundance
S	Second
S	Siemen (electrical conductance)
V	Volts
xg	times gravity (centrifugal force)
M	Micro

Chapter 1

Introduction

1.1 *Burkholderia pseudomallei* and melioidosis

Burkholderia pseudomallei was first isolated in Yangon, Myanmar (then known as Rangoon, Burma). It was described by Whitmore (1913) as Gram negative, motile, facultatively anaerobic, bacillus bacteria. *B. pseudomallei* is still of scientific interest today due to its ongoing impact on human health as the aetiological agent of melioidosis (Whitmore, 1913; Limmathurotsakul *et al.*, 2016) and classification as a potential bioterror threat (Gilligan, 2002; Butler, 2012).

1.1.1 Phylogeny and genome

The genus *Burkholderia* was established by Yabuuchi *et al.* (1992). Seven species of *Pseudomonas*, including *Pseudomonas pseudomallei*, were reclassified into the new genus based on 16S rRNA sequencing, DNA-DNA homology (including guanosine and cytosine (GC)-richness) and phenotypic observations such as fatty acid composition, motility and colony morphology. The *Burkholderia* genus contains more than 90 species, ranging in diversity from the non-pathogenic plant symbiont *B. tuberum* (Angus *et al.*, 2014) to the highly infectious human pathogen and bioterror threat *B. pseudomallei* (Nelson *et al.*, 2011). *B. mallei*, the closest relative of *B. pseudomallei*, is also considered a potential bioterror agent and causes Glanders in solipeds, which posed a significant threat during World War I (Whitlock *et al.*, 2007). Another close relative of *B. pseudomallei* is *B. thailandensis*, which is often used as an 'avirulent' model organism (Haraga *et al.*, 2008). However, there have been rare incidents of *B. thailandensis* pathogenicity in humans reported (Glass *et al.*, 2006). Other human pathogens found in the genus include a group of species known as the '*B. cepacia* complex', often isolated from cystic fibrosis patients (Vandamme *et al.*, 1997; Mahenthiralingam *et al.*, 2005). The diversity of these species has

prompted the recommendation of a separate genus, *Paraburkholderia*, to distinguish between pathogenic and non-pathogenic strains. This differentiation is currently achieved by dividing the genus into two clades based 16S rRNA sequencing (Figure 1; Sawana *et al.*, 2014).

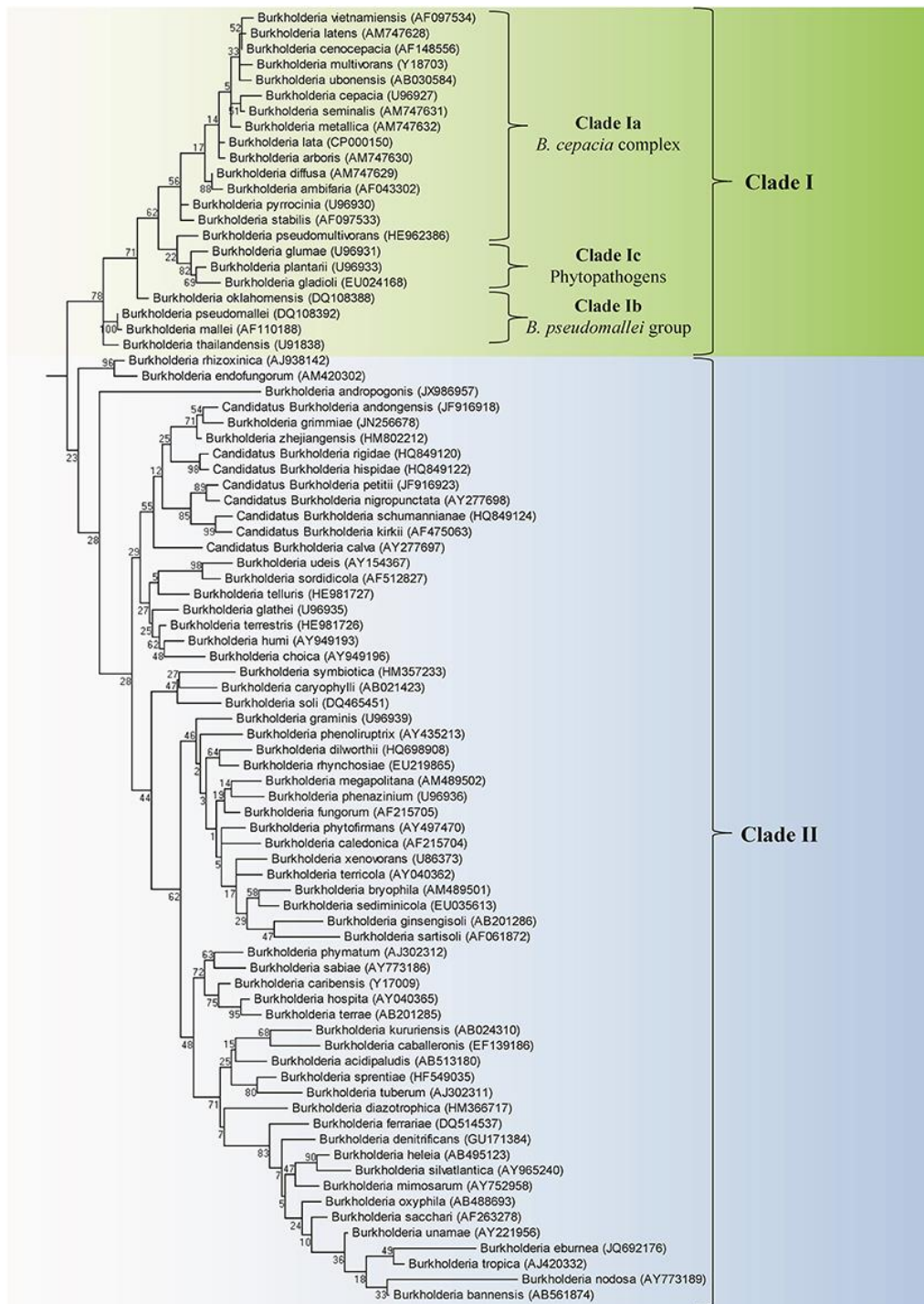


Figure 1: *Burkholderia* phylogenetic tree. Maximum likelihood tree based on 16S rRNA sequences of the species within the *Burkholderia* genus. Showing demarcation of Clade I (pathogenic) and Clade II (non-pathogenic) species. Clade I is further devolved into subclades by pathogenic function. Reproduced from Sawana *et al.* (2014).

The complete genome of *B. pseudomallei* strain K96243, originally isolated from an infected human, was sequenced by Holden *et al.* (2004). The genome is large (7.25 Mbp) and GC-rich (67 %) in comparison to other Gram negative bacteria. For example, the *Escherichia coli* K12 MG1655 (EMBL/GenBank: U00096) genome is 4.64 Mbp in size and has a 51 % GC content (Blattner *et al.*, 1997). The *B. pseudomallei* genome is divided between two chromosomes. Chromosome 1 (EMBL/GenBank: BX571965) is 4.07 Mbp in size and encodes many essential core functions, such as amino acid metabolism. Chromosome 2 (EMBL/GenBank: BX571966) is smaller at 3.17 Mbp and encodes accessory functions for the adaptation to atypical conditions (Holden *et al.*, 2004; Sim *et al.*, 2008). Variation between species and strains occur partly due to the presence of genomic islands (GIs). These non-native sequences of DNA are obtained by lateral transfer (often via bacteriophages) and have been identified within both chromosomes. They frequently encode virulence factors and adaptation mechanisms. For instance, mutation of the GI gene BPSS2053 (*fhaB*) on chromosome 2 caused a reduction in bacterial adherence to epithelial cells (Sim *et al.*, 2008). Ten genes involved in lipopolysaccharide (LPS) synthesis, also a virulence factor, were found distributed over three distinct GI regions on chromosome 2 (Sim *et al.*, 2008).

1.1.2 Epidemiology of *B. pseudomallei*

1.1.2.1 Distribution in the environment

Originally described as a soil-dwelling saprophyte (Whitmore, 1913; Dance, 1991), *B. pseudomallei* is naturally found throughout the environment. It is abundant in the rhizosphere, particularly in exotic grasses and may be disseminated by grazing animals (Berg *et al.*, 2005; Kaestli *et al.*, 2012). Present

in both soil and water, rice paddy fields have been identified as a substantial risk for harbouring the pathogen, posing a threat to human health (Brook *et al.*, 1997; Vongphayloth *et al.*, 2012). In a study by Rattanavong *et al.* (2011), soil samples from nine rice paddy fields throughout Laos showed 44 % of sites were contaminated with viable *B. pseudomallei*. Of the multiple samples taken from one site in the study, 94 % tested positive for the pathogen with a mean burden of 464 colony forming units (CFU) per gram of soil. A similar investigation in Thailand found 28 % of samples from a single site tested positive for *B. pseudomallei* (Wuthiekanun *et al.*, 2009). These studies show that where present, *B. pseudomallei* is ubiquitous in the environment and in high abundance.

1.1.2.2 Global distribution

B. pseudomallei and melioidosis were historically associated with Thailand and Northern Australia (Limmathurotsakul *et al.*, 2010; Currie *et al.*, 2010; McRobb *et al.*, 2014). However, epidemiological studies have revealed a global distribution throughout tropical and sub-tropical regions including Africa (Cuadros *et al.*, 2011; Morosini *et al.*, 2013; Wiersinga *et al.*, 2015), South and Central America (Inglis *et al.*, 2006), Oceania (Currie *et al.*, 2000b; Baker *et al.*, 2011a) and South Asia (Jesudason *et al.*, 2003). The global distribution of *B. pseudomallei* and melioidosis are discussed in-depth by Currie *et al.* (2008) and Limmathurotsakul *et al.* (2013; 2016) (Figure 2).

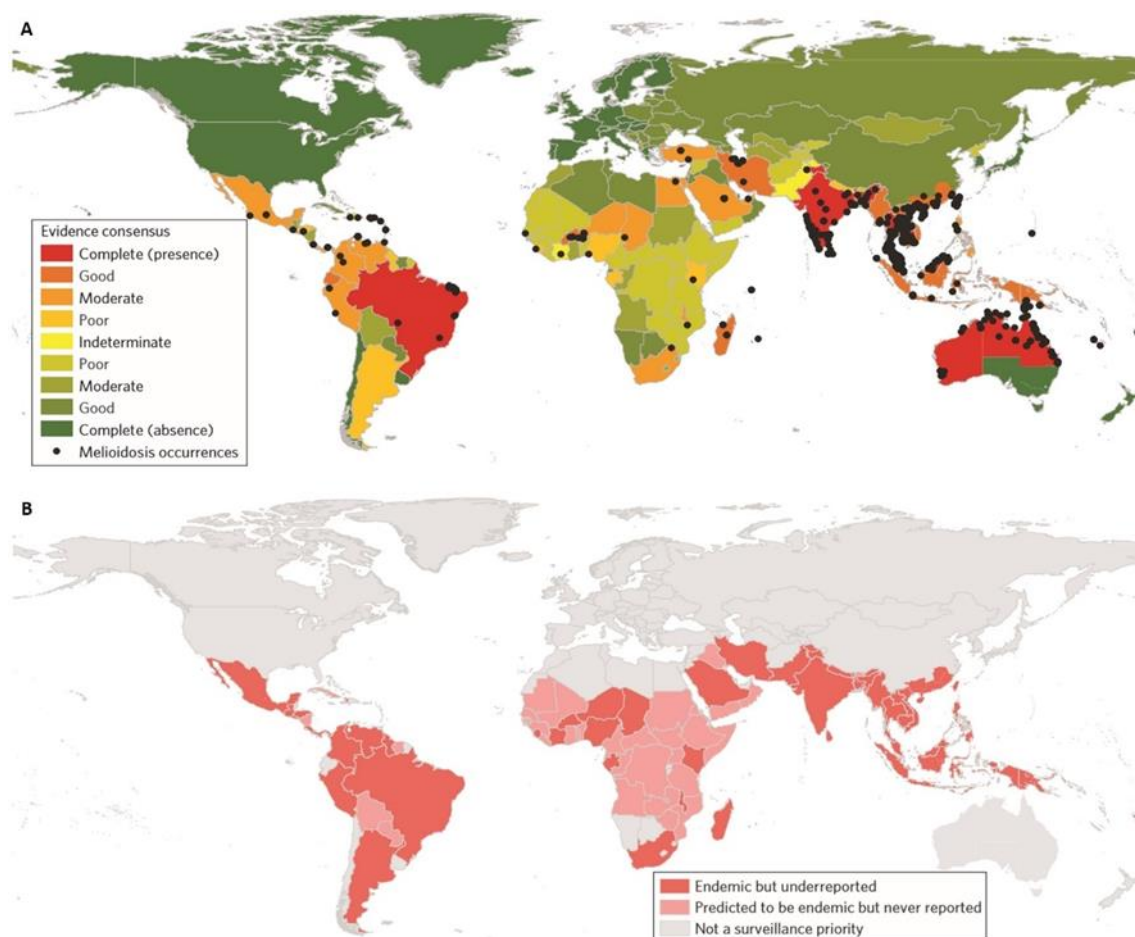


Figure 2: Global distribution of *Burkholderia pseudomallei* and incidence of melioidosis. (A) Presence or absence of *B. pseudomallei* and melioidosis per country, based on reported cases during 1901-2014. Colour gradient dependant on consensus of collated evidence. (B) Surveillance priorities for *B. pseudomallei* and melioidosis based on predicted of incidence of *B. pseudomallei* infection. Reproduced and adapted from Limmathurotsakul *et al.* (2016).

South East Asia and Northern Australia show the highest incidence of melioidosis, with fewer cases in Central and South America and Africa (Figure 2). However, it is suggested that *B. pseudomallei* is likely endemic throughout the tropics in South Asia, Oceania, Central and Southern America and Africa (Figure 2); Limmathurotsakul *et al.*, 2016). Due to a lack of awareness of the pathogen outside South-East Asia and Northern Australia, misdiagnosis is likely causing an underreporting of cases. This has been highlighted as a key problem for the prevention and treatment of melioidosis globally (Wilde & Suankratay, 2007). Consequently, there is a growing interest in a more robust surveillance

system for *B. pseudomallei* infection, in addition to improved antibiotic and vaccine development strategies (Limmathurotsakul *et al.*, 2016).

Isolated incidents have also been reported in Europe, North America and the Middle East. In these areas, melioidosis is usually observed in people who have travelled to endemic regions (Schülin & Steinmetz, 2001; Visca *et al.*, 2001; Currie, 2003; Cuadros *et al.*, 2011; Darazam *et al.*, 2011; Morosini *et al.*, 2013; Amadasi *et al.*, 2015). However, domestically-acquired *B. pseudomallei* infection appears to be emerging in Central and Northern America (Engelthaler *et al.*, 2011; O'Sullivan *et al.*, 2011, Doker *et al.*, 2014; Truong *et al.*, 2015).

1.1.3 *B. pseudomallei* infection in humans

1.1.3.1 Route of transmission

Transmission of *B. pseudomallei* from the environment to a human host can occur by inoculation via the skin, water-borne ingestion or aerosol inhalation.

Percutaneous inoculation

Percutaneous infection occurs when small abrasions are sustained, breaking the epidermal barrier. The infective dose for percutaneous inoculation is thought to be low. Experimental data show as few as 26 CFU are required for subcutaneous infection and progression to melioidosis in marmosets (*Callithrix jacchus*) (Nelson *et al.*, 2014). Percutaneous inoculation is a common route of infection, described in 25 % of cases in humans from Northwest Australia and Papua New Guinea (Currie *et al.*, 2000b; Cheng & Currie, 2005). Agricultural work has been identified as a risk factor for *B. pseudomallei* infection, due to exposure to the environment and likelihood of injury. Assessment of 337 melioidosis patients in Thailand and Malaysia indicated 228 (67.7 %) were agricultural, forestry or fishery workers

(Suputtamongkol *et al.* 1999; Hassan *et al.* 2010). However, any outdoor activity in an endemic area poses a risk; one report describes a patient who developed melioidosis after sustaining an abrasion on a sports field in Australia (Hill *et al.*, 2013).

Ingestion

Infection with *B. pseudomallei* may occur via ingestion of contaminated drinking water. Water supplies testing positive for *B. pseudomallei* have been identified in Thailand (Limmathurotsakul *et al.*, 2014; Thaipadungpanit *et al.*, 2014) and Northern Australia (Inglis *et al.*, 2000a; Currie *et al.*, 2001; Draper *et al.*, 2010). Two cases of melioidosis probably caused by contaminated drinking water have been observed (Limmathurotsakul *et al.*, 2014). Two other incidents involved a combination of inhalation and ingestion during near-drowning accidents (Lee *et al.*, 1985; Pruekprasert & Jitsurong, 1991).

Inhalation

Transmission of *B. pseudomallei* by aerosol was observed in soldiers inhaling the updraft caused by helicopters during the war in Vietnam. Cases of reactivated melioidosis and *B. pseudomallei* carriage have been subsequently reported in veterans (Chodimella *et al.*, 1997; Currie *et al.*, 2000a). A more recent case of infection was reported in a tourist following a recreational helicopter flight in Singapore (Amadasi *et al.*, 2015). An increased rate of infection is also observed during monsoon/typhoon season, when turbulent weather disturbing the environment may cause an increase in airborne microbes aerosolised from the environment (Ko *et al.*, 2007; Chen *et al.*, 2014; 2015).

The infective dose for this route of transmission is suspected to be very low. A study in marmosets (*C. jacchus*) could not determine the intranasal dose required to cause mortality in 50 % test subjects (LD₅₀), as it was below 10 CFU (Nelson *et al.*, 2011). This is taken to indicate that a single bacterium is potentially sufficient to colonise a host and establish an infection via this route.

Human-human transmission and zoonosis

Although human-human transmission and zoonosis of *B. pseudomallei* are rare, they remain potential routes of infection. Transmission between humans is most likely to occur between mother and baby. A case of neonatal infection in India (Noval *et al.*, 2009) and two cases in breastfed babies in Northern Australia (Ralph *et al.*, 2004) have been reported.

Despite *B. pseudomallei* being isolated from a wide range of animals (Sprague & Neubauer, 2004), there have been few incidents of zoonotic transmission to date (Choy *et al.*, 2000). However, two cases of imported pet iguanas infected with *B. pseudomallei* and presenting clinical symptoms of melioidosis have been reported recently in Czech Republic and USA (Elschner *et al.*, 2014; Zehnder *et al.*, 2014). The movement of animals carrying *B. pseudomallei* could increase global dissemination and pose a threat to both human and animal health.

1.1.3.2 Risk factors for infection in humans

Several risk factors that increase susceptibility to *B. pseudomallei* infection and melioidosis have been identified, including immune status, co-morbidity, lifestyle choices and seasonal rainfall.

Occupation

As *B. pseudomallei* naturally occurs in the environment, occupations that expose workers to contaminated soil and water present a heightened risk of infection. For example, a survey in Thailand reported 85 % of melioidosis patients were rice farmers (Suputtamongkol *et al.*, 1999). A study of melioidosis patients in Malaysia found 18.6 % cases occurred in farm, forestry and fishery workers and 18.6 % in the unemployed, who are likely to survive by tending their own crops and gardening. In comparison, 6.9 % of cases occurred in administrative workers (Hassan *et al.*, 2010).

Immune Status

Four studies analysed data from melioidosis patients in Thailand, Papua New Guinea and Australia, India and Malaysia (Suputtamongkol *et al.*, 1999; Currie *et al.*, 2000b; Vidyalakshmia *et al.*, 2012; Hassan *et al.*, 2010). Compiling data from the 646 described cases reveals the average age of patients was 50.3 years and that 69.4 % were male. Of the patients, 57.2 % were diabetic, 19.9 % suffered alcohol misuse and 13.1 % had chronic renal conditions. Co-morbidity with chronic pulmonary or hepatic diseases, thalassaemia, cancer and smoking were identified as risk factors. These data demonstrate that immunocompromised adults, particularly those with diabetes mellitus, are most at risk of developing melioidosis.

Rainfall

An environmental risk factor is increased rainfall and subsequent groundwater runoff (Currie & Jacups, 2003; Currie *et al.*, 2010). Increased rates of infection and mortality are typically recorded during monsoon/typhoon seasons (Baker *et*

al., 2011b; Chen *et al.*, 2014). Whilst higher rainfall increases the risk of infection, the same effect is not observed with faster wind speed and dissemination is not influenced by wind direction (Chen *et al.*, 2014) This suggests it is the physical impact of rain on the ground that causes aerosolisation of *B. pseudomallei* in the environment.

1.1.4 Melioidosis

1.1.4.1 Clinical symptoms

Melioidosis in humans usually manifests as fever and other non-specific symptoms. *B. pseudomallei* has been isolated from most tissues in the human body and symptoms that develop are often dependent upon the route and localisation of infection. For example, inhalation can cause a chronic lung infection and pneumonia (Meumann *et al.*, 2012), whereas percutaneous inoculation can result in acute skin infection and lesions (Gibney *et al.*, 2008). Other pathologies include, but are not limited to, osteomyelitis (Vollmar *et al.*, 2014), arthritis (Rajadhyaksha *et al.*, 2012), liver, spleen and lung abscesses (Currie *et al.*, 2010), suppurative parotitis (Shivbalan *et al.*, 2010) and sepsis (Bart & Cheng, 2006).

In addition to the non-specific symptoms, many cases involve several pathologies simultaneously. For example, a patient in Venezuela (only the second case in the country at the time of recording) presented with osteomyelitis in the right tibia, right parietal bone and sternum, arthritis in the left shoulder and right knee and subperiosteal abscesses, before developing sepsis (Redondo *et al.*, 2011). As these symptoms appeared unrelated and the disease is largely unknown in the region, the correct diagnosis of melioidosis took years to obtain.

twice as quickly in human macrophages than those from type II or type III colonies (Tandhavanant *et al.*, 2010). However, bacteria from type III colonies are more resistant to antimicrobial peptides and more likely to undergo morphology switching (Tandhavanant *et al.*, 2010). A wrinkled colony morphology may be caused by the expression of capsular polysaccharide (CPS), which is an important virulence factor. *B. thailandensis* is usually acapsular and has smooth colonies. However, atypical *B. thailandensis* E555 expresses the *B. pseudomallei* CPS also exhibits some colony wrinkling (Sim *et al.*, 2010). Understanding the underlying mechanisms that determine colony morphologies could have huge implications for diagnosis of melioidosis. It may be used as a predictor for chronic or persistent infection or to inform the type and duration of antibiotic therapy required.

In cases where a culture-positive diagnosis is unattainable, diagnostic PCR (Novak *et al.*, 2006) and antigen detection by lateral flow assay (Houghton *et al.*, 2014; Robertson *et al.*, 2015) are acceptable alternative methods. However, unlike the culture method, these techniques do not provide evidence of live, viable *B. pseudomallei*.

1.1.4.3 Treatment

B. pseudomallei is resistant to many antibiotics including aminoglycosides, chloramphenicol, fluoroquinolones, macrolides, tetracycline and trimethoprim. Therefore, options for antimicrobial therapy are limited. Treatment usually consists of 10 – 14 days' intravenous administration of ceftazidime, followed by 3 – 6 months' oral trimethoprim–sulfamethoxazole. The lengthy and expensive treatment programme can be challenging in terms of patient compliance and affordability (Wiersinga *et al.*, 2012). However, if properly administered,

treatment with ceftazidime reduces mortality rates by 50 % (White *et al.*, 1989).

1.1.4.4 Mortality rates

Reported melioidosis mortality rates range from 16 % in Singapore (Heng *et al.*, 1998) and Northern Australia (Currie *et al.*, 2004) to 43 % in Thailand (Limmathurotsakul *et al.*, 2010) and 44% in Malaysia (How *et al.*, 2009). In Thailand, melioidosis is the third most common cause of death from infection, after AIDS and tuberculosis (Limmathurotsakul *et al.*, 2010).

1.1.5 Virulence factors

B. pseudomallei has a range of virulence factors that enable the bacteria to adapt and survive in both extracellular and intracellular environments, but can also cause pathology in a host. *B. pseudomallei* infection and melioidosis has been observed in a range of species including plants (Lee *et al.*, 2010), amoebae (Inglis *et al.*, 2000b), nematodes (O'Quinn *et al.*, 2001), insects (Fisher *et al.*, 2012), birds (Hampton *et al.*, 2011), reptiles (Elschner *et al.*, 2014; Zehnder *et al.*, 2014), fish, amphibians and marsupials (Choy *et al.*, 2000). The following discussion concerns virulence factors pertinent to mammalian hosts. Experimental animal models are discussed in section 1.3.1.

1.1.5.1 Intracellular survival

Naturally found in soil and water, *B. pseudomallei* increases its ability to survive in the environment by invading amoebae (*Acanthamoeba* species) (Inglis *et al.*, 2000b). Inside the amoebae, the bacteria are protected from the extracellular environment, posing a significant challenge for the decontamination of drinking water (Howard & Inglis, 2005). This mode of survival is mirrored in the human

body, as *B. pseudomallei* invades host cells (Figure 4). The intracellular habitat protects bacteria from the immune system and antibiotics. *B. pseudomallei* invades fibroblasts, epithelial cells and immune cells, including macrophages, polymorphonuclear leukocytes (e.g. neutrophils) and mononuclear leukocytes (monocytes) (Pruksachartvuthi *et al.*, 1990; Jones *et al.*, 1996).

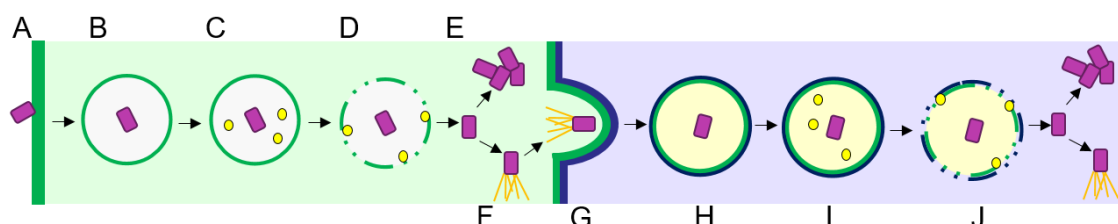


Figure 4: Intracellular lifecycle of *B. pseudomallei*. (A) Bacteria (purple rectangles) enter host cell by attachment and invasion or phagocytosis. (B) Internalised bacteria are encapsulated in a phagosome/vesicle within the host cell cytosol (green). (C) Bacteria release proteins (yellow circles) using secretion systems. (D) secreted proteins degrade vesicle membrane, releasing bacteria into the host cell cytosol. (E) Bacteria can replicate and (F) form actin tails (orange) to gain directed motility. (G) Bacteria move to new host cell (blue) via adjacent cell membranes. (H) They are encased in a vesicle consisting of a double membrane from the original and new host cells. (I) Degradative proteins are secreted by the bacteria. (J) Vesicle double membrane degrades, releasing the bacteria into the cytosol and the lifecycle begins again with (E) replication and (F) directed motility. Adapted from Ray *et al.* (2009).

Few other species of bacteria are known to survive intracellular destruction by phagocytes. These include *Shigella flexneri*, *Listeria monocytogenes* and *Francisella tularensis*. Reviews by Ray *et al.* (2009) and Willcocks *et al.* (2016) provide detailed overviews of intracellular survival mechanisms.

B. pseudomallei either actively invades epithelial cells or is passively phagocytosed by macrophages and neutrophils (Figure 4A/B). Once internalised, the bacteria quickly escape the phagosome or vesicle to avoid destruction by reactive nitrogen and oxygen species and acidification via lysosome fusion (Yates *et al.*, 2005; Haas, 2007). A type III secretion system (T3SS) is used to release proteins which enable the bacteria to escape from the vesicle (Figure 4C/D). *B. pseudomallei* trapped in the vesicle exhibit a greatly reduced

rate of replication (Stevens *et al.*, 2002; Pilatz *et al.*, 2006). Upon escaping into the more favourable conditions of the cytosol, *B. pseudomallei* restores its ability to replicate by adapting 22 % of its transcriptome to the cytosolic environment (Figure 4E). Genes with functions in metabolism, flagellar motility and cell envelope synthesis are down-regulated, whereas genes involved in anaerobic respiration are upregulated to adapt to the low-oxygen conditions (Chieng *et al.*, 2012). *B. pseudomallei* acquires phosphate, iron and magnesium from the host cell cytosol and synthesises purine, histidine and *p*-aminobenzoic acid which are necessary for growth (Pilatz *et al.*, 2006). In the cytosolic environment, *B. pseudomallei* also achieves directed motility by actin polymerisation (Figure 4F; Stevens *et al.*, 2006). This enables the bacteria to spread from cell to cell via adjacent membranes, preventing exposure to the immune system by avoiding an extracellular phase (Figure 4G). In the new host cell, *B. pseudomallei* is encased in a vesicle formed of the two host cell membranes from which it must escape to return to the more favourable cytosolic environment (Figure 4H-J). Spreading between host cells also causes formation of multinucleated giant cells (MNGCs) (Kespichayawattana *et al.*, 2000; Suparak *et al.*, 2005). This phenomenon is uniquely observed in *B. pseudomallei*, *B. mallei* and *B. thailandensis* and is controlled by a type VI secretion system (T6SS) (Burtnick & Brett, 2013; Schwarz *et al.*, 2014). It is comparable to granuloma formation in *Mycobacterium tuberculosis* infections and may prevent access to the infecting pathogen by the immune system, reducing opsonisation, phagocytosis and pathogen clearance (Lay *et al.*, 2006).

1.1.5.2 Flagellar motility

First descriptions of *B. pseudomallei* revealed the bacteria are flagellated and capable of directed motility (Whitmore, 1913). Later observations demonstrated constitutive expression of flagella *in vitro* and *in vivo*, indicating their importance for survival in the environment and within a host (Wikraiphath *et al.*, 2009; Tuanyok *et al.*, 2006). In addition to motility, flagella participate in the adherence to and invasion of cells. Inglis *et al.* (2003) demonstrated aflagellate *B. pseudomallei* lose the ability to invade amoebae (*Acanthamoeba astronyxis*). However, flagella are not universally essential for the colonisation of a host. DeShazer *et al.* (1997) used the Tn-0T182 transposon to create knockouts of *fliC*, a structural protein integral to the flagella filament. Despite the mutation, *B. pseudomallei* maintained virulence in Syrian hamster and diabetic rat models. This was later confirmed by Tuanyok *et al.* (2006). The importance of flagella as a virulence factor appears to be dependent on the model of infection. A study by Chua *et al.* (2003) created an aflagellate mutant of *B. pseudomallei* KHW by deleting the *fliC* gene. The non-motile bacteria were able to invade and replicate within human lung epithelial cells and to kill roundworm (*Caenorhabditis elegans*) with the same virulence as the wild type. However, the mutant *B. pseudomallei* KHW were avirulent in both intranasal and intraperitoneal infection of BALB/c mice. Bacterial load in the spleens and lungs of challenged mice was extremely low. Virulence was restored upon *fliC* complementation (Chua *et al.*, 2003).

1.1.5.3 Host cell adhesion

Host-cell adhesion is an important virulence factor; initiating the intracellular lifecycle in non-phagocytic cells (Figure 4A). *B. pseudomallei* is significantly more efficient at attaching to and colonising human epithelial cells than

B. thailandensis, which is considered avirulent under most circumstances (Kespichayawattana *et al.*, 2004).

B. pseudomallei expresses pili (fimbriae) on the surface of the cell. These hair-like structures attach to host cell surfaces. Eight genetic loci encoding seven type IV pili subunits and additional accessory proteins have been identified in *B. pseudomallei* K96243 (Essex-Lopresti *et al.*, 2005). Deletion of the *pilA* gene, which encodes a fimbrial precursor protein, resulted in reduced attachment of *B. pseudomallei* to three human epithelial cell lines *in vitro* compared to wild type K96243 (4.0 % and 21.3 % adherence rates respectively). The same study also demonstrated the *pilA* deletion mutant had reduced virulence in *C. elegans* (99 % mortality achieved at 58 hours with the mutant compared to 40 hours with the wild type K96243). However, the mutant was only weakly attenuated in BALB/c mice following a low-dose intranasal challenge compared to wild type K96243 (Essex-Lopresti *et al.*, 2005). In mammalian models, other virulence factors and adhesion mechanisms may be sufficiently upregulated to establish a pathogenic infection.

There are also several adhesins involved in host cell attachment (Balder *et al.*, 2010; Lafontaine *et al.*, 2014; Adler *et al.*, 2015). These proteins interact with both the bacteria (often the pili) and the host cell surface, forming an attachment between pathogen and host. Individual deletion mutations of genes encoding trimeric autotransporter adhesins (*boaA*, *boaB*, *bpaA*, *bpaB*, *bpaC*, *bpaD*, *bpaE*, *bbfA* and *bimA*) resulted in attenuation of *B. pseudomallei* K96243 in BALB/c mice. The median lethal dose (MLD) increased between 10-fold for *boaA* and *bimA* mutants to 95-fold for the *bpaA* mutant when compared to wild type K96243 (Adler *et al.*, 2015). No attenuation was observed for deletions of *bpaB* or *bpaD*.

1.1.5.4 Secretion systems

Secretion systems are important virulence mechanisms in a broad range of pathogenic bacteria. They are used to deliver effector proteins to a host cell, performing functions such as host cell attachment, nutrient acquisition, vesicle/phagosome escape, intercellular dissemination and host cell toxicity. Green & Meccas (2016) recently reviewed bacterial secretion systems.

Type 3 secretion systems

Three different T3SS have been identified in *B. pseudomallei*. The second and third T3SS are also found in *B. mallei* and *B. thailandensis*, however only *B. pseudomallei* possesses the first T3SS (Rainbow *et al.*, 2002). The first and second T3SS share homology with secretion systems found in the plant pathogens *Ralstonia solanacearum* and *Xanthomonas* species (Winstanley *et al.*, 1999; Attree & Attree, 2001). Predictably, these T3SS are virulence determinants in *B. pseudomallei* infection of tomato plants (Lee *et al.*, 2010). The third T3SS is encoded on the Bsa locus and shares homology with the Inv/Mxi-Spa-T3SS in *Shigella* and *Salmonella* species (Stevens *et al.*, 2002). Whilst the first and second T3SS have limited influence on virulence of *B. pseudomallei* 1026b derivative in Syrian golden hamsters, the third Bsa T3SS was described as a virulence determinant in the same experiment (Warawa & Woods, 2005).

The Bsa T3SS is required for full virulence in the intracellular lifecycle of *B. pseudomallei*. It secretes several effector proteins with functions in the invasion of host epithelial cells (Balder *et al.*, 2010), immune system evasion and escape from the vesicle/phagosome (Gong *et al.*, 2011) and intercellular dissemination (Suparak *et al.*, 2005). *B. pseudomallei* 576 with deletions in Bsa T3SS effector protein BipD, BopE, BopA or BopB were attenuated in BALB/c

mice following intraperitoneal challenge (Stevens *et al.*, 2004). The median survival time (MST) of the mice was extended from day 20 post-challenge with wild type to day 21, 30, 32 and 50 post-challenge with *bopE*, *bopB*, *bopA* and *bipD* defective mutants respectively. The attenuation was significant in all the deletion mutants except for *bopE*. However, all mice inoculated with the most attenuated *bipD* mutant succumbed to the disease by day 63 post-infection, suggesting these proteins contribute to virulence but are not individually essential. A more recent study using a higher dose of *B. pseudomallei* K96243 investigated the BipC protein (Kang *et al.*, 2015). All BALB/c mice challenged with the wild type succumbed to infection by day 2 post-inoculation. In comparison, all mice challenged with the *bipC* deletion mutant survived until the end of the study (day 35 post-inoculation). This suggests BipC has an essential role in virulence using this model of infection.

Another virulence factor secreted by the T3SS, translocator protein BipB, is important in the formation of MNGCs (Suparak *et al.*, 2005). The percentage of macrophage nuclei in MNGC formation reduced from 96.5 % when challenged with wild type *B. pseudomallei* K96243 to 15.1 % when challenged with a *bipB* deletion mutant.

The BsaZ and BipD effector proteins, also secreted by the T3SS, are essential for actin-dependent movement within the cytosol and membrane protrusion for movement between cells (Stevens *et al.*, 2002).

Type VI secretion system

B. pseudomallei expresses a T6SS which is involved in intercellular dissemination and MNGC formation. Deletion of T6SS structural protein Hcp1 caused attenuation of *B. pseudomallei* K96243 in Syrian hamsters (3-fold

increase in intraperitoneal LD₅₀). The mutant *B. pseudomallei* also showed reduced replication and MNGC formation in RAW 264.7 macrophages compared to wild type (Burtnick *et al.*, 2011).

Other secretion systems

Other secretion systems have been identified in *B. pseudomallei* but have not been shown to influence virulence determination. For example, deletion of type II secretion system (T2SS) proteins protease, lipase, and phospholipase C are not attenuating in Syrian hamsters (DeShazer *et al.*, 1999).

1.1.5.6 Polysaccharides

Polysaccharides are usually expressed on the bacterial cell surface or secreted into the extracellular environment. They can be significant virulence factors and potent immunogens. Two predominant exopolysaccharides found in all *B. pseudomallei* strains are the CPS and LPS. They are known as type I and type II O-antigens respectively. Other polysaccharides have been identified in *B. pseudomallei* but in lower abundance, are not present in all strains and have less influence in virulence determination. A chapter discussing the polysaccharides of *B. pseudomallei* can be found in section 1.2.

1.1.5.7 Antibiotic resistance

A huge challenge for the treatment of melioidosis is the inherent resistance of *B. pseudomallei* to a wide range of antibiotics. This is largely due to the expression of three efflux pump systems, AmrAB-OprA, BpeAB-OprB and BpeEF-OprC (Moore *et al.*, 1999; Chan *et al.*, 2004; Kumar *et al.*, 2006; Mima *et al.*, 2011). Recently, mutations in β -lactamase PenA indicated resistance to

the frontline drug ceftazidime may be emerging (Rholl *et al.*, 2011; Randall *et al.*, 2015). Isolates of ceftazidime and amoxicillin-clavulanic acid-resistant *B. pseudomallei* have also been identified (Wuthiekanun *et al.*, 2011). As antibiotic resistance continues to increase and discovery of novel therapeutics remains slow, there is an increasing need to develop effective prophylactic vaccines rapidly and economically (French, 2010; Norrby *et al.*, 2005). Another virulence factor that reduces antibiotic efficacy is isocitrate lyase, which is an exacerbating factor in persistent infections (van Schaik *et al.*, 2009).

1.1.5.8 Latent, persistent and chronic infections

Latent, persistent and chronic infections pose significant challenges for the diagnosis and treatment of melioidosis. Latent infections can occur several years after the initial inoculation event. In one instance, cutaneous melioidosis manifested in a former prisoner of war 62 years post-exposure to *B. pseudomallei* (Ngauy *et al.*, 2005). Latent infections are likely kept under control by the host immune system until there is a traumatic event (e.g. sustaining a wound) or a change in immune status (e.g. aging or developing diabetes mellitus).

Persistent infections are established when a small proportion of the pathogen population tolerates antibiotic therapy. Upon cessation of treatment the bacteria repopulate and the disease recurs. Despite the different phenotypes, persister cells exist within a clonal population and are genetically identical to normal cells. This is demonstrated by the percentage of persister cells remaining the same following regrowth of a bacterial population and maintaining a heterologous phenotype (Figure 5; Biggar, 1944).

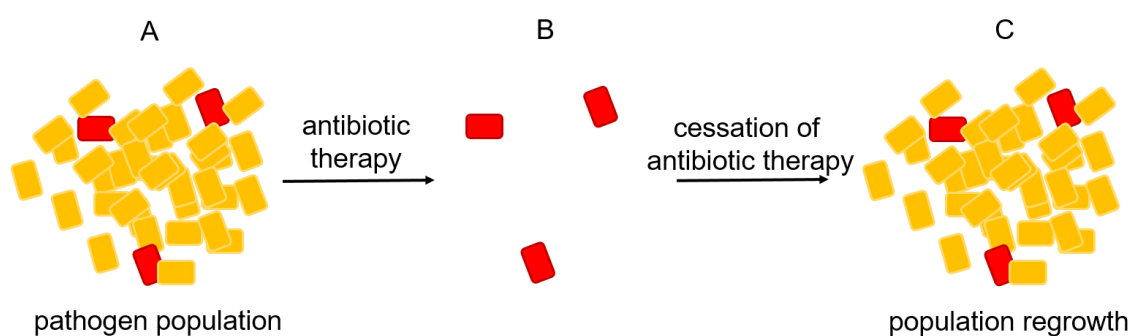


Figure 5: Persister cell tolerance to antibiotics. (A) Bacterial population contains normal cells (orange) and a small percentage of persister cells (red). (B) Persister cells tolerate exposure to antibiotics and the bacterial population become homogeneous. (C) Upon cessation of antibiotic exposure, the bacterial population regrows. However, the percentage of persister cells remains consistent with pre-exposure levels and heterologous phenotype is restored.

As well as antibiotic tolerance, persister cells capable of withstanding osmotic pressure and nutrient starvation (Nierman *et al.*, 2015). The mechanisms of persister cell development and survival are not understood, though recent studies have identified toxin antitoxins as a potential factor. A study by Butt *et al.* (2014) showed increasing levels of HicA toxin expression concurrently increased tolerance to ciprofloxacin and ceftazidime. The same study showed deletion of the *hicAB* locus decreased the percentage of persister cells within a *B. pseudomallei* K96243 population.

If a patient survives initial acute melioidosis, chronic disease symptoms can occur. A study of four chronic cases of melioidosis lasting six year found *B. pseudomallei* continues to evolve within the host, possibly increasing its antibiotic resistance and host survival mechanisms. Alterations in the genome ranged from point mutations to a 330 kb deletion (Hayden *et al.*, 2012).

1.1.5.9 Other virulence factors

B. pseudomallei is robust and highly adaptable and many other factors contribute to its virulence. It expresses three acylhomoserine lactone quorum sensing

systems, including the BpsI-BpsR system (Gamage *et al.*, 2011). Quorum sensing plays an important role in the formation of biofilms which can prevent successful treatment with antibiotics. Quorum sensing and biofilm formation are particularly problematic for the treatment of chronic infections in cystic fibrosis patients. As this is more pertinent for other *Burkholderia* species (*B. cenocepacia* for example), a good review can be read by Eberl (2006).

Virulence factors can simply enable the bacteria to adapt to hostile or atypical environments. RelA and SpoT regulate stress response during starvation, are important for stationary-phase maintenance and for replication in macrophages. *B. pseudomallei* deletion mutants of RelA and SpoT also exhibited reduced virulence in the wax moth (*Galleria mellonella*) (Müller *et al.*, 2012).

Many of the *B. pseudomallei* virulence factors are poorly understood and characterisation of the pathogen is still ongoing. Recently, a transcription accessory factor, Tex, was identified as a potential virulence determinant and immunogen (Moule *et al.*, 2015).

1.2 Polysaccharides in *B. pseudomallei*

Polysaccharides are important structures found throughout nature in both eukaryotic and prokaryotic cells; including Gram-positive and Gram-negative bacteria. They are sugar polymers that can reach high molecular weights (up to 10^6 Da). They may be modified (acetylated or methylated, for example) and can be linear or consist of complex side-chains and branching. They can be further adapted by linkage with other molecules such as proteins and lipids.

Polysaccharides, usually either secreted or attached to a bacterial cell surface, perform many functions in bacterial survival and virulence. These include protection from desiccation (Roberson & Firestone, 1992; Ophir & Gutnick, 1994),

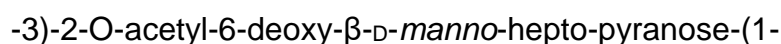
giving structure and cohesion to biofilms (Danese *et al.*, 2000; Ahimou *et al.*, 2007), protection from antimicrobials (Tamaki *et al.*, 1971; Roantree *et al.*, 1977; Snyder & McIntosh, 2000; Ding *et al.*, 2003; Campos *et al.*, 2004; Papo & Shai, 2005; Mohanram & Bhattacharjya, 2014), preventing complement-mediated killing and phagocytosis by the immune system (Tomas *et al.*, 1986; Marques *et al.*, 1992; Merino *et al.*, 1992; Clements *et al.*, 2008; Hyams *et al.*, 2010) and enabling attachment to host cells (Courtney *et al.*, 1996; Schragger *et al.*, 1998; Okamoto *et al.*, 2004; Hyams *et al.*, 2008).

1.2.1 Capsular polysaccharide

Initially, the CPS of *B. pseudomallei* was erroneously identified as a serotype of LPS and was designated Type I O-antigen (Knirel *et al.*, 1992; Perry *et al.*, 1995). However, CPS was later observed on the surface of LPS-deficient *B. pseudomallei* using immunogold electron microscopy (DeShazer *et al.*, 1998). It was subsequently determined a CPS by subtractive hybridisation (Reckseidler *et al.*, 2001) and isolated from LPS (Isshiki *et al.*, 2001). It is now recognised as an important virulence determinant and immunogen.

1.2.1.1 Structure

The CPS O-antigen is an unbranched homopolymer with the structure:



It has a 1 → 3 glycosidic link and is acetylated on the second carbon (Figure 6; Knirel *et al.*, 1992; Perry *et al.*, 1995).

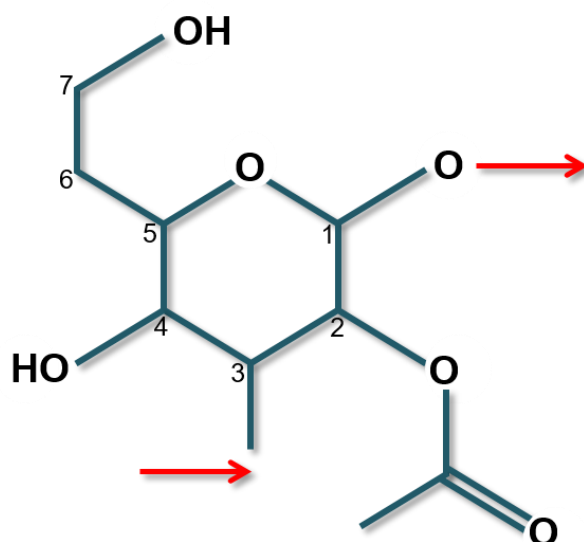


Figure 6: Structure of the repeated *manno*-heptose in *B. pseudomallei* CPS. Numbering for carbon atoms 1 to 7 is indicated. Red arrows indicate location of 1→3 glycosidic link that joins two monomers. There is an acetyl group on the second carbon.

The O-antigen is anchored to the outer membrane by a lipid. Though the structure of the lipid has not been elucidated, undecaprenyl phosphate is one proposed candidate (Cuccui *et al.*, 2012). The CPS is not associated with lipid A, differentiating it from LPS (Isshiki *et al.*, 2001). It is also likely that the lipid is primed with a mannose sugar before the *manno*-heptose polymer is assembled (Cuccui *et al.*, 2012). Therefore, three reactions are required for expression of the full CPS molecule; priming of the lipid anchor with a mannose sugar, linking of a *manno*-heptose to the mannose primer and finally polymerisation of *manno*-heptose (Figure 7).

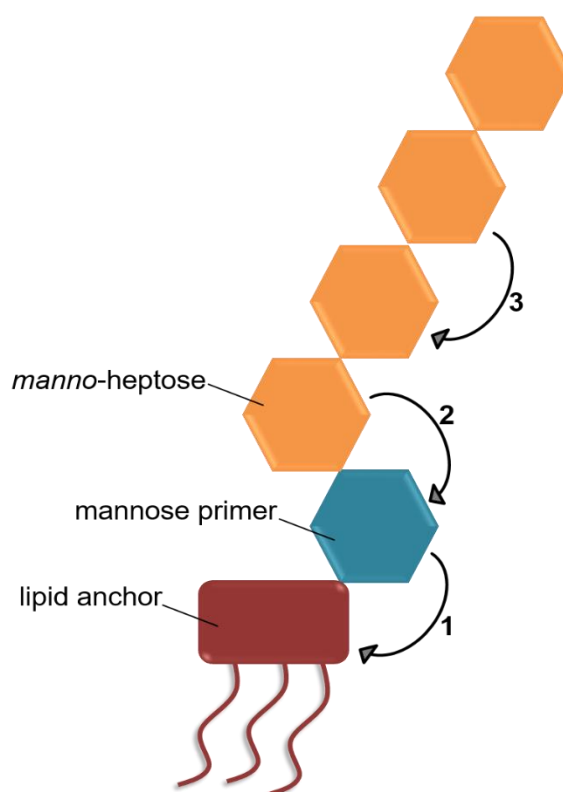


Figure 7: Structure of *B. pseudomallei* CPS. A lipid (red) anchors the polysaccharide to the outer membrane of the cell. There is evidence for a mannose priming sugar (blue) linking the lipid to the manno-heptose polymer (orange). Arrows indicate reactions in the proposed assembly of CPS. A mannose primer is added to the lipid (1), followed by linkage of *manno*-heptose to the mannose primer (2) and finally polymerisation of the *manno*-heptose (3).

The CPS is a high molecular weight molecule of around 200 kDa (Reckseidler-Zenteno *et al.*, 2005). It is expressed in high abundance, evenly distributed, across the cell surface of *B. pseudomallei* (Figure 8; DeShazer *et al.*, 1998). This viscous substance effectively coats the cell, forming a protective capsule.

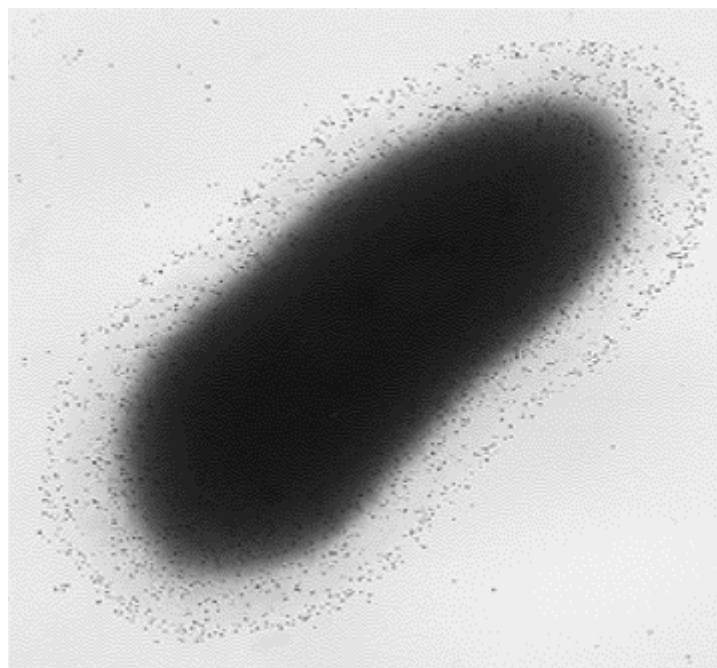


Figure 8: Immunogold electron microscopy of CPS expression in *B. pseudomallei* 1026b. The bacterial cell is surrounded by a thick layer of CPS. Each dot represents antibody bound to the CPS following exposure to rabbit anti-serum, washing, then goat anti-rabbit IgG immunogold conjugate. Magnification x 30,000. (Adapted and reproduced from DeShazer *et al.*, 1998).

1.2.1.2 Function

The CPS is required for full virulence in *B. pseudomallei* and several studies have demonstrated attenuation of acapsular mutant strains. The intraperitoneal LD₅₀ in Syrian golden hamsters was increased from 10² CFU for wild type *B. pseudomallei* 1026b to 10⁷ CFU for a CPS-deficient mutant (Reckseidler-Zenteno *et al.*, 2005). Similarly, intraperitoneal LD₅₀ increased from 10⁴ CFU for wild type *B. pseudomallei* 1026b to 10⁹ CFU for a CPS-deficient mutant in BALB/c mice (Wikraiphath *et al.*, 2009). Another study showed an acapsular strain was outcompeted by wild type *B. pseudomallei* 576 in Outbred Porton mice, with no mutant bacteria recoverable compared to 10⁹ CFU wild type. The acapsular mutant was cleared from mouse peritoneum by day 24 post-challenge, whilst wild type continued replicating (Atkins *et al.*, 2002b). However, all of the mice succumbed to infection by day 19 post-intravenous challenge, suggesting the acapsular strain was not fully attenuated (Atkins *et al.*, 2002b). This finding is

supported by the fact *B. thailandensis* E555, an unusual strain that expresses the *B. pseudomallei* CPS, was completely cleared from BALB/c mice 3-days post-intraperitoneal challenge (Scott *et al.*, 2013). Therefore, acquisition of the CPS alone is not sufficient to confer wild type *B. pseudomallei* virulence in *B. thailandensis*.

The capsule forms a physical barrier between the bacterial cell and the external environment, which includes the immune system of a host. The CPS was found to be important for survival and persistence in the blood (Reckseidler-Zenteno *et al.*, 2005). A 10^2 dose of CPS-deficient mutant and wild type *B. pseudomallei* 1026b was administered intraperitoneally to Syrian golden hamsters. Whilst the numbers of recoverable wild type bacteria continued to increase in all tissues sampled (blood, lung, liver and spleen), the numbers of CPS-deficient mutant remained low. After 48 h, recovery *B. pseudomallei* ranged from 10^4 CFU in the lungs to 10^6 CFU in the spleen for wild type compared to 10 CFU in the blood to 10^2 CFU in the spleen for CPS-deficient mutant (Reckseidler-Zenteno *et al.*, 2005). This is thought to be due to the capsule preventing C3b deposition on the bacterial cell surface. The C3b protein triggers the complement pathway of immune system which is highly effective at pathogen killing.

Reckseidler-Zenteno *et al.* (2005) examined C3b deposition on wild type *B. pseudomallei* 1026b and a CPS-deficient mutant. They observed 3.5-fold and 2.5-fold increases in C3b deposition on the CPS-deficient strain in 10 % and 30 % NHS respectively compared to wild type *B. pseudomallei*. Deposition of C3b also occurred more quickly on the CPS-deficient strain than the wild type bacteria. Furthermore, the capsule can also prevent killing by phagocytosis. Uptake of CPS-deficient *B. pseudomallei* by polymorphonuclear leukocytes was 51.7 % compare to 35.9 % for wild type *B. pseudomallei* in 10 % NHS. Similarly, in 30 %

NHS, the rate of phagocytosis was 82.3 % for the CPS-deficient mutant compared to 59.3 % for wild type *B. pseudomallei* (Reckseidler-Zenteno et al., 2005). This is probably due to the physical size and viscosity of the CPS making the process of phagocytosis more challenging as well as reducing opsonisation of the bacteria which recruits professional phagocytes

Another function that may be mediated by CPS is adhesion to the surface of host epithelial cells. Ahmed *et al.* (1999) observed a polysaccharide layer attaching *B. pseudomallei* to host pharyngeal epithelial cells by electron microscopy. *B. pseudomallei* is also more able to adhere to cells than *B. thailandensis*, which normally lacks the capsule (Kespichayawattana *et al.*, 2004). Additionally, Sarkar-Tyson *et al.* (2007) found the CPS influences the surface hydrophobicity of *B. pseudomallei* which may have importance in attachment to surfaces including host cells (Vercauteren *et al.*, 1993).

1.2.1.3 Coding locus

Genes encoding the CPS in *B. pseudomallei* are clustered on a single 34.5 kb locus on chromosome 1 (Holden *et al.*, 2004; Cuccui *et al.*, 2007; 2012). Homologous genes from the locus are present in pathogenic *B. mallei* (DeShazer *et al.*, 2001) but normally absent from 'avirulent' *B. thailandensis* (Kespichayawattana *et al.*, 2004; Ong *et al.*, 2004; Yu *et al.*, 2006). However, *B. thailandensis* E555, which is genetically distinct from other *B. thailandensis* strains, possesses the CPS coding locus and expresses the capsule (Sim *et al.*, 2010).

There are 26 genes on the coding locus, with functions in the biosynthesis of a nucleotide activated sugar precursor and the lipid anchor, glycosyltransferases

for polymer assembly and chaperones/transporters for localisation of the CPS (Figure 9; Cuccui *et al.*, 2012).

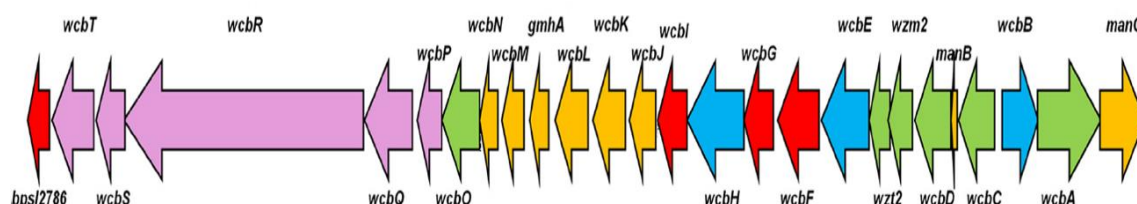


Figure 9: CPS coding locus in *B. pseudomallei*. Genes have functions in nucleotide sugar biosynthesis (yellow); polymer assembly (glycosyltransferases; blue); lipid biosynthesis (purple) and transportation (green). There are three genes of unknown function (red). Adapted and reproduced from Cuccui *et al.* (2012).

Previously, attempts were made to clone the entire operon from *B. pseudomallei* and express it in *E. coli* (data not shown). However, this approach was not successful in producing any CPS. For this project, only essential genes from the operon were selected for cloning into *E. coli*, with the aim being to exploit existing mechanisms in *E. coli* to express the CPS. The genes chosen were *wcbJ*, *wcbK*, *wcbL*, *gmhA*, *wcbM* and *wcbN* for the biosynthesis of the activated sugar nucleotide precursor GDP-6-deoxy-D-manno-heptose and glycosyltransferases *wcbB*, *wcbE* and *wcbH* for polymer assembly.

1.2.1.4 GDP-6_dHep biosynthetic pathway

Seduheptulose-7-phosphate (S7P) is diverted from the pentose phosphate pathway into the production of GDP-6_dHep (Figure 10).

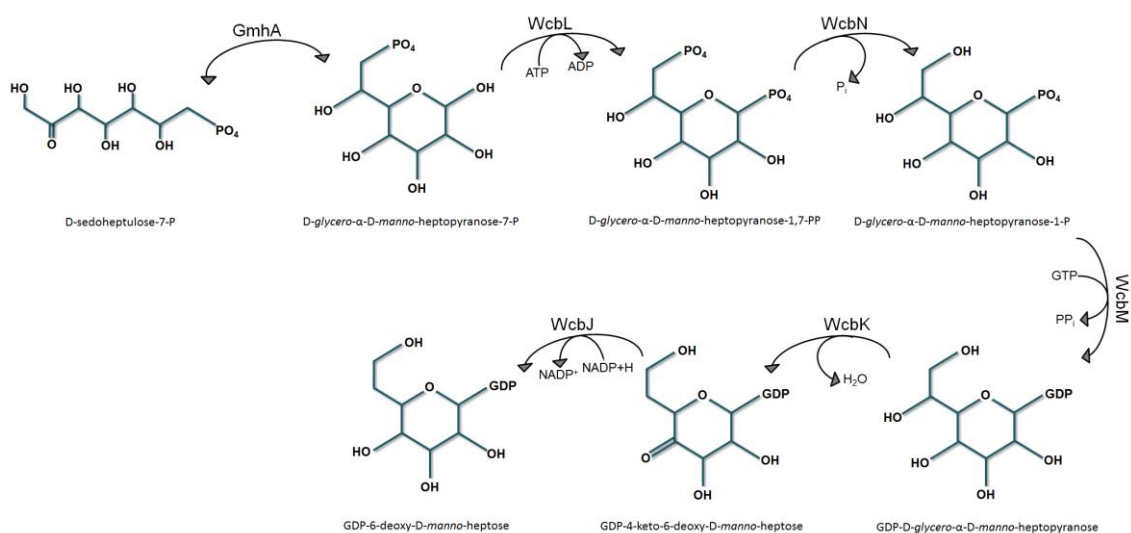


Figure 10: Biosynthetic pathway for GDP-6dDHep. Six biosynthetic enzymes convert sedoheptulose-7-phosphate into the activated nucleotide sugar precursor for *B. pseudomallei* CPS production. Intermediate sugars are shown, the enzyme responsible for each reaction and any by-products are also indicated. P_i = inorganic phosphate; PP_i = pyrophosphate.

Orthologous pathways are found in *Y. pseudotuberculosis* H892/87 and *Campylobacter jejuni* (Pacinelli *et al.*, 2002; Karlyshev *et al.*, 2005). Enzymes from the different organisms can be used in the same reaction pathway to produce GDP-6dDHep from S7P (Butty *et al.*, 2009) GmhA orthologous are found in other virulent species including *C. jejuni*, *Vibrio cholerae*, *Pseudomonas aeruginosa*, *E. coli* and *Y. pseudotuberculosis* (Pacinelli, *et al.*, 2002; Karlyshev, *et al.*, 2005; Seetharaman *et al.*, 2006; Taylor *et al.*, 2008).

Figure 10 depicts the biosynthetic pathway described briefly here: The first step is catalysed by GmhA, a metalloenzyme, requiring zinc in the active site (Harmer, 2010). It interconverts S7P with D-glycero- α -D-manno-heptopyranose-7-phosphate. Next, heptokinase WcbL converts M7P into D-glycero- α -D-manno-heptopyranose-1,7-bisphosphate with a 1-O-phosphosphorylation (Vivoli *et al.*, 2015). WcbN, a GmhB orthologue, then dephosphorylates carbon 7 of D-glycero- α -D-manno-heptopyranose-1,7-bisphosphate to produce D-glycero- α -D-manno-heptopyranose-1-phosphate and releases an organic phosphate (Karlyshev, *et*

al., 2005; Butty *et al.*, 2009). WcbM, a GmhC orthologue, is a guanosyltransferase, adding GDP to carbon 1 and producing pyrophosphate (Karlyshev, *et al.*, 2005; Butty *et al.*, 2009). The resulting GDP-D-*glycero*- α -D-*manno*-heptopyranose is then dehydrated by WcbK, which is orthologous to DmhA, producing GDP-4-keto-6-deoxy-D-*manno*-heptose and water (Butty *et al.*, 2009). Finally, WcbJ, a DmhB reductase orthologue, generates GDP-6dDHep and converts NADP+H to NADP⁺ (Butty *et al.*, 2009).

1.2.1.5 Glycosyltransferase mediated assembly of CPS

Once the GDP-6dDHep precursor has been synthesised, it is polymerised into the final CPS structure. (This may occur before or after the 2-O-acetylation step. There is some indication that WcbI may be the acetyltransferase (Vivoli *et al.*, 2014). Assembly of the CPS is mediated by glycosyltransferases; an important class of enzyme found in abundance throughout nature. Glycosyltransferases covalently attach activated (high energy) donor sugars to acceptor molecules via a glycosidic link. Leloir-type glycosyltransferases, which use nucleotide sugars as donor substrates, are involved in the biosynthesis of CPS and other polysaccharides (Leloir, 1971; Frey, 1996; Timmons & Jackeman, 2008) Acceptor molecules may be other sugars (including simple monomers to a complex branched/modified polysaccharides), lipids, proteins or structures such as flagella. There are 96 glycosyltransferase families, representing the diversity of reactions they perform (Campbell *et al.*, 1997; Coutinho *et al.*, 2003)..

There are three putative glycosyltransferases in the *B. pseudomallei* CPS coding locus; *wcbB*, *wcbE* and *wcbH*. They are annotated as either belonging to glycosyltransferase family 1 or family 4, depending on the database (UniProt (<http://www.uniprot.org/>) and CAZy (<http://www.cazy.org/>) respectively)

(Leinonen *et al.*, 2004; Lombard, V. *et al.*, 2013). However, there is consensus that structurally, they contain a GT-B fold, which is characterised by two similar $\beta/\alpha/\beta$ Rossmann-like domains facing together to form the active site (Breton *et al.*, 2006; Albesa-Jové *et al.*, 2014).

Their proposed functions are firstly to attach a mannose priming sugar to a lipid anchor, to link a *manno*-heptose (GDP-6dDHep precursor) to the mannose primer and to finally polymerise the GDP-6dDHep (Figure 7). The activities for these enzymes have not been elucidated, though pairwise BLAST analysis suggests WcbE may be a mannosyltransferase and therefore responsible for attaching the priming sugar to the lipid anchor (Cuccui *et al.*, 2012).

1.2.1.6 Capsule assembly

As there are no chain length determinants or polymerases on the coding locus, Cuccui *et al.* (2012) proposed a group 3 method of CPS assembly (Whitfield & Roberts, 1999; Whitfield, 2006) (Figure 11). GDP-6dDHep is synthesised in the cytosol and an unknown lipid is likely synthesised by WcbP/Q/R/S/T. Putative transporters WcbA/WcbO then direct glycosyltransferases WcbB, WcbE and WcbH to assemble nascent polysaccharide onto the lipid anchor on the inside of the inner membrane. Transporters wzt2 and wzm2 relocate the nascent CPS to the outer leaflet of the outer membrane, assisted by WcbD and WcbC. The lipid anchors the CPS to surface of the cell. (Genes for putative enzymes are annotated on the CPS coding locus in Figure 9).

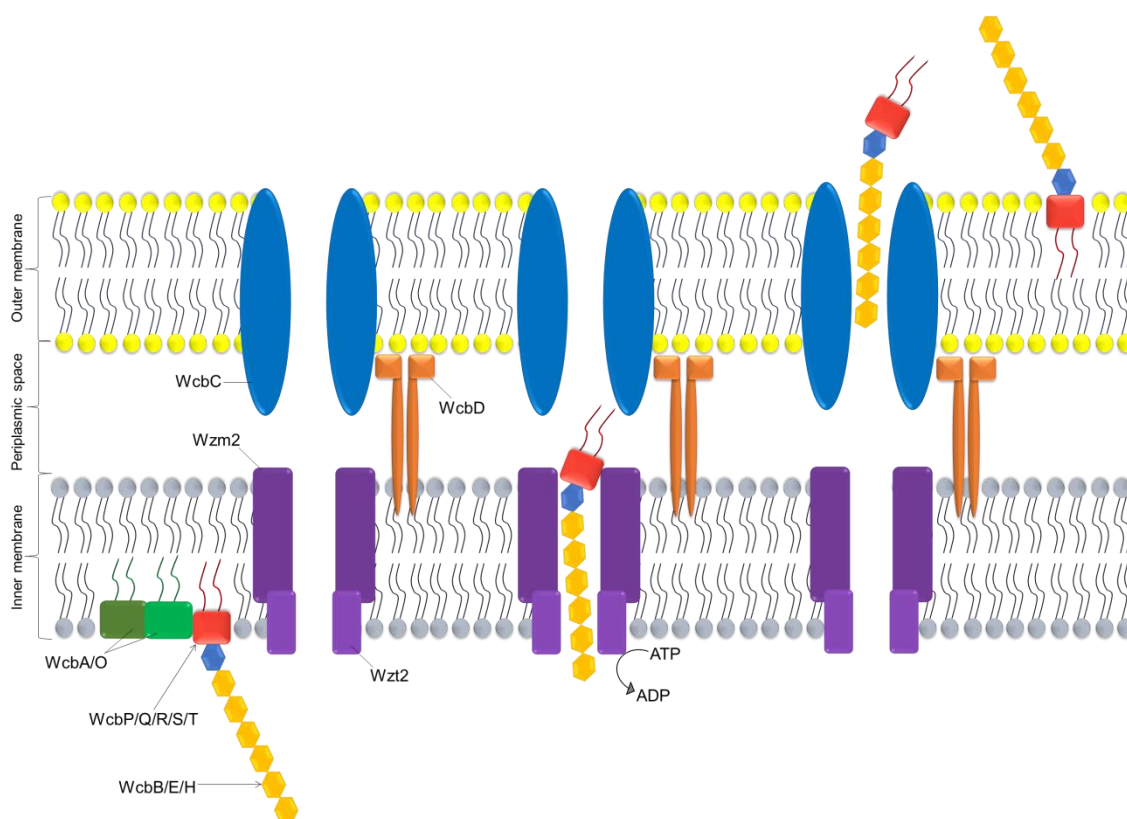


Figure 11: Proposed model for the assembly of CPS on the surface of *B. pseudomallei*. Nascent CPS is synthesised in the cell and then transported across the inner and outer membranes and is anchored to the outside of the cell. Adapted from Cuccui *et al.* (2012).

1.2.2 Lipopolysaccharide

LPS is an integral component of the outer membrane in Gram negative bacteria. It is found in high abundance in the outer leaflet and is exposed to the extracellular environment. LPS consists of a Lipid A anchor embedded in the membrane, a core domain (which may have different inner and outer components) and an O-antigen polysaccharide (Figure 12; Kastowsky *et al.*, 1992).

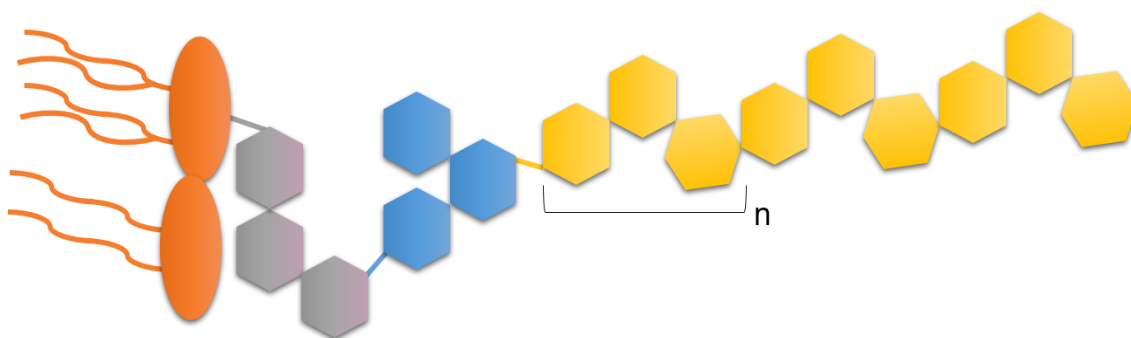


Figure 12: Generic structure of LPS. Lipid A (orange) anchors LPS into the outer membrane of Gram negative bacteria. A core domain, consisting of inner (grey) and outer (blue) fragments, connects the Lipid A to the O-antigen (yellow). The O-antigen is a polymer of repeated monosaccharides (n).

Within a species, the lipid A and core saccharide domains are usually well conserved. (There are exceptions; modifications in lipid A of *Neisseria meningitidis* and *S. flexneri* greatly influence the immune response to infection (Fransen *et al.*, 2010; Paciello *et al.*, 2013)). However, the O-antigen is often highly variable in length and composition as observed in *E. coli* and *Salmonella enterica* (Heinrichs *et al.*, 1998). It can also be absent from the LPS altogether (known as rough as opposed to smooth LPS) as exhibited by *Chlamydia* species (Lukáčová *et al.*, 1994). Unusually, there have only been a small number of LPS variations observed in *B. pseudomallei*. Designated Type II O-Polysaccharide, the most abundant LPS O-antigen is an unbranched polymer of D-glucose and L-talose monosaccharides (Knirel *et al.*, 1992; Perry *et al.*, 1995). Observed variations occur only in the L-talose monomers. Position 2 can be unsubstituted, acetylated or methylated, position 3 is either glycosylated to β -glucose or methylated and position 4 can be unsubstituted or acetylated (Figure 13; Perry *et al.*, 1995; Heiss *et al.*, 2013).

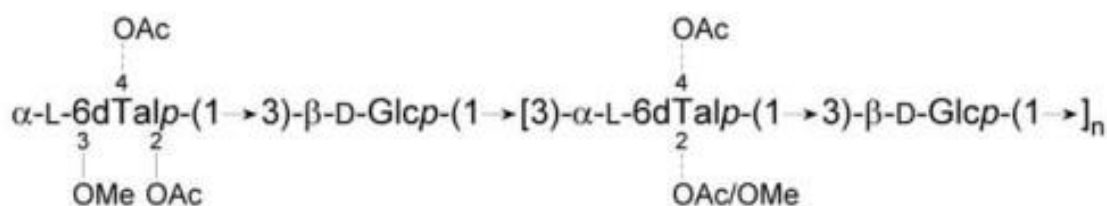


Figure 13: O-antigen of *B. pseudomallei* LPS showing possible sites of acetylation or methylation (From Heiss *et al.*, 2013).

The standard O-antigen was detected in 97 % of 1,327 isolates screened by Anuntagool *et al.* (2006). They found an atypical smooth serotype in 2 % of the isolates and a rough serotype in the remaining 1 %. The atypical smooth and rough serotypes were almost exclusively recovered from clinical samples, suggesting these modifications occur as an adaptation to the host environment Anuntagool *et al.* (2006). A further atypical smooth serotype, observed in only 7 clinical isolates, has also been identified (Tuanyok *et al.*, 2012). In addition to the effect on virulence, variation in the O-saccharide of LPS can have an effect on colony morphology (Wikraiphath *et al.*, 2015). Understanding the link between colony morphology and virulence could be an important tool in future diagnosis of melioidosis (see section 1.1.4.2).

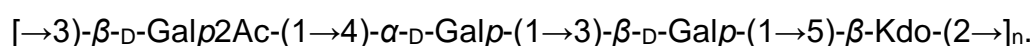
LPS is a virulence factor with roles in evading the innate immune system. LPS-deficient *B. pseudomallei* lost its resistance to killing when exposed to 30 % normal human serum (NHS) compared to wild type 1026b (DeShazer *et al.*, 1998). The LPS-deficient strain was subsequently found to have a 10-fold higher LD₅₀ in Syrian golden hamsters and guinea pigs and 100-fold higher LD₅₀ in diabetic rats compared to wild type *B. pseudomallei* 1026b. (DeShazer *et al.*, 1998). This study showed that LPS plays an important role in resistance to serum killing and is LPS-deficient *B. pseudomallei* is attenuated in three model organisms. LPS-deficient *B. pseudomallei* were also less able to invade and

replicate in murine and human macrophages than wild type 1026b. The mutant was also more susceptible to killing by NHS, antimicrobial peptides and reactive nitrogen and oxygen intermediates (Arjcharoen *et al.*, 2007; Wikraiphath *et al.*, 2009).

As well as an important virulence factor, LPS is also extremely immunogenic. Known as endotoxin, it can cause toxic shock in a host with septicaemia (Galanos & Freudenberg, 1993). The lipid A domain has been highlighted as the major factor in toxic shock (Khan *et al.*, 1998) Use of LPS as a potential vaccine candidate is explored in section 1.3.4.2.

1.2.3 Minor exopolysaccharides

Several other polysaccharides have been observed in *B. pseudomallei*. In separate studies, Masoud *et al.* (1997) and Nimtz *et al.* (1997) identified a linear, high molecular weight tetrasaccharide consisting of three galactose monomers (one acetylated) and 3-deoxy-D-manno-octulosonic acid (Kdo):



Nimtz *et al.* (1997) suggest the exopolysaccharide is expressed by all *B. pseudomallei* strains. Masoud *et al.* (1997) identified the polysaccharide in a clinical isolate of *B. pseudomallei* 304b in Thailand, suggesting the polysaccharide may contribute to virulence. More recently, the same polysaccharide was isolated from *B. cepacia* in a cystic fibrosis patient (Cescutti *et al.*, 2003). However, the genes responsible for this polysaccharide are yet to be identified (Sarkar-Tyson *et al.*, 2007).

Gene clusters encoding two additional polysaccharides, Type III O-polysaccharide and Type IV O-polysaccharide, have been located on chromosome 2 (Holden *et al.*, 2004).

The Type III O-polysaccharide is not a virulence determinant. Sarkar-Tyson *et al.* (2007) demonstrated that BALB/c mice challenged with Type III O-polysaccharide-deficient *B. pseudomallei* did not have an improved survival rate compared with mice challenged with wild type K96243. This was later verified by Reckseidler-Zenteno *et al.* (2010) who demonstrated there was no increase in LD₅₀ for Type III O-polysaccharide-deficient *B. pseudomallei* compared to wild type 1026b in Syrian golden hamsters. In contrast, deletion of the CPS caused the LD₅₀ to increase from <10 CFU for wild type *B. pseudomallei* 1026b to 3.5 x 10⁵ CFU in the mutant.

The Type IV O-polysaccharide may contribute to virulence in *B. pseudomallei*. In the same study by Sarkar-Tyson *et al.* (2007), *B. pseudomallei* deficient in Type IV O-polysaccharide were attenuated in BALB/c mice. The mean time to death (MTTD) increased from 3 days to 11.6 days when challenged with wild type *B. pseudomallei* K96243 compared to the mutant respectively. Neither Type III or Type IV O-polysaccharides affected the hydrophobicity of *B. pseudomallei*, suggesting any influence on virulence they may have is limited (Vercauteren *et al.*, 1993; Sarkar-Tyson *et al.*, 2007).

Finally, glycosylation of other structures within *B. pseudomallei* may affect survival and virulence determination. Scott *et al.*, 2011 showed that glycosylation of flagella in *B. pseudomallei* is required for full motility (2.52 cm halo diameter in non-glycosylated mutant compared to 6.92 cm halo diameter in wild type K96243). The glycan modification was only 291 Da in size, suggesting there may

be many other examples of functional glycosylations in *B. pseudomallei* that have not yet been identified Scott *et al.* (2011).

1.3 Vaccine development

Vaccination has proven to be an effective method of reducing the burden of disease across the globe. Rigorous immunisation programmes have resulted in the eradication of smallpox (Henderson, 2011) and the eradication of polio is imminently achievable. Only 25 cases of poliomyelitis were reported in 2015, isolated to Afghanistan and Pakistan (Cochi *et al.*, 2016). Vaccination has also successfully morbidity and mortality caused by other viral pathogens including measles, mumps, rubella, varicella, hepatitis A and hepatitis B (Roush & Murphy, 2007). As viruses are obligate intracellular pathogens, potential reservoirs and routes of transmission are limited. This enables dissemination of the disease to be monitored and controlled more easily than for opportunistic bacterial pathogens from the environment. However, the success of these vaccines demonstrates immunisation can provide effective prophylaxis against intracellular pathogens. Vaccines against bacterial pathogens have also proven effective at reducing the burden of disease. Examples include vaccines against *N. meningitidis*, *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Bordetella pertussis*, *Clostridium tetani* and *Corynebacterium diphtheria* (Roush & Murphy, 2007).

Treatment of *B. pseudomallei* infection is challenging due to inherent resistance to a broad spectrum of antibiotics, recurring persistent and chronic infections and the emerging resistance to frontline antibiotic ceftazidime (Randall *et al.*, 2015). The limited therapeutic options and a low-infective dose, particularly via aerosol, has resulted in the classification of *B. pseudomallei* as a potential bioterror agent

(Gilligan, 2002; Butler, 2012). Melioidosis is also a significant public health risk, endemic in tropical regions across the globe (Limmathurotsakul *et al.*, 2016). For these reasons, there is considerable interest in the development of a prophylactic vaccine for both military and civilian populations.

Several factors must be considered in the design and development of a vaccine, including the host immune response and the antigen used. Here, some of the factors in vaccine development and the potential strategies for immunisation against *B. pseudomallei* are examined.

1.3.1 Animal models of infection

The animal model used in testing a novel antigen or vaccine formulation must be carefully considered, as susceptibility and immune response can differ between species and subspecies. Several animal models have been used for *B. pseudomallei* research, alongside *in vitro* experiments and human serological studies.

The most commonly used models for *B. pseudomallei* infection are the BALB/c and C57BL/6 murine models. BALB/c mice are highly susceptible to *B. pseudomallei*, with an LD₅₀ of 4 to 45 CFU, high bacterial loads recoverable from the lungs, spleen, liver and blood, a strong gamma interferon (IFN- γ) response and low survival times post-inoculation (Leakey *et al.*, 1998; Liu *et al.*, 2002). In contrast, C57BL/6 mice are more resistant, with an LD₅₀ of 4.8×10^3 – 2.5×10^4 CFU, lower bacterial loads recoverable from tissues and blood, protective antibody response and asymptomatic survival to week 6 post-inoculation (Leakey *et al.*, 1998; Liu *et al.*, 2002). For these reasons, BALB/c mice were characterised as a model for acute infection and C57BL/6 mice as a model for chronic infection.

Other, less commonly used mammalian models are guinea-pigs, Syrian golden hamsters and diabetic rats (DeShazer *et al.*, 1997; 1999; Warawa & Woods, 2005; Tuanyok *et al.*, 2006). Of these, diabetic rats are most resistant to *B. pseudomallei* infection via intraperitoneal inoculation, with a 10-fold higher LD₅₀ than guinea-pigs and hamsters (DeShazer *et al.*, 1999). Less commonly used models are goats (Gonzalez-Juarrero *et al.*, 2013) and marmosets (Nelson *et al.*, 2011; 2014)

Non-mammalian models for *B. pseudomallei* infection include single eukaryotic cells such as macrophages and amoebae (*A. astronyxis* and *Dictyostelium discoideum*) for examining the mechanisms of phagocytosis and phagosome escape (Inglis *et al.*, 2000b; 2003; Hasselbring *et al.*, 2011; Wand *et al.*, 2011). Simple multicellular eukaryotes such as nematodes (*C. elegans*) and the wax moth larvae (*Galleria mellonella*) can be used to examine the relative virulence of different strains and mutants of *B. pseudomallei* and related species such as *B. thailandensis* (O'Quinn *et al.*, 2001; Wand *et al.*, 2011).

1.3.2 Routes for vaccine administration

The route of administration greatly influences the infectious dose, host immune response and type of disease caused if an infection is established. For *B. pseudomallei*, studies by Nelson *et al.* (2011; 2014) determined infective doses of 26 CFU when administered subcutaneously and fewer than ten CFU via intranasal aerosol in marmosets. Similarly, Lever *et al.* (2009) demonstrated BALB/c mice are more susceptible to *B. pseudomallei* infection via the aerosol route (5 CFU MLD) than intraperitoneally (54 CFU MLD). The increased susceptibility to infection via the intranasal route and the threat of use of aerosolised *B. pseudomallei* as a biowarfare agent has led to an interest in the

development of an intranasal vaccine. Whilst most vaccines are administered intramuscularly or orally, a live-attenuated intranasal vaccine against influenza is used to immunise young children (Belshe *et al.*, 2007). An important feature of immunisation via aerosol is the production of immunoglobulin (Ig)A; a high-affinity antibody secreted in the mucous membranes. Increased IgA production was observed in BALB/c mice challenged with *B. pseudomallei* via aerosol (Liu *et al.*, 2002). This may be vital for preventing *B. pseudomallei* invasion of epithelial cells in the respiratory tract

1.3.3 Host immune response

The immune system is a complex matrix of interactions between innate and adaptive responses for recognising and eliminating an invading pathogen. Here, a very brief overview is provided in the context of *B. pseudomallei* infection. Murphy & Casey (2016) provide a more in-depth review of the innate and adaptive immune systems.

1.3.3.1 Innate immunity

Upon entering the host, a pathogen is confronted by the innate immune system. This frontline defence is non-specific to the pathogen and has no memory of previous encounters. However, it can be very effective as it generates an immediate response upon recognition of a pathogen and can rapidly recruit an arsenal of leucocytes to neutralise a perceived threat. The innate immune system can also activate and mediate the adaptive immune response.

Pathogen detection

The innate immune system is activated by recognition of pathogen-associated molecular patterns (PAMPs). Common PAMPs include LPS, lipid A, peptidoglycan, flagellin and T3SS, all of which are present in *B. pseudomallei* (Wiersinga *et al.*, 2006) PAMPs are detected by receptors, which may be secreted or expressed by immune cells. For example, toll-like receptors (TLRs) are expressed on the surface of splenic tissue and peripheral leucocytes. Increased levels of TLR-1, TLR-2 and TLR-4 expression on monocytes and granulocytes has been observed in melioidosis patients compared to healthy individual (Wiersinga *et al.*, 2007). It was subsequently shown that recognition of *B. pseudomallei* lipid A and LPS results in TLR-4-dependent activation of both murine and human natural killer cell lines *in vitro* (West *et al.*, 2008).

Leucocyte recruitment

Activation of a PAMP receptor results in the release of chemokines and cytokines that recruit leucocytes (such as macrophages, neutrophils, monocytes and granulocytes) to the site of infection. Recognition of *B. pseudomallei* LPS by TLR-4 results in release of the cytokine tumour necrosis factor alpha (TNF- α), which predominately recruits macrophages (West *et al.*, 2008). Another important cytokine, IFN- γ , recruits macrophages and neutrophils. Easton *et al.* (2007) demonstrated IFN- γ knock-out C57BL/6 mice generated lower levels of alveolar neutrophils post-intranasal challenge with *B. pseudomallei* 576 than wild type mice. This resulted in all the deficient mice succumbing to the infection on day 3 post-challenge, whereas 50 % of wild type mice survived to day 40 post-inoculation.

Pathogen destruction

Recruited neutrophils are responsible for the further release of cytokines, thereby proliferating the immune response. Expression of cytokines TNF- α , IFN- γ and interleukin 6 (IL-6) is reduced 98 %, 88 % and 91 % respectively in neutrophil-depleted mice (Easton *et al.*, 2007). In addition to cytokine release, neutrophils can form extracellular traps that create a chromatin barrier, reducing dissemination of the pathogen and causing their destruction by release of granular proteins (Riyapa *et al.*, 2012)

Macrophages, dendritic cells, monocytes and neutrophils are able to internalise pathogens via phagocytosis. The process is often assisted by opsonisation of the pathogen via the complement system. However, *B. pseudomallei* resists phagocytosis by the expression of CPS, which reduces complement deposition on the bacterial cell surface (Reckseidler-Zenteno *et al.*, 2005). Once internalised, the pathogen is usually destroyed via reactive oxygen and nitrogen species and NAPH-oxidase pathways (Breitbach *et al.*, 2006; Willcocks *et al.*, 2016). However, *B. pseudomallei* is able to escape from the phagosome using the T3SS and survive within the cytosol. Gong *et al.* (2011) describe evidence of autophagy in the macrophage RAW 264.7 cell-line as a defence against escape of *B. pseudomallei* from the phagosome. If the pathogen is successfully degraded, the phagocyte can process antigen and present it on the cell surface to activate the adaptive immune response.

1.3.3.2 Adaptive immune system

The adaptive immune response is highly pathogen-specific, effective at pathogen killing and can generate long-term protection from subsequent exposure to an encountered pathogen. However, unlike the innate immune system which

constantly patrols and monitors the host environment for potential targets, the adaptive immune system is slow to activate upon first encounter. Moreover, it is heavily influenced by assistance provided by the innate immune response, which regulates cytokine release and antigen presentation.

Antigen presentation

A pathogen phagocytosed by an antigen presenting cell (e.g. macrophage) becomes trapped within a phagosome and is degraded by oxidative pathways. Fragments from the destroyed pathogen are then packaged and transported to the cell surface for presentation by the major histocompatibility complex (MHC). Different MHC classes and co-stimulants are expressed depending upon the source (intracellular or extracellular) of the pathogen. Antigens are primarily presented to T lymphocytes, which mature and can perform a number of functions. The two principal types of mature T lymphocyte are helper and cytotoxic cells. Neefjes *et al.* (2011) have reviewed the process of antigen presentation via MHC.

B. pseudomallei is capable of escaping the phagosome to evade the antigen presentation process (Yates *et al.*, 2005; Haas, 2007; Gong *et al.*, 2011). However, antibodies against *B. pseudomallei* are found in the sera of melioidosis patients and in human populations in endemic regions (Charuchaimontri *et al.*, 1999; Vasu *et al.*, 2003; Cheng *et al.*, 2008; Rolim *et al.*, 2011). This suggests there is scope for an effective vaccine against *B. pseudomallei*.

Humoral immune response

Presentation of antigen from an extracellular pathogen via MHC class II, with co-stimulation, causes maturation of helper T (CD4+) lymphocytes. These matured

lymphocytes assist in the differentiation of B lymphocytes into plasma (antibody-producing) and memory B cells. This is crucial for generation of long-lasting, pathogen-specific immunity (Kurosaki *et al.*, 2015). The MST of CD4+ lymphocyte-depleted C57BL/6 mice challenged intraperitoneally with *B. pseudomallei* 576 was increased significantly from 22 days to 58 days when supplemented with antibody (Haque *et al.*, 2006a). This shows the humoral immune response is important for protection against melioidosis.

B lymphocytes can also sample antigen, independent of helper T lymphocytes, and differentiate into plasma and memory cells. This is particularly important in recognition of non-protein antigen, such as polysaccharides and DNA (Puangpetch *et al.*, 2014). A T lymphocyte-independent response is thought to be produced by crosslinking of B cell receptors by the polysaccharide, stimulating a strong signal. The B lymphocytes can then differentiate to plasma and memory cells (Defrance *et al.*, 2011; Mitchell *et al.*, 2014).

There are five main classes of antibody, each structured to perform different functions (Figure 14; Murphy & Casey, 2016). The most important antibodies in the humoral immune response to *B. pseudomallei* are IgM, IgG and IgA (Chenthamarakshan *et al.*, 2001; Vasu *et al.*, 2003).

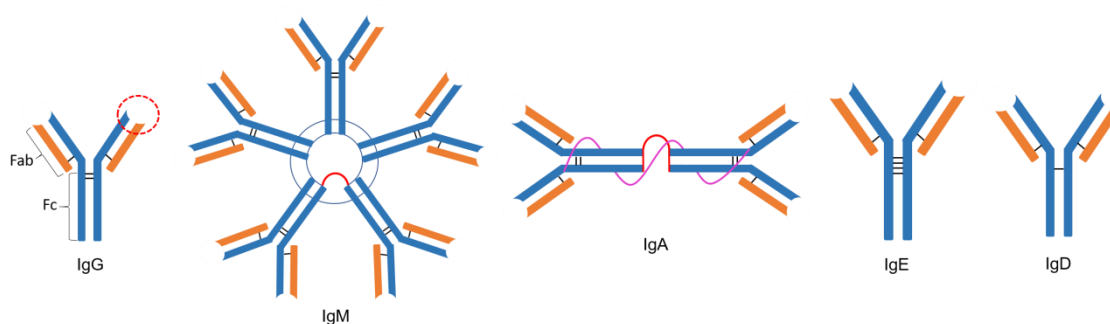


Figure 14: Basic structure of the five major antibody classes. Showing the heavy (blue) and light (orange) chains, interchain disulphide bonds (black), joining (J)-chain (red) and secretory component (purple). The highly variable antigen binding site (red dotted circle), variable antibody-binding (Fab) and constant (Fc) regions are indicated on IgG.

The first antibody produced against a newly-encountered pathogen is the IgM pentamer (Figure 14). It is the largest antibody expressed and is comparatively poor at opsonisation. Although IgM has low affinity to specific antigen, its high avidity is useful for binding antigens in a naïve immune system. Conversion of B cells into antigen-specific plasma cells results in the production of high affinity monomeric IgG (Figure 14; Nutt *et al.*, 2015). IgG antibody is excellent at opsonisation, causing agglutination and immobilisation of pathogens in addition to inducing phagocytosis. Titres of both IgM and IgG are significantly raised in patients with septic, localised and clinically-suspected melioidosis (Chenthamarakshan *et al.*, 2001a; 2001b).

The third antibody class important in melioidosis is dimeric IgA (Figure 14). It is secreted in mucosal tissues, including the respiratory tract and is important in resistance to intranasal pathogens. Raised IgA titres have been detected in melioidosis patients (Chenthamarakshan *et al.*, 2001a; Vasu *et al.*, 2003). IgA is an important consideration in developing a vaccine as aerosolised *B. pseudomallei* is a military concern (Gilligan, 2002; Butler, 2012).

To date, there is no evidence to suggest either IgE or IgD have significant functions in the immune response to *B. pseudomallei* (Figure 14). IgE is present on the skin and is important for parasite opsonisation, whilst IgD is expressed on mast cells and is comparatively poorly characterised (Murphy & Casey, 2016).

Cellular immune response

Intracellular pathogens, such as *B. pseudomallei*, increase survival by evading the humoral immune response. To eliminate these invaders, infected host cells must be identified and either induced to destroy the pathogen (e.g. neutralisation in macrophages; Breitbach *et al.*, 2006) or to trigger apoptosis. Presentation of

antigen from an intracellular pathogen via MHC class I results in maturation of cytotoxic (CD8+) T lymphocytes (Neefjes *et al.*, 2011). The matured lymphocytes identify host cells presenting antigen specific to intracellular pathogens and induce apoptosis (Zhang & Bevan, 2011).

To propagate the cellular immune response, helper T lymphocytes can be differentiated into Th1 cells which produce IFN- γ , a potent cytokine for macrophages and cytotoxic T lymphocytes. Exposure to *B. pseudomallei* antigens stimulated IFN- γ production by natural killer cells in samples from seropositive humans (Tippayawat *et al.*, 2009).

Cellular immune response has been observed in melioidosis patients and in healthy individuals with previous exposure to *B. pseudomallei*. Re-exposure to *B. pseudomallei* antigens resulted in increased IFN- γ levels and greater proliferation of helper and cytotoxic T lymphocytes and in both groups compared with naïve control subjects (Ketheesan *et al.*, 2002; Govan & Ketheesan, 2004; Tippayawat *et al.*, 2009).

Lymphocyte proliferation and IFN- γ levels were found to be greater in healthy seroconverted individuals when compared to current melioidosis patients, suggesting the cellular response is integral to generating protective immunity (Barnes *et al.*, 2004). A more recent study observed generation of memory effector T cells that can mediate activation of the cellular immune response upon re-exposure to *B. pseudomallei* antigen (Tippayawat *et al.*, 2009). These data suggest the cellular immune response is vital to protect against *B. pseudomallei* infection and melioidosis and must be considered carefully in the development of a prophylactic vaccine.

1.3.4 Immunisation strategies

1.3.4.1 Whole-cell vaccination

Traditional approaches to immunisation against bacteria comprised of whole-cell killed or live-attenuated vaccines. Notable successful examples include killed *Yersinia pestis* (Titball & Williamson, 2004) and attenuated *Salmonella enterica* serovar Typhi (Rao, 1991) vaccines. The benefit of whole-cell vaccines is that a large number of antigens are presented to the immune system, generating immune responses to several targets. However, they also present a risk of excessive or inappropriate immune responses and, for attenuated vaccines, the potential to revert to pathogenicity.

Inactivated/killed whole cell

A study by Sarkar-Tyson *et al.* (2009) used heat killed *B. pseudomallei* (strains K96243 and 576), *B. mallei* and *B. thailandensis* to immunise BALB/c mice. Relative antibody titres recovered from the mice were 100 fold higher for both *B. mallei* and *B. thailandensis* compared with both strains of *B. pseudomallei*. The mice were subsequently challenged with an intraperitoneal dose of *B. pseudomallei* (K96243 or 576). All the vaccinated mice had a significantly improved MTTD compared with naïve mice. Improvement of MTTD ranged from 7.4 days in mice immunised with *B. thailandensis* to 17.2 days in mice immunised with *B. mallei*, compared to 5.9 days in naïve mice.

Live-attenuated whole cell

Attenuated pathogens are severely limited in their ability to survive and replicate within a host. However, the transient presence of live microbes within a host may enable a stronger immune response to develop against the

pathogen. In BALB/c mice, immunisation with live attenuated *B. pseudomallei* NCTC 13179 provided better protection (81.8 % survival rate) than immunisation with a heat-killed preparation (0 % survival rate) against intravenous challenge with highly virulent *B. pseudomallei* NCTC 13178 (Barnes & Ketheesan, 2007).

An effective strategy for attenuating a pathogen is to engineer auxotrophic mutants. Disrupting a metabolic pathway can prevent the biosynthesis of essential metabolites that cannot be sequestered from the environment *in vivo*. Several auxotrophic *B. pseudomallei* have been constructed and assessed as potential vaccine strains. An *ilvI* deletion mutant is auxotrophic for branched chain amino acids leucine, isoleucine and valine (Atkins *et al.*, 2002a). Immunisation of BALB/c mice with this strain conferred *B. pseudomallei*-specific immunity up to 6,000 x MLD (offering no protection against *F. tularensis*). T-lymphocyte antigen specificity was developed to T3SS effector proteins BopE and BipD. However, no cytotoxic T lymphocytes were detected, suggesting poor stimulation of a cell-mediated immune response (Atkins *et al.*, 2002; Haque *et al.*, 2006b).

Another attenuated *B. pseudomallei* strain was developed by Propst *et al.* (2010). Deletion of *purM* resulted in auxotrophy for purines adenine and thiamine. Immunisation with this attenuated strain provided protection for 100 % of C57BL/6 mice and 80 % of BALB/c mice, stimulating an IgM and IgG response to surface antigens (Propst *et al.*, 2010; Silva *et al.*, 2013).

A less successful attempt was *B. pseudomallei* with an *aroC* deletion, which resulted in a strain auxotrophic for aromatic compounds such as tryptophan (Srilunchang *et al.*, 2009). Whilst all BALB/c and C57BL/6 mice survived immunisation with the mutant strain, demonstrating the bacteria were

attenuated, none of the BALB/c mice survived subsequent challenge with wild type *B. pseudomallei* A2. Survival rates in the more resistant C57BL/6 model ranged from 80 % when dosed with 6×10^3 CFU to 20 % when dosed with 6×10^5 CFU wild type *B. pseudomallei* A2. These levels of protection are less than desirable for effective vaccination.

Live carrier species

An alternative method is to use a related species to deliver the antigen. *B. mallei* was attenuated by deletion of *tonB*, which is required for iron acquisition (Mott *et al.*, 2013). As *B. mallei* expresses the same CPS as *B. pseudomallei*, this attenuated strain could be used to delivery CPS antigen to the host. In BALB/c mice, a humoral IgM and IgG response was detected. Subsequent challenge with *B. pseudomallei* K96243 resulted in a 75 % protection rate on day 36 post inoculation. However, pathology in the form of splenomegaly was observed and although all the *B. pseudomallei* was cleared from the host, the *B. mallei* attenuated strain persisted (Mott *et al.*, 2015). In another study, *B. thailandensis* E555 expressing the *B. pseudomallei* CPS was used to immunise BALB/c mice (Scott *et al.*, 2013). Mice immunised with *B. thailandensis* E555 were protected from infection when challenged with up to 6,000 times MLD *B. pseudomallei* K96243, whilst those immunised with an acapsular strain were not protected. Significant levels of IgG specific for *B. pseudomallei* CPS were detected in the protected mice.

1.3.4.2 Subunit vaccines

More recent strategies have focussed on identifying antigens that stimulate protective immunity and are highly specific to the pathogen of interest. A vaccine

with fewer, characterised antigens reduces the risk of an inappropriate immune response that may cause adverse side effects or anaphylaxis. Most subunit vaccines in use are toxoid vaccines that produce an immune response to toxins secreted by the pathogen, such as those against *C. diphtheriae* and *C. tetani* (Pichichero *et al.*, 2005). A vaccine containing protective antigen and lethal factor, which combine to form the lethal toxin of *Bacillus anthracis*, is also available to at-risk members of the UK Ministry of Defence (Baillie *et al.*, 2004). Non-toxoid subunit vaccines usually contain antigens expressed on the surface of the pathogen, such as the surface antigen protein used in the Hepatitis B vaccine (Cupps *et al.*, 1984) and the CPS of *S. pneumoniae* (Smit *et al.*, 1977; Shapiro *et al.*, 1991).

Protein subunits

Several virulence factors have been identified as potential antigens for immunisation against *B. pseudomallei*, including the T3SS, flagella and outer membrane proteins (OMPs).

The T3SS is an important virulence factor that consists of a number of structural, translocator and effector proteins. Stevens *et al.* (2004) observed a 60 % survival rate in BALB/c mice immunised with BipD-deficient *B. pseudomallei* 576 then subsequently challenged intraperitoneally with wild type *B. pseudomallei* K96243. However, BALB/c mice immunised with purified BipD only achieved a 20 % survival rate. This suggests that although BipD is a virulence factor it does not elicit a protective immune response. Similarly, a study by Druar *et al.* (2008) reported BALB/c mice immunised with purified BipD, BipB-N, BipC-C or BipC-N protein had no prolonged survival rates when challenged with wild type

B. pseudomallei Ashdown, compared to non-immunised mice (100 % mortality by day 4 post-challenge).

Another virulence factor, flagella, enables motility of *B. pseudomallei* in the extracellular environment. The structural protein flagellin has been investigated as a potential virulence target. Polyclonal IgG antibody raised in rabbits were able to reduce motility of the bacteria *in vitro*. However, passive immunisation of diabetic rats with the antibodies offered only 67.5 % survival rate compared to 40 % in non-immunised rats when subsequently challenged with *B. pseudomallei* (Brett *et al.*, 1994).

A number of OMPs have been tested as potential vaccine candidates. A study by Hara *et al.* (2009) found antibody isolated from melioidosis patients against Omp3 and Omp7. They used these proteins to immunise BALB/c mice but only achieved a 50 % survival rate when subsequently challenged with *B. pseudomallei* D286. Su *et al.* (2010) found immunisation with Omp85 gave 70 % protection in a similar study with BALB/c mice. They also observed a strong antibody response that induced complement-mediated killing and phagocytosis via opsonisation. Casey *et al.* (2016) observed 75 % protection of BALB/c mice immunised with OmpW and a strong IgG response. These results suggest the OMPs are immunostimulatory and are good candidates for subunit vaccines (arguably more so than T3SS and flagella proteins). However, the vaccine formulation needs to be improved to increase the protection efficacy.

Polysaccharides

LPS and CPS in *B. pseudomallei* elicit an immune response. Antibodies to both polysaccharides are found in melioidosis patients in endemic regions, though only anti-LPS antibodies were associated with reduced septicaemia and mortality

(Charuchaimontri *et al.*, 1999). Passive intravenous immunisation of BALB/c mice with antibodies against LPS and CPS increased the number of survivors and survival time compared to naïve BALB/c mice when challenged intraperitoneally with *B. pseudomallei* 4845 (Jones *et al.*, 2002). Passive immunity was also observed in BALB/c mice challenged with *B. pseudomallei* K96423 via the intranasal route. A combination of antibodies against LPS and CPS provided protection to 88 % of the immunised mice (AuCoin *et al.*, 2012).

A study by Nelson *et al.* (2004) analysed LPS and CPS as potential vaccine candidates. The study recorded no protective immunity in immunised mice when challenged via the intranasal route with *B. pseudomallei* NCTC 4845^T. However, when challenged via the intraperitoneal route, MTTD was extended from 2.6 days in naïve mice (100 % mortality day 11) to 17.6 days in mice immunised with LPS (50 % mortality day 35) and 10.5 days for mice immunised with CPS (100 % mortality, day 28)

B. thailandensis strain E264 expresses a *B. pseudomallei*-like LPS with an unmodified talose moiety (Ngugi *et al.*, 2010). Despite the structural difference, immunisation with LPS from *B. thailandensis* E264 conferred a 50 % survival rate on BALB/c mice challenged with *B. pseudomallei* K96423, compared to 60 % survival in those immunised with *B. pseudomallei* LPS. MTTD in naïve mice was 2.2 days. This increased to 32.8 days for mice immunised with *B. thailandensis* E264 LPS compared to 31.5 days for mice immunised with *B. pseudomallei* LPS (Ngugi *et al.*, 2010). As *B. thailandensis* can be handled at containment level 2, there is more scope for purifying the LPS directly from this organism than from highly virulent *B. pseudomallei*. However, as with all LPS preparation, care must be taken to ensure lipid A is removed to prevent potential cases of endotoxic shock.

Another atypical *B. thailandensis* is the encapsulated E555 strain expressing *B. pseudomallei* CPS. Immunisation of BALB/c mice via the intraperitoneal route with *B. thailandensis* was fully protective against intraperitoneal challenge with *B. pseudomallei* K96243 up to 6,000 MLD (Scott *et al.*, 2013). *B. thailandensis* E555 also cleared from tissues within three days post-inoculation and could therefore be used as an attenuated whole cell vaccine. However, it is more likely to be exploited for production and purification of *B. pseudomallei* CPS at a lower containment level.

Capsular polysaccharide vaccines have been developed against several pathogens including *Salmonella typhi*, *N. meningitidis* and *S. pneumoniae* (Artenstein *et al.*, 1970; Butler *et al.*, 1993; Klugman *et al.*, 1996). However, as they stimulate a B-lymphocyte response independent of helper T-lymphocytes, the ability of the vaccines to stimulate long-lasting immunity can be low, requiring booster immunisation (Jones *et al.*, 2005). Conjugation of a polysaccharide to a protein adjuvant improves the memory B-cell response against the saccharide antigen.

Glycoconjugates

Conjugating a polysaccharide to a carrier protein elicits a memory B cell response to the polysaccharide, generating long-lasting immunity. This is vital for the success of a vaccine programme against *B. pseudomallei*, as the immune system must be primed to recognise the pathogen before it enters the intracellular lifecycle and evades opsonisation.

The precise mechanisms by which glycoconjugates elicit an immune response to the polysaccharide have been challenging to define. The latest model was proposed by Avici, *et al.* (2012) (Figure 15). Briefly, they suggest that the

polysaccharide fragment of the binds to a B lymphocyte receptor and induces cross-linking. This triggers the B lymphocyte to endocytose the glycoconjugate before degrading it with reactive oxygen species. Once denatured, the resultant oligosaccharides and peptides form a new link and become glycan_p-peptides. The hydrophilic oligosaccharide is presented by MHC class II to the αβ T-lymphocyte receptor which is receptive to polysaccharide presented in this manner, with co-stimulatory signals. Upon recognition of the oligosaccharide, the αβ receptor activates the T-lymphocyte, which produces cytokines such as IL-2 and IL-4. These cytokines stimulate maturation of the B-lymphocyte into an antibody-producing memory cell (Avici *et al.*, 2012; Figure 15).

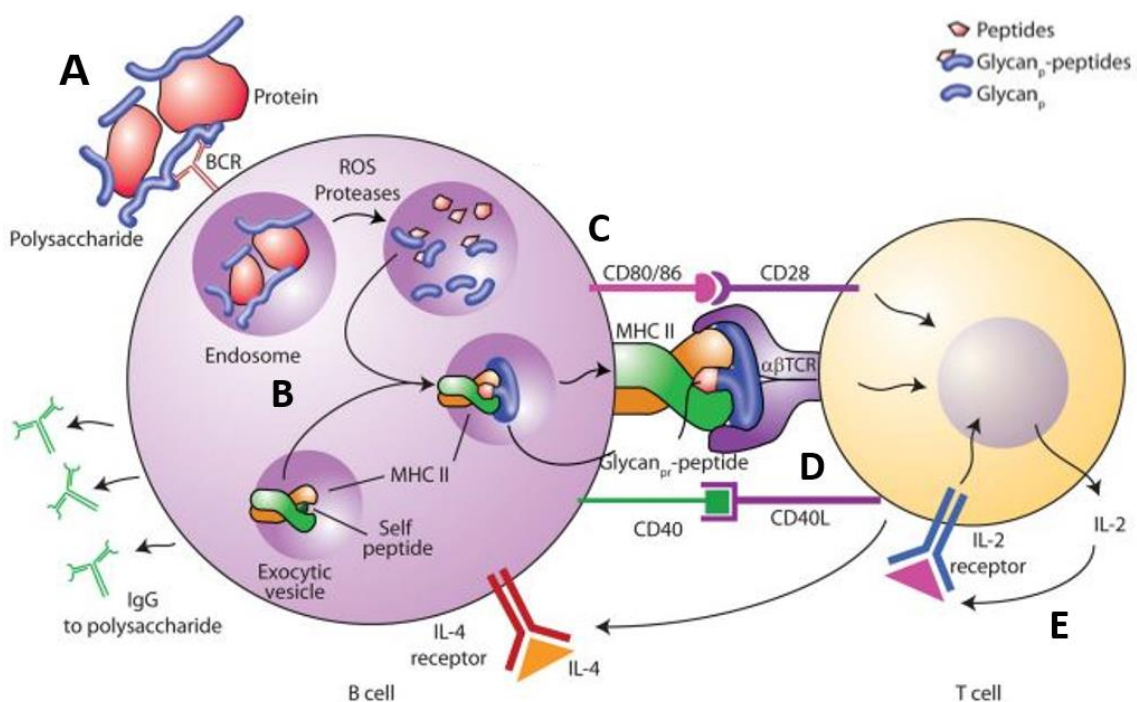


Figure 15: Avici model for the mechanism of T-cell stimulation by glycoconjugate vaccines. (A) polysaccharide antigen on the conjugate cross links B-cell receptor and initiates endocytosis of the glycoconjugate. (B) The glycoconjugate is denatured by reactive oxygen species into smaller oligosaccharides and peptides. These are processed to form a glycan_p-peptide ~10 kDa in size (C) binding of MHC class II to the peptide moiety of the glycan_p-peptide and presents the hydrophilic oligosaccharide to the CD4⁺ T lymphocyte αβ receptor. (D) The αβ receptor recognises the oligosaccharide in the context of MHC class II presentation (plus co-stimulation) and activates the T cell. (E) The T-cell produces IL-4 and IL-2, which induce the B-cell to mature into an anti-polysaccharide memory/plasma cell (Adapted and reproduced from Avici *et al.*, 2012).

Glycoconjugates for both LPS and CPS in *B. pseudomallei* have shown efficacy as protective immunogens. Passive immunity with IgG antibody raised against LPS conjugated to flagellin increased the LD₅₀ of *B. pseudomallei* 316c 100-fold in immunised diabetic rats compared with naïve rats (Brett & Woods, 1996). An study found LPS conjugated to gold nanoparticles extended MTTD and survival rates in BALB/c mice challenged with *B. mallei*. This could also stimulate immunity in *B. pseudomallei* as it expresses this same LPS as *B. mallei* (Gregory *et al.*, 2015).

Scott *et al.* (2014) conjugated CPS with transmembrane transporter protein LolC formulated with Alhydrogel-CpG adjuvant. This vaccine stimulated a strong IgG response in rats. It also protected 70 % BALB/c mice to day 35 post-intraperitoneal challenge with *B. pseudomallei* K96243 (Scott *et al.*, 2014). In comparison, all naïve mice succumbed to disease by day 2 post-challenge.

Glycoconjugate vaccine formulations - carrier proteins and adjuvants

In addition to linking a polysaccharide to a carrier protein, immunostimulatory adjuvants can be included in a vaccine formulation to boost the immune response. Priming the immune system (i.e. making it more active) can increase the level of protection achieved by a vaccine. Potential carrier proteins and adjuvants include include; cBSA (cationic bovine serum albumin) (Burtnick *et al.*, 2012; Scott *et al.*, 2014), CpG DNA (Scott *et al.*, 2014), liposomes (Puangpetch *et al.*, 2012), CRM₁₉₇ (non-toxic tetanus toxin mutant), virus-like particles (internal data, not shown) or *B. pseudomallei* proteins LolC (Scott *et al.*, 2014), FliC and Hcp1 (Gregory *et al.*, 2015).

Outer membrane vesicles

Outer membrane vesicles (OMVs) are naturally-occurring structures that have been adapted for use as an antigen delivery mechanism for vaccine development. Although the precise reasons and mechanisms for OMV formation are not fully understood, it is known to be a normal process and not the result of instability in the cell wall (McBroom *et al.*, 2006). The Gram-negative cell wall is a robust and complex structure, consisting of two phospholipid bilayers separated by peptidoglycan (Figure 16A).

Detailed reviews of OMV formation can be read by Ellis & Kuehn (2010) and Schwechheimer & Kuehn (2015). Briefly, the process begins with the distention of the outer membrane away from the peptidoglycan layer (Figure 16B). It eventually forms a discrete vesicle and sheds from the cell. The vesicle coat is comprised of the outer membrane bilayer, including the associated lipids, proteins and polysaccharides (Figure 16C). The vesicle lumen can encapsulate a variety of periplasmic and cytosolic molecules including DNA, receptors, enzymes, toxins and misfolded proteins (Figure 16C). OMVs are thought to perform several functions including cell wall maintenance, biofilm formation, microbial killing and resistance to bacteriophage (Kulkarni & Jagannadham, 2014).

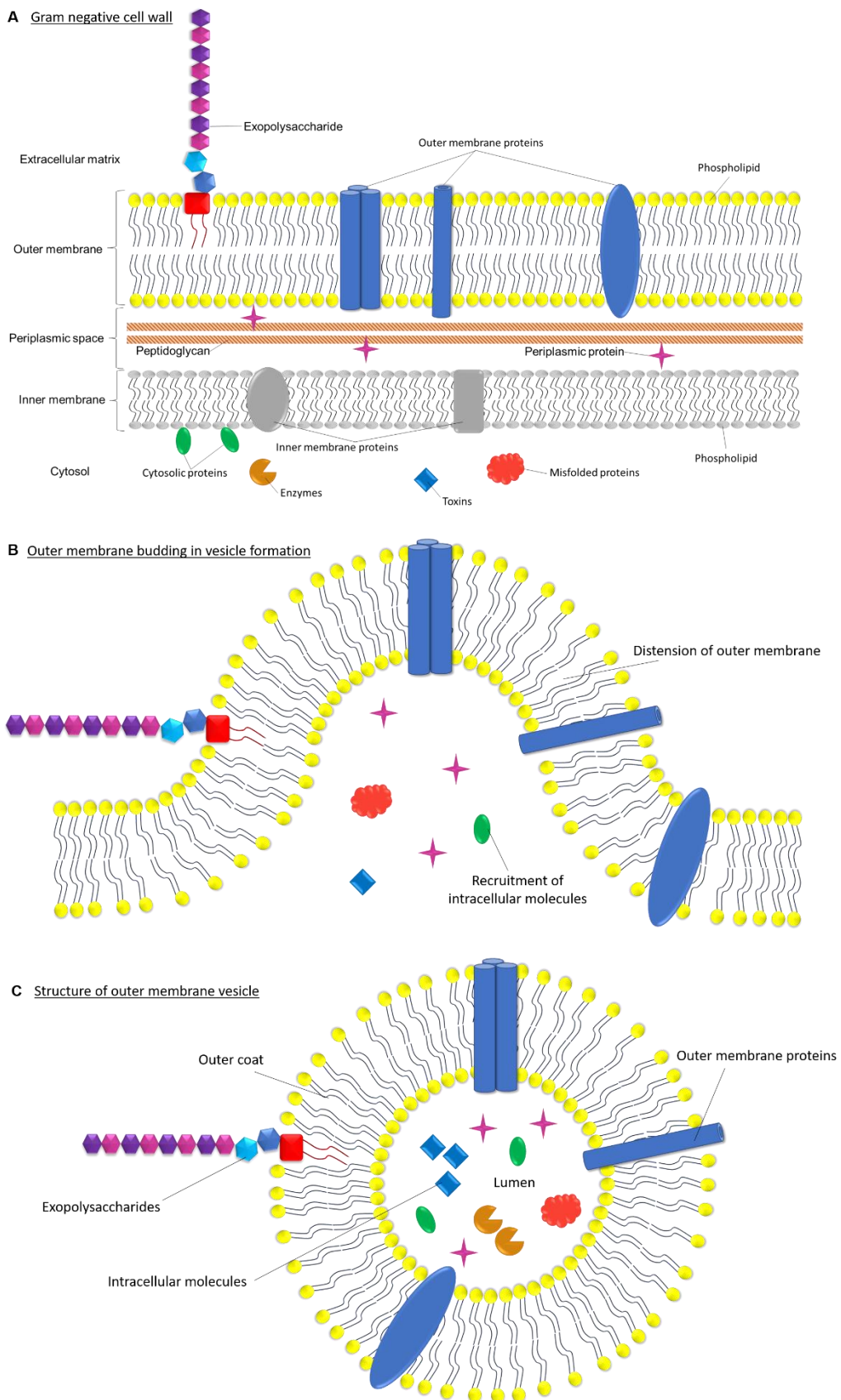


Figure 16: Outer membrane vesicle formation. (A) structure of the Gram-negative cell membrane. (B) budding of the outer membrane to form an outer membrane vesicle (OMV). (C) The OMV, formed of the outer membrane phospholipid bilayer, encapsulates intracellular molecules, including cytosolic and periplasmic proteins, enzymes, toxins and misfolded proteins. The OMV displays proteins and polysaccharides associated with the outer membrane.

As OMVs display many native antigens usually expressed on the surface of pathogenic Gram negative bacteria, their ability to elicit a protective immune response has been examined.

The first vaccine composed of OMVs was developed against *N. meningitidis* serotype B. It was trialled in schools in Norway, but only provided a 57.2% protection rate and was therefore discontinued (Bjune *et al.*, 1991). However, development of a vaccine for *N. meningitidis* serovar B proved challenging due to the CPS O-antigen polysialic acid closely resembling that found on human neural cells. The cross-reactivity of antibody posed a risk of either an autoimmune reaction or no protective immune response to potential vaccines (Finne *et al.*, 1987). The OMV strategy was therefore revisited. A novel multivalent vaccine containing three protein subunits (factor H binding protein, *Neisseria* adhesin A protein and *Neisseria* heparin binding antigen) and OMVs displaying the OMP PorA was developed (Perrett *et al.*, 2015). This combined protein and OMV vaccine is routinely administered against *N. meningitidis* serovar B as part of the UK childhood vaccine schedule.

OMVs have also shown to stimulate a protective immune response against other pathogens including *N. meningitidis* serogroups A and W-135 (Norheim *et al.*, 2012), *F. tularensis* (Brudal *et al.*, 2015) and *B. pertussis* (Bottero *et al.*, 2016).

OMVs have also shown promise as a potential vaccine candidate for *B. pseudomallei* (Nieves *et al.*, 2011; 2014). BALB/c mice were immunised subcutaneously with a preparation of multivalent *B. pseudomallei* OMVs containing CPS, LPS and proteins (17 putative periplasmic proteins and 12 predicted outer membrane proteins were identified). Following intranasal challenge with *B. pseudomallei* 1026b, 60% immunised mice survived compared with none of the naïve mice. However, the same protection was not observed in

mice immunised via the intranasal route despite strong IgG and IgA response (Nieves *et al.*, 2011). Another study showed that subcutaneous immunisation with OMVs derived from one strain of *B. pseudomallei* (1026b) can protect against intraperitoneal challenge with a different strain (K96243) (Nieves *et al.*, 2014). The immunised mice produced strong IgM and IgG responses specific to the OMV, LPS and CPS.

OMVs can also be engineered to express a desired polysaccharide. Immunisation in BALB/c mice has demonstrated a polysaccharide-specific immune response, similar to that observed in vaccination with a glycoconjugate (Price *et al.*, 2016; Valentine *et al.*, 2016). Trials have included expression of *N*-heptasaccharide from *C. jejuni*, serotype 14 O-antigen from *S. pneumoniae* and polysialic acid and Thomsen-Friedenreich antigen from *N. meningitidis* (Price *et al.*, 2016; Valentine *et al.*, 2016).

1.4 Engineering *B. pseudomallei* CPS in *E. coli*

1.4.1 Synthetic biology for polysaccharides

Synthetic biology is the practice of designing a pathway, circuit or system using functions that exist in nature to create a novel or practical product. An important feature of synthetic biology is the characterisation of components that are also becoming more accessible open source. Choices made in the design stage can influence the effectiveness of the system, so the availability of characterised components to inform design is of vital importance. This applies to the genetic elements (for example, the strength of a promoter) and the functional proteins (for example, an the activity of an enzyme). Unfortunately, there is still much work to be done in characterising glycosyltransferase activity for the biosynthesis of polysaccharides.

Currently, polysaccharides for research and industrial applications can be chemically synthesised (Gao & Guo, 2016) or expressed and purified from the natural organism (Gonçalves *et al.*, 2003). Chemical synthesis of polysaccharides is very complex due to number of possible structural isomers that exist, as well as branching and chain modifications that can occur. The process often requires several protection and de-protection steps. For example, chemical synthesis of oligosaccharides from the surface of *N. meningitidis* requires 15 reaction steps and has an overall yield of 14.5 %. In comparison, three glycosyltransferase reactions can synthesise the same product with a yield of 89.4 % (Figure 17) (Seeberger *et al.*, 2009).

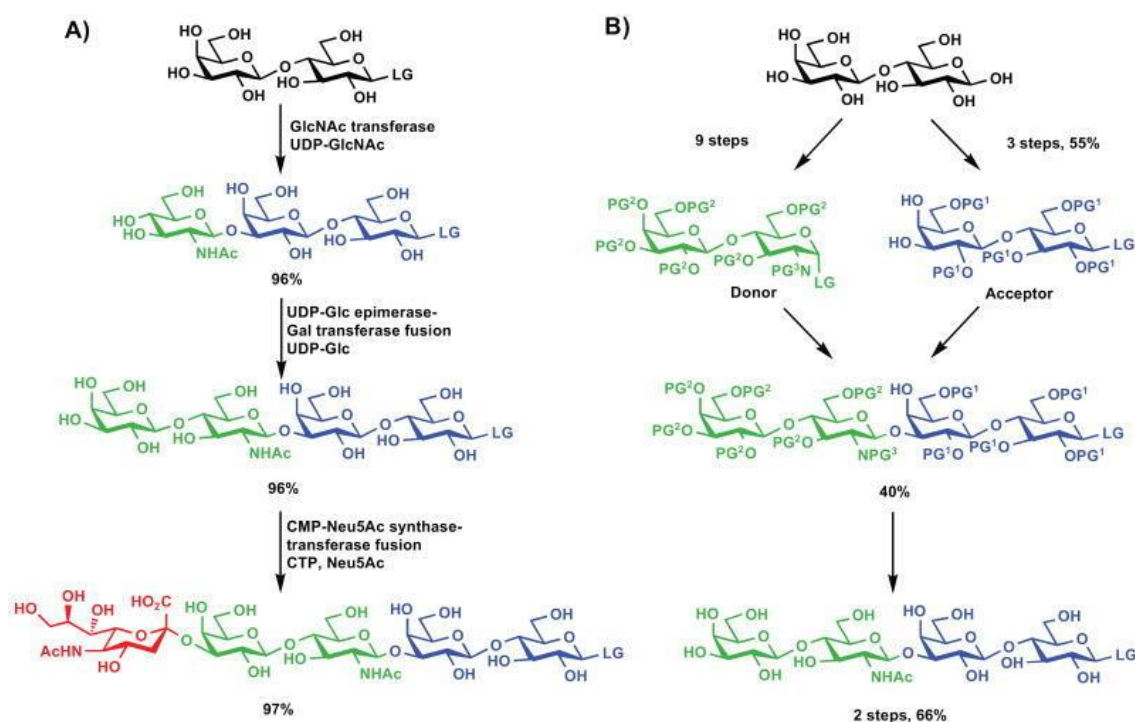


Figure 17: Enzymatic (A) and chemical (B) synthesis of oligosaccharides from *Neisseria meningitidis*. Reproduced from Seeberger *et al.* (2009).

Expression in the natural organism can be a viable option, however in many cases yields can be low (10 – 15 mg per litre *B. pseudomallei* lacking the O-antigen of the LPS (Burtnick *et al.* 2012)). Furthermore, it is not desirable to

culture large quantities of an ACDP hazard group 3 pathogen for the purpose of harvesting its capsule. There may also be a requirement to modify the polysaccharide. Modifications can include isolating an O-antigen from endotoxic lipid A, adding immunostimulatory moieties or antigen-binding sites, conjugating to a carrier protein to enhance the immune response or tagging with a signal peptide for secretion into the extracellular matrix.

The *B. pseudomallei* CPS is of interest as a potential antigen for a glycoconjugate vaccine. It is also a comparatively simple polysaccharide structure, a homopolymer of 2-O-acetyl-6-deoxy- β -D-manno-heptopyranose (Knirel *et al.*, 1992; Perry *et al.*, 1995). This provides an excellent opportunity to develop a platform for polysaccharide biosynthesis with a relatively uncomplicated polymer, before expanding the capability to enable biosynthesis of any bespoke polysaccharide desired.

Ideally, the nascent CPS would be conjugated to a carrier protein in the cell for direct inclusion in a glycoconjugate vaccine. It would then either be expressed on the cell surface and extracted (by traditional methods or enzymatic cleavage) or tagged with a signal peptide and actively secreted into the extracellular matrix for purification. This project examines the possibility of expressing the CPS on truncated polysaccharides in *E. coli*. In the process of engineering *E. coli* to express the *B. pseudomallei* CPS, a platform for expression of bespoke polysaccharides will develop.

1.4.2 Operon assembly for biosynthesis of *B. pseudomallei* CPS precursor

As well as traditional PCR, there are several methods for operon assembly available which broadly fit into two strategies. The first is based on flanking DNA fragments with defined sequences and fusing the overlaps using PCR and DNA

repair enzymes. Perhaps the most frequently used technique is Gibson cloning (Gibson *et al.*, 2009). Alternatives of this method include circular polymerase extension cloning (Quan & Tian, 2009), ligation independent cloning (Savitsky *et al.*, 2010), sequence-ligation independent cloning (Li & Elledge, 2012) and seamless ligation cloning extract (Zhang *et al.*, 2012). The benefits of these methods are that multiple fragments can be simultaneously assembled in a single 'one-pot' reaction and there are no restrictions on the sequences that can be included. However, these methods are limited by the specificity of each assembly which is not interchangeable with other constructs due to the requirement of complementary overhanging flanks. They are low-throughput methods that are ideal for the synthesis of an optimised operon for an isolated project. However, as previously discussed, synthetic biology is moving towards a more modular approach, where constituent parts are standardised and characterised.

The second strategy is the assembly of operons by the digestion and ligation of fragments via defined restriction sites. Not only is this a modular approach but also eliminates the requirement for PCR, which is time-consuming and can introduce point mutations. These approaches include the BioBricks™ 3A assembly method (Shetty *et al.*, 2008) and its derivatives (BglBricks and In-fusion BioBrick™ assembly (Anderson *et al.*, 2010; Sleight *et al.*, 2010)), Golden Gate cloning (Engler *et al.*, 2008) and the standard European vector architecture database (Silva-Rocha *et al.*, 2013). These approaches enable sequential assembly of an operon and allows each new combination to be characterised and optimised. It also enables any non-functioning or toxic combinations to be identified repaired or replaced. The main limitation of these methods is that the fragments cannot contain any of the specific restriction sites used for assembly.

1.4.2.1 BioBricks 3A assembly method

The method chosen to construct the CPS coding locus for expression in *E. coli* was BioBricks™ 3A assembly (Shetty *et al.*, 2008). The BioBricks™ Registry of Biological parts is an open source repository of genetic components (eg. promoters, terminators, ribosome binding sites (RBS), regulators, vectors). Some are well-characterised, others remain undefined and await evaluation. This allows plenty of choice and adaptability in the design of the CPS operon to ensure optimal expression in *E. coli*. In developing a new platform for bespoke polysaccharide biosynthesis, it would be beneficial to retain each intermediate combination of fragments. The intermediates could be used for subsequent biosynthesis of different bespoke polysaccharides or may benefit from alteration of the combination. It also easily allows each gene added to be tested for cytotoxicity. For example, if ligating and expressing gene A and gene B together is successful but the addition of gene C results in cytotoxicity, the cause of the problem is immediately identifiable. If genes A, B and C are simultaneously assembled together or the whole operon synthesised and cytotoxicity is observed, identifying the problem could be more challenging.

Whilst it was accepted that there are limitations to the method, the BioBricks™3A assembly protocol was chosen for its adaptability and for the potential benefits of a modular approach in the development of a bespoke polysaccharide expression platform.

1.5 Project aims

The main objective of this project was to better understand the biosynthesis of polysaccharides and to develop a novel synthetic biology platform for producing bespoke polysaccharides as required.

To achieve this, the CPS from *B. pseudomallei* was chosen as proof-of-principle due to its comparatively simple structure. It is also of interest because it has been identified as a potential antigen for a glycoconjugate vaccine against *B. pseudomallei*.

The aims of this project were to:

- identify the essential genes required for *B. pseudomallei* CPS synthesis from the CPS coding locus
- construct and express an operon for GDP-6dDHep biosynthesis in *E.coli*
- characterise the GDP-6dDHep biosynthetic pathway by expression *in vitro*
- develop a novel assay for the characterisation of glycosyltransferases
- exploit native polysaccharide biosynthetic pathways in *E. coli* for expression of *B. pseudomallei* CPS

The purpose of these aims is to contribute towards better understanding the biosynthesis of polysaccharides and how to exploit natural mechanisms to produce them for research and industrial uses.

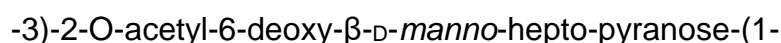
Chapter 2

Purification of biosynthetic GDP-6-deoxy- α -D-manno- heptose from genetically engineered *E. coli*

2.1 Introduction

The capsular polysaccharide (CPS) of *B. pseudomallei* is a virulence determinant (Ahmed *et al.*, 1999; Reckseidler-Zenteno *et al.*, 2005) and potential candidate for a glycoconjugate vaccine (REF). As production of polysaccharides by chemical synthesis is complex and expensive, there is a requirement for more efficient, economically viable methods. One approach is to utilise natural enzymatic pathways to produce the CPS biosynthetically. This can be achieved by genetically engineering *E. coli* to over-express the CPS.

The Type I O-polysaccharide CPS is found in all strains of *B. pseudomallei* (Knirel *et al.*, 1992; Perry *et al.*, 1995). It is an unbranched homopolymer with the structure:



Due to the simplicity and non-variability of its structure, *B. pseudomallei* CPS is an ideal candidate for the development of a new platform for biosynthesis of non-native, novel and bespoke polysaccharides.

The genes for expression of CPS are clustered on a single locus on chromosome 1 of the *B. pseudomallei* genome (Cuccui *et al.*, 2012). As previous attempts at cloning the entire operon into *E. coli* were unsuccessful, the approach used here was to select only essential genes to integrate into the native pathways of *E. coli*.

The first step in biosynthesis of *B. pseudomallei* CPS is the expression of nucleotide activated sugar precursor GDP-6-deoxy- α -D-manno-heptose (GDP-6_DHep). Six biosynthetic genes (*gmhA*, *wcbN*, *wcbM*, *wcbL*, *wcbJ* and *wcbK*) divert S7P from the pentose phosphate pathway into GDP-6_DHep production (Butty *et al.*, 2009; Cuccui *et al.*, 2012).

2.1.1 Aims

The aim of this chapter is to clone the six genes for GDP-6dbHep biosynthesis from *B. pseudomallei* and express them in *E. coli*. The second aim is to purify GDP-6dbHep from *E. coli* and to verify the structure was correct.

The purpose of these aims is to demonstrate that *E. coli* can be engineered to express a non-native nucleotide activated sugar precursor. This is the important first step in developing an adaptable platform for the biosynthesis of any desired polysaccharide. It is also required for synthesis the synthesis of *B. pseudomallei* CPS, which could contribute to the design of a glycoconjugate vaccine against melioidosis.

2.2 Materials and methods

All reagents were purchased from Sigma-Aldrich unless otherwise stated.

2.2.1 Bacterial strains and reagents

The strain of *E. coli* used was Subcloning Efficiency™ DH5α™ (Invitrogen) throughout this work.

2.2.2 Growth, storage and competent cell preparation of *E. coli*

2.2.2.1 Growth and storage conditions

Broth cultures of *E. coli* were grown in Luria-Bertani (LB) medium (1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 170 mM NaCl, pH 7.0) (Green & Sambrook, 2012) with appropriate antibiotics and aeration at 37 °C for 18 to 24 h or until a specific OD₆₀₀ was reached as stated. Cultures were 5 mL in volume, with antibiotic concentrations at 50 µg mL⁻¹ unless otherwise indicated.

LB agar (2 % w/v agar) plates with the relevant antibiotics at 50 µg mL⁻¹ (unless otherwise stated) were streaked with 180 µL bacterial culture and incubated at 37 °C for 18 to 24 h. Plates with colonies were stored at 4 °C for up to one month before discarding.

Glycerol stocks were made from overnight broth cultures of *E. coli* by adding 200 µL sterile glycerol to an 800 µL aliquot of culture and storing at -80 °C.

2.2.2.2 Preparation of chemically competent cells

A small scrape from an *E. coli* glycerol stock was cultured overnight in 30 mL LB broth at 37 °C with aeration. From this, 8 mL was added to 200 mL LB and returned to the same incubation conditions until OD₆₀₀ 0.4 was reached. The culture was divided into four 50 mL aliquots and chilled on ice for 15 min. The

cells were pelleted by centrifugation (3,000 $\times g$ for 15 min at 4 °C.) The supernatant was discarded and the pellet resuspended in 16 mL ice-cold buffer (100 mM RbCl, 50 mM MnCl₂, 10 mM CaCl₂, 15 % glycerol (v/v), 1 M potassium acetate, pH 5.8 with acetic acid). The cells were chilled on ice for 15 min and then centrifuged (3,000 $\times g$ for 15 min at 4 °C). The supernatant was discarded and the pellets resuspended in 4 mL buffer (10 mM RbCl, 75 mM CaCl₂, 15 % glycerol (v/v), 10 mM MOPS, pH 6.8). The resuspended pellets were pooled and then divided into 250 μ L aliquots. These were frozen in liquid nitrogen and stored at -80 °C until use.

2.2.3 Plasmid construction for GDP-6ddHep biosynthesis

2.2.3.1 Gene synthesis

Six *B. pseudomallei* genes for the biosynthesis of GDP-6ddHep (*gmhA*, *wcbL*, *wcbN*, *wcbM*, *wcbK* and *wcbJ*) were codon optimised for expression in *E. coli* and synthesised by Eurofins MWG Operon. The constructs included the pBAT4 ribosome binding site (RBS), C-terminal Myc-tag and were flanked by the *EcoRI*, *XbaI*, and *SpeI* (*BcuI*) and *PstI* restriction sites for BioBricks™ cloning (Figure 18). See Appendices 3-5 for full sequences. The genes were synthesised in pEX plasmids. Upon receipt, they were resuspended in 10 mM Tris-HCl (pH 8.0) to a final concentration of 100 ng μ L⁻¹ and kept at -20 °C until use.

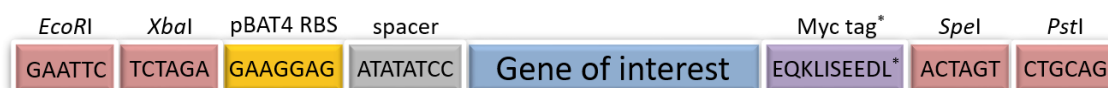


Figure 18: Synthetic gene constructs for BioBricks™ 3A assembly. Blue = gene of interest; Red = restriction site; Orange = ribosome binding site; Purple = c-Myc-tag (*amino acid sequence shown; followed by a stop codon); Grey = spacer sequence before start codon.

2.2.3.2 Plasmid transformation

A 2 μL aliquot of plasmid was incubated on ice with 50 μL freshly defrosted competent *E. coli* cells for 30 min. The cells were heat shocked in a 42 $^{\circ}\text{C}$ water bath for 45 s, then returned immediately to ice for 1 min. Next, 150 μL of LB medium was added to the cells, which were then incubated at 37 $^{\circ}\text{C}$ for 1 h. The cells were then plated on LB agar containing 50 $\mu\text{g mL}^{-1}$ of the appropriate antibiotic and incubated at 37 $^{\circ}\text{C}$ for 18-24 h.

2.2.3.3 Plasmid purification

For plasmid purification, 5 mL overnight cultures were inoculated from either a glycerol stock or single colony after transformation. The cells were pelleted by centrifugation (3,000 $\times g$, 30 min at 4 $^{\circ}\text{C}$) and the plasmid extracted using the GeneJet Plasmid Miniprep kit (Thermo Fisher) following the manufacturer's instructions. Concentrations of plasmid were calculated by absorbance at 260 nm using a Nanodrop ND-2000c (Thermo Fisher). Purified plasmids were stored at -20 $^{\circ}\text{C}$.

2.2.3.4 Polymerase chain reaction

Primers were synthesised by Invitrogen (Appendix 1). Lyophilised primers were resuspended in deionised water (dH_2O) to a final concentration of 100 μM . These were further diluted in dH_2O to make working stocks of 10 μM . Primers were kept at -20 $^{\circ}\text{C}$ until use. Polymerase chain reaction (PCR) reagents were obtained from Thermo Fisher unless otherwise indicated.

Each reaction contained 5x GC enhancer (New England BioLabs), 5x GC buffer, 250 μM dNTPs, 0.5 μM forward primer, 0.5 μM reverse primer, 10 mU μL^{-1} Phusion® polymerase and nuclease-free water to 20 μL .

PCR conditions were: initial denaturation at 98 °C for 30 s, then 35 cycles of 98 °C for 20 s, 55 °C for 15 s, 72 °C for 45 s and a final extension time at 72 °C for 45 s per kbp (minimum 90 s). PCR products were kept at 4 °C until use.

When required, PCR products were purified either directly or following gel extraction, using the GeneJET PCR Purification Kit or GeneJET Gel Extraction Kit respectively (Thermo Fisher) following the manufacturer's instructions.

2.2.3.5 Restriction digests

Reagents for restriction digest (RD) were from Thermo Fisher unless otherwise stated. Each reaction contained 5 µL plasmid DNA, 2 µL 10x FastDigest buffer (New England BioLabs®), 2.5 % (v/v) bovine serum albumin (BSA), 0.25 U µL⁻¹ of each restriction enzyme and nuclease-free water to 20 µL. (Total enzyme used was 5 U per 20 µL reaction, from a 10 U µL⁻¹ stock). The reactions were incubated at 37 °C for 30 min and 80 °C for 20 min to heat kill the enzymes for downstream applications.

RD products were purified by gel extraction using the GeneJET Gel Extraction Kit (Thermo Fisher) following the manufacturer's instructions as required.

2.2.3.6 Ligations

All reagents for ligations were from Thermo Fisher. Reactions used 0.25 U µL⁻¹ T4 DNA ligase, 1 µL 10x T4 DNA ligase buffer, 2 µL of each DNA component and nuclease-free water to 10 µL. The reactions were incubated at 16 °C for 30 min and 80 °C for 20 min to heat kill enzymes for downstream applications.

2.2.3.7 Nucleic acid visualisation

PCR and RD products were run on 1 % (w/v) agarose gels in TAE buffer (40 mM Tris base, 20 mM acetic acid, 1 mM ethylenediaminetetraacetic acid (EDTA)) with Midori Green stain (Nippon Genetics) at 100 V for 40 min. DNA markers, HyperLadder™ 1 kb or 50 bp (BioLine), were added for comparison. Images were taken on a bench top UV transilluminator (UVP).

2.2.3.8 Subcloning into a BioBricks™ vector

To move the synthesised genes into a BioBricks™ vector (e.g. one containing a promoter to express a protein), it was necessary for the destination vector to have a different antibiotic resistance to the plasmid containing the gene. Both the plasmid and destination vector were digested with *EcoRI* and *PstI*. The digested parts were ligated and transformed into *E. coli* DH5α and plated onto agar with the antibiotic selection for the destination vector. Colonies were selected for plasmid purification and the success of the subcloning was verified by PCR.

2.2.3.9 BioBricks™ 3A assembly

Genes for ligation were synthesised in vectors with a different antibiotic resistance as the destination vector and with specific flanking restriction sites (Figure 18). Regulatory components (e.g. promoters) in plasmids with the same restriction site configuration were picked from the 2012 BioBricks™ kit by resuspending in 10 µL nuclease-free water. The plasmids were transformed into *E. coli* and purified. Two components and a destination vector were RD using 250 ng of each plasmid per reaction. The combination of RD enzymes used were as follows: the first component was digested with *EcoRI* and *SpeI*, the second with *XbaI* and *PstI*, and the destination vector with *EcoRI* and *PstI* (Figure 19).

The digested components were then ligated (Figure 19). Assuming equal RD efficiency, 2 μL of each component was used per 10 μL reaction. The ligation product was then transformed into *E. coli* DH5 α . Colonies were selected for plasmid preparation and confirmation of ligation success by PCR.

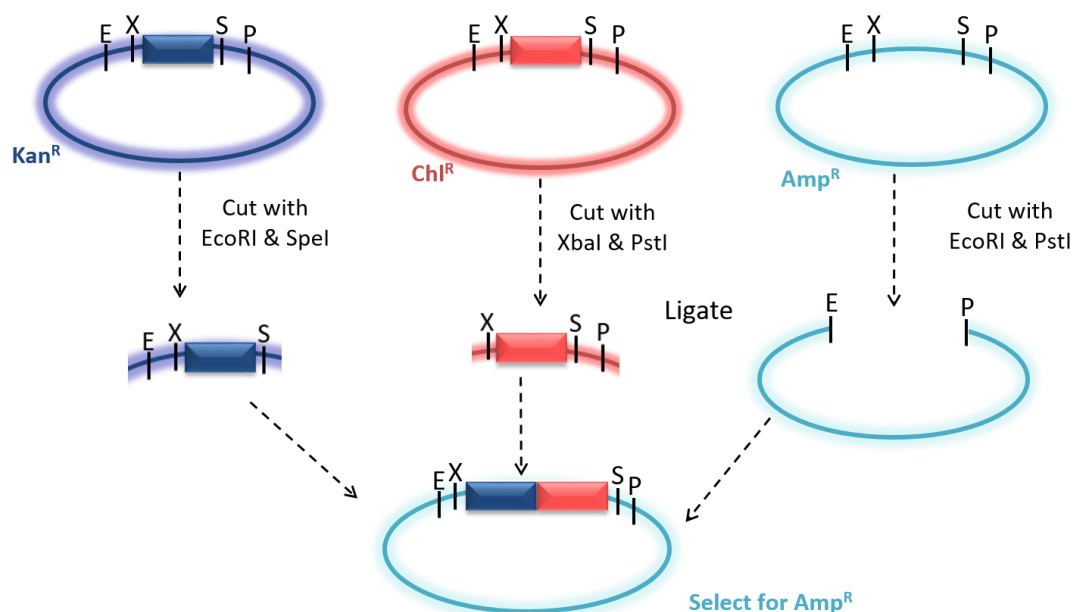


Figure 19: BioBricks™ 3A assembly protocol. Dark blue = component 1; Red = component 2; Light blue = destination vector. The three plasmids are cut with the restriction enzymes indicated and the three parts ligated. The novel plasmid has the antibiotic resistance of the destination vector backbone. The flanking BioBricks™ restriction sites are conserved, allowing the ligated construct to be used as a component in downstream BioBricks™ 3A assembly cloning. The *SpeI* and *XbaI* restriction sites have a common sticky end, but the ligated product is not a target for either enzyme.

2.2.4 Verification of plasmid expression

2.2.4.1 *E. coli* growth curve

Overnight cultures of *E. coli* were used to inoculate 50 mL ZYM-5052 broth containing 50 $\mu\text{g}/\text{mL}$ chl to a final OD_{600} of 0.05. These were incubated at 37 $^{\circ}\text{C}$ plus aeration for 1,440 min (24 h), with the OD_{600} recorded on an hourly basis (for 9 hours) to monitor *E. coli* growth. A final reading was taken at the 1,440 min (24 h) time point.

2.2.4.2 Plasmid sequencing

Plasmids were purified as described in section 2.2.3.3. They were subsequently sequenced by Source Bioscience using Sanger sequencing. This required 5 μL template DNA (100 ng μL^{-1}) and 5 μL each primer (3.2 pmol μL^{-1}) per reaction, measured by Nanodrop ND-2000c (Thermo Fisher). The sequencing data were analysed using SnapGene v. 3.1.

2.2.4.3 RNA extraction

For RNA extraction, the PCR hood, pipettes, filter tips and nuclease-free water were decontaminated with 30 min UV exposure and RNaseZap® (Thermo Fisher) to remove ribonucleases (RNases). Overnight cultures of *E. coli* were diluted with LB to 0.25 OD₆₀₀ in a final volume of 500 μL for use with the RNeasy Protect Bacteria Mini Kit RNA extraction kit (Qiagen), following the manufacturer's instructions. After extraction, the samples were treated with the Ambion TURBO DNA-free™ Kit (GE Healthcare) multiple times to ensure full removal of any contaminating DNA, following the manufacturer's instructions. PCR (section 2.2.3.4) was used to confirm purity of the RNA extracts. Expression of the housekeeping gene for GAPDH in wild type *E. coli* was used as a control.

2.2.4.4 Reverse transcriptase PCR

Purified RNA samples were analysed by reverse transcriptase PCR (RT-PCR) using the OneTaq® RT-PCR Kit (New England Biolabs). The recommended instructions for first strand complementary DNA (cDNA) synthesis and reverse transcription were followed. However, the final PCR was conducted as described in section 2.2.3.4.

2.2.4.5 Small volume protein extraction

E. coli cells were pelleted from 5 mL cultures by centrifugation at 3,000 xg for 30 min at 4 °C. The supernatant was removed and the cells resuspended in 250 µL buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl). The samples were either used immediately or stored at -20 °C.

The cells were then lysed using a freeze-thaw method, which involved alternating between immersion in liquid nitrogen and a 42 °C water bath five times. To the lysed cells, 0.25 U µL⁻¹ Benzonase® nuclease (Merck Millipore) was added and incubated at room temperature for 20 min until the sample became less viscous. A 20 µL aliquot was kept for the untreated 'neat' sample. The remaining volume was centrifuged at 14, 000 xg at 4 °C for 30 min. The supernatant was removed and retained and the pellet resuspended in 40 µL buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl).

2.2.4.6 Purification of c-Myc tagged proteins

The lysate from small cultures of *E. coli* DH5α transformed with empty pSB1C3 vector and the GDP-6dHep biosynthesis operon were applied to an anti-c-Myc agarose column to extract the c-Myc tagged proteins (GmhA, WcbL, WcbN, WcbM, WcbK and WcbJ). A Pierce c-Myc Tag IP/Co-IP Kit (Thermo Fisher) was used, as per the manufacturer's instructions.

2.2.4.7 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

To examine protein expression, prepared samples were analysed using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). The gels were either made in-house (10 % acrylamide; Table 1) or purchased pre-cast (ExpressPlus™ PAGE Gel 4-12 % acrylamide; GenScript®).

Table 1: In-house SDS- and BN-PAGE gel composition. Proportions of chemicals required for casting the stacking and separating components of in-house gels for SDS-PAGE and BN-PAGE analysis of proteins.

	SDS-PAGE	Native
Stacking	4 % (v/v) 19:1 acrylamide/bis-acrylamide 360 mM bis-Tris-HCl (pH 6.6) 0.1 % (w/v) ammonium persulphate 0.1 % (v/v) TEMED	3.5 % (v/v) 29:1 acrylamide/bis-acrylamide 100 mM Tris-HCl (pH 6.8) 0.1 % (w/v) ammonium persulphate 0.1 % (v/v) TEMED
Separating	10 % (v/v) 19:1 acrylamide/bis-acrylamide 360 mM bis-Tris-HCl (pH 6.6) 0.2 % (w/v) ammonium persulphate 0.2 % (v/v) TEMED	15 % (v/v) 29:1 acrylamide/bis-acrylamide 700 mM Tris-HCl (pH 8.8) 0.2 % (w/v) ammonium persulphate 0.2 % (v/v) TEMED

Samples were prepared by taking 20 μ L aliquots of protein (neat, pellet and supernatant), adding 20 μ L loading buffer (100 mM Tris-HCl pH 6.8, 4 % (w/v) SDS, 20 % (v/v) glycerol, 2 % (v/v) β -mercaptoethanol, 0.2 % (w/v) bromophenol blue; Green & Sambrook, 2012) and boiling for 5 min to denature the proteins. Depending on expected protein concentration, 2 - 20 μ L of each sample was loaded onto the gel with 5 μ L protein ladder for comparison (Spectra™ Multicolor Broad Range Protein Ladder (Thermo Fisher), unless otherwise indicated). The gels were run at 200 V for 50 min. In-house running buffer (250 mM MOPS, 250 mM Tris base, 5 mM EDTA, 0.5 % (w/v) SDS) or premade Tris-MOPS buffer (GenScript®) was used as appropriate. The gels could be imaged or used for Western blotting (section 2.2.4.10). For imaging, the gels were stained in Instant Blue™ (Expedeon) for 30 min, destained in dH₂O overnight, then photographed using a bench top UV transilluminator (UVP) with a white light filter.

2.2.4.8 Blue native-PAGE

Blue native (BN)-PAGE (Fiala *et al.*, 2011) separates intact proteins by molecular weight to charge ratio. After preparation of the protein, 4 μ L BN loading buffer (310 mM Tris-HCl pH 6.8, 0.1 % (w/v) bromophenol blue,

50 % (v/v) glycerol) was added to 16 μL aliquots of each sample and marker protein. Marker proteins (Table 2) used instead of a protein ladder were kindly gifted by Sheera Abdulla (University of Exeter).

Table 2: Marker proteins for BN-PAGE. Concentration (mg mL^{-1}), molecular weight (kDa) and subunit molecular weight (kDa) of marker proteins for BN gel electrophoresis.

Marker Protein	Protein concentration (mg mL^{-1})	Molecular weight (kDa)	Subunit molecular weight (kDa)
Carbonic anhydrase	3	30	-
Albumin	10	66.5	-
Alcohol dehydrogenase	5	150	-
β -amylase	4	206	56
Apo ferritin	10	440	19
Thyloglobulin	8	660	-

A 10 μL aliquot of each sample was then loaded onto 15 % native gels as prepared in Table 1 with 10 μL of each marker protein. The inner cassette was filled with cathode buffer (15 mM bis-Tris-HCl, 50 mM tricine, 0.02 % (w/v) Coomassie Brilliant Blue G-250 (BioRad) pH 7.0) and the tank filled with anode buffer (50 mM bis-Tris-HCl pH 7.0). The gels were run for 75 min at 150 V. The gels could then either be used for second dimension (2D)-PAGE (section 2.2.4.9) or Western blotting (section 2.2.4.10).

2.2.4.9 Second dimension-PAGE

To enhance the resolution of protein separation, 2D-PAGE was used. This method combines the molecular weight to charge ratio separation of BN PAGE with size separation of denatured proteins under SDS-PAGE conditions.

For 2D-PAGE separation, the lane containing sample proteins from a BN gel was excised and soaked in SDS buffer (12.5 mM Tris-HCl pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 0.02 % (w/v) bromophenol blue). The strip was microwaved in the buffer for 30 s then cooled at room temperature. A second SDS-PAGE gel was prepared (Table 1). However, for the stacking segment a large gap was made to accommodate the native gel sample strip, with a standard well for a protein ladder. The excised BN gel strip was placed in the gap in the separating gel and overlaid with SDS buffer. The second separation was run under normal SDS-PAGE conditions (section 2.2.4.7) and could be imaged or used for Western blotting (section 2.2.4.10).

2.2.4.10 Western blot

Gels from SDS-, BN- and 2D-PAGE were used in Western blotting to examine protein expression by antibody detection of the Myc- or His-tag.

The gel, two filter papers, nitrocellulose (GE Healthcare) and two sponges were soaked in Towbin buffer (27.5 mM Tris-HCl pH 8.6, 200 mM glycine, 20 % (v/v) methanol, 0.01 % (w/v) SDS; without SDS for BN gels) (Link, 1999) for 10 min. For transferring proteins to the nitrocellulose membrane, a sponge, filter paper, nitrocellulose membrane, gel, filter paper and sponge were layered respectively in a cassette which was placed in a tank filled with Towbin buffer. The transfer ran for 75 min at 180 V. The nitrocellulose was then removed and stained with Ponceau S to assess quality of the transfer. The membrane was rinsed in TBST (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1 % Tween 20) then blocked with 5 % (w/v) skim milk overnight at 4 °C with gentle rocking. The nitrocellulose was gently rinsed with TBST then incubated for 1 hr at 4 °C in primary antibody (Table 3) in TSBT plus 5 % (w/v) skim milk (TSBT/milk). The membrane was then washed

three times in TBST/milk to remove unbound primary antibody. Next, the membrane was incubated with the secondary antibody (Table 3) in TBST/milk at room temperature for 1 hr. Finally, it was washed three times in TBST/milk to remove unbound secondary antibody.

Table 3: Antibodies for Western blotting against His- and c-Myc-tagged proteins. Detailing the host and target organism, cell lineage, step used, conjugate, dilution used, supplier and proteins targeted in this study. *HRP = horseradish peroxidase.

Antibody	6x-His Epitope Tag Antibody (HIS.H8)	c-Myc Epitope Tag Antibody	Goat anti-Mouse IgG (H+L) Poly-HRP* Secondary Antibody
Host	Mouse	Mouse	Goat
Target	Tag	Tag	Mouse
Lineage	Monoclonal	Monoclonal	Polyclonal
Step	Primary	Primary	Secondary
Conjugate	Unconjugated	Unconjugated	HRP*
Dilution	1 in 1,000	1 in 2,500	1 in 5000
Supplier	Thermo Fisher	Thermo Fisher	Thermo Fisher
Proteins	Apyrase, WcbH	WcbJ, WcbK, WcbL, WcbM, WcbN, GmhA	IgG (heavy and light chain)

The blots were processed with the SuperSignal™ West Pico Chemiluminescent Substrate kit as instructed (Thermo Fisher). They were then wrapped in cling-film and placed in a cassette for exposure to X-ray film (Kodak).

2.2.5 Total metabolite extraction from *E. coli*

Overnight 5 mL *E. coli* cultures were harvested by centrifugation (3,000 $\times g$ for 30 min at 4 °C) immediately after incubation. The supernatant was discarded and the pellets resuspended in 500 μ L 80 % (v/v) methanol kept at -80 °C. The cells were lysed by five freeze-thaw cycles in liquid nitrogen and a 42 °C water bath. The lysed cells were centrifuged at 14,500 $\times g$ for 30 min at 4 °C. The supernatant was removed, taking care not to disturb the pellet, and filtered at 20 μ m. The extracted metabolites were kept at -80 °C until use.

2.2.6 GDP-6dbHep identification by mass spectrometry

2.2.6.1 Liquid chromatography for mass spectrometry

Liquid chromatography (LC) followed the same method before either quadrupole time-of-flight (Q-TOF) or triple quadrupole (QQQ) mass spectrometry (MS) analysis. LC was performed on an Agilent Technologies 1200 Rapid Resolution system. Hydrophilic interaction chromatography was performed using an Xbridge™ Amide 3.5 μm (2.1 x 150 mm) column (Waters). The buffers used were: A = 90 % (v/v) acetonitrile (ACN) and 0.1 % (v/v) ammonia; B = 10 % (v/v) ACN, 0.1 % (v/v) ammonia and 5 mM ammonium formate. The gradient was 0 - 54 % buffer B for 17 min, 54 % buffer B for 6 min, 54 - 0 % buffer B for 1 min and 10 min post time, with a 0.25 mL min⁻¹ flow rate and column temperature held at 35 °C throughout. The injection volume was 5 μL for standards and 20 μL for samples.

2.2.6.2 Untargeted LC-MS metabolite profiling (Q-TOF)

Metabolic profiling was conducted on an Agilent Technologies 6520 Accurate mass Q-TOF MS system coupled with LC as previously described (section 2.2.6.1). Source conditions for electrospray ionisation were: gas temperature 325 °C with drying gas flow rate of 9 L min⁻¹ and a nebuliser pressure of 35 psig. The capillary voltage was ± 3.5 kV, the fragmentor was 115 V and skimmer 70 V. Scanning was performed using the auto MS/MS function at 3 scans s⁻¹ with a sloped collision energy of 3.5 V/100 Da and an offset of 5 V.

2.2.7 Purification of GDP-6dbHep

2.2.7.1 Extraction of GDP-6dbHep

To purify GDP-6dbHep, a modified protocol was followed (Räbinä *et al.*, 2001). Metabolites were extracted as previously described (section 2.2.5). Ammonium bicarbonate was then added to each sample to a final concentration of 10 mM. Samples were applied twice to 250 mg (3 mL) Supelclean™ ENVI-Carb Solid Phase Extraction (SPE) columns which had been conditioned with 3 mL 80 % (v/v) ACN plus 0.1 % (v/v) trifluoroacetic acid (TFA) and flushed with 2 mL dH₂O. Loaded samples were washed with 2 mL dH₂O, then 2 mL 25 % (v/v) ACN and finally 2 mL 50 mM triethylammonium acetate (TEAA) (pH 7.0). The samples were eluted in 3 mL 50 mM TEAA and diluted to 5 mL volume with ammonium bicarbonate at a final concentration of 10 mM. (For larger-scale purification, 1 g (12 mL) SPE columns were also used and volumes were increased as appropriate.) Anion exchange chromatography was then performed using an ÄKTA pure system, with a flow rate of 0.5 mL min⁻¹ used throughout. A diethylaminoethanol (DEAE) Sepharose Fast Flow column (GE Healthcare) was equilibrated with 10 mM ammonium bicarbonate. The 5 mL samples were applied to the column and washed with 5 mL 10 mM ammonium bicarbonate. The sugar nucleotides were eluted with 250 mM ammonium bicarbonate. The first 750 µL were discarded, and the remaining 1.25 mL kept at 4 °C until use. The samples were dried by vacuum centrifugation, dissolved and dried with 200 µL dH₂O twice and with 20 mM TEAA (pH 6.0) once to remove the ammonium bicarbonate salt. Finally, the samples were resuspended in 200 µL 20 mM TEAA (pH 6.0) and kept at 4 °C until analysis with high performance liquid chromatography (HPLC) and LC-MS.

2.2.7.2 Purification of GDP-6dbHep by HPLC

HPLC was performed on a 1260 Infinity system (Agilent Technologies) using a 250 x 2.0 mm Polaris 3 μm C18-Amide column (Agilent Technologies). The buffers were: buffer A = 20 mM TEAA (pH 6.0); and buffer B = 4 % (v/v) ACN in 20 mM TEAA (pH 6.0). For analysis, 30 μL of sample was injected, with the remaining 170 μL injected for purification. The column was equilibrated with buffer A. The gradient used was 0 % buffer B for 20 min at 0.15 mL min^{-1} , 0 - 50 % buffer B for 20 min at 0.2 mL min^{-1} , 0 % buffer B for 7 min at 0.2 mL min^{-1} . Post-time = 10 min. Fractions were collected for targeted MS-MS.

2.2.8 Verification of GDP-6dbHep structure

2.2.8.1 Sample purity analysis (Q-TOF)

To examine the purity of the collected GDP-6dbHep fractions, the samples were analysed by LC-MS Q-TOF under the same conditions as detailed in sections 2.2.6.1 and 2.2.6.2.

2.2.8.2 Targeted LC-MS metabolite screening (QQQ)

Targeted MS-MS was performed on an Agilent technologies 6410 QQQ MS system coupled with LC as previously described (section 2.2.6.1). Source conditions were as follows: gas temperature 350°C, drying gas flow rate 11 L min^{-1} , nebuliser pressure 35 psig, capillary voltage ± 4 kV and skimmer 70 V. The fragmentor was 135 V and collision energies were 25 V/100 Da for multiple reaction monitoring (MRM) targeting fragments GDP (m/z: 442.0150) and 6-deoxy- α -D-*manno*-heptose (176.0705) and 0 V/100 Da targeting intact GDP-6dbHep (m/z: 618.0855). Additional identifying fragments were also screen for; m/z: 78.9594, 158.9263, 177.0760, 344.0407 and 424.0057.

2.2.8.3 Gas chromatography-MS structural analysis

Purified samples of GDP-6dBHep and several standards (Table 4) were prepared for analysis by gas chromatography (GC)-MS.

Table 4: Samples and standards for GC-MS analysis. All samples were of unknown concentration. All standards were at 1 mM except for 6-deoxy-*manno*-heptose, for which concentration was unknown. *kindly provided by Prof Rob Field's group, John Innes Centre.

Name	Sample/Standard	Source
GDP-6dBHep	Sample	Purified from <i>E. coli</i> (section 2.2.7)
GDP-6dBHep	Sample	Synthesised <i>in vitro</i> (section 3.2.3.3)
GDP-4K6Hep	Sample	Synthesised <i>in vitro</i> (section 3.2.3.3)
GDP-6dBHep	Sample	Crude extract from <i>E. coli</i>
GDP-mannose	Standard	Sigma-Aldrich
Mannose	Standard	Sigma-Aldrich
Sorbitol	Standard	Sigma-Aldrich
Mannitol	Standard	Sigma-Aldrich
6-deoxy- <i>manno</i> -heptose	Standard	Synthesised by JIC*
M7P	Standard	Carbosynth
S7P	Standard	Carbosynth
M1P	Standard	Synthesised <i>in vitro</i> (section 3.2.3.1)
Blank	Standard	No compound negative control

For purified GDP-6dBHep, the dry weight of the powdered form (~ 1 mg in each case) was resuspended in 100 μ L 20 mM TEAA (pH 6.0). For GDP-6dBHep, GDP-4K6Hep and M1P synthesised *in vitro* the whole reaction volume (200 μ L) was mixed to final concentration 80 % (v/v) methanol and centrifuged at 3,000 $\times g$ for 30 min at 4 $^{\circ}$ C to remove the proteins. The supernatant was decanted, dried by vacuum centrifugation and the resultant pellet resuspended in 100 μ L 20 mM TEAA (pH 6.0). Each commercially-obtained standard was resuspended in 100 μ L 20 mM TEAA (pH 6.0) to a final concentration of 1 mM. The crude extract was prepared as described in section 2.2.5 and a 100 μ L aliquot was taken. The aliquot was dried by vacuum centrifugation and the pellet resuspended in 100 μ L

20 mM TEAA (pH 6.0). No compound was added to the blank negative control. Samples and standards were prepared in duplicate. One replicate was hydrolysed with TFA then derivatised and the other was derivatised without the hydrolysis step.

For hydrolysis, TFA was added to the samples to a final concentration of 1 M and heated at 100 °C for 1 h. The samples were then thoroughly dried by vacuum centrifugation and kept in a desiccator overnight.

For derivatisation, 20 µL 20 mg mL⁻¹ methocyanine dissolved in pyridine was added to the samples, which were then transferred to glass MS vials. The vials were sealed with rubber-topped lids and incubated at 37 °C for 2 h. After incubation, 40 µL N-methyl-N-(trimethylsilyl)trifluoroacetamide (MSTFA) plus 12.5 % (v/v) mixed alkanes internal standard (1.2mg mL⁻¹ decane, tetradecane and octadecane) was added to the samples by injecting through the rubber lids. Finally, the samples were incubated at 37 °C for 30 min.

After derivatisation, the samples were immediately analysed on an Agilent 7200 series GC-Q-TOF with 7890A GC. The following method was developed by Debbie Salmon, University of Exeter: Injection volume for both samples and standards was 0.2 µL. The split vent ratio was 10:1 and inlet source temperature 250 °C. The emission current was 35 µA and emission voltage 70 eV. A 3.0 mL min⁻¹ septum purge flow was applied. A ZB Semi-volatiles GC column from Phenomenex (30 µm x 250 µm x 0.25 µm) was used with a constant carrier gas helium flow of 1.2 mL min⁻¹. Temperature was held at 70 °C for 4 min, then ramped up to 310 °C at a rate of 15 °C min⁻¹, with a final hold at 310 °C for 6 min. Ions within the range 50-600 atomic mass unit were detected with an acquisition rate of 5 spectra s⁻¹.

2.2.8.4 Nuclear magnetic resonance

A 1.4 mg freeze-dried sample of purified GDP-6ddHep was kindly analysed with ^1H and ^{31}P Nuclear magnetic resonance (NMR) by Dr. Martin Rejzek at the John Innes Centre, Norwich, UK. Firstly, a heavy water (D_2O) blank was performed for both ^1H and ^{31}P NMR GDP-6ddHep was then compared with GDP-mannose, guanosine monophosphate (GMP), triethylammonium acetate (a solvent present in the sample), methyl-6-deoxy- α -D-*manno*-heptoside and D-*manno*-heptose (Figure 48).

2.3 Results

2.3.1 Plasmid construction for GDP-6dBep biosynthesis

2.3.1.1 Synthetic genes for plasmid construction

Six genes are responsible for the biosynthesis of the initial sugar nucleotide, GDP-6dBep, for CPS assembly in *B. pseudomallei*. The genes, *gmhA*, *wcbL*, *wcbN*, *wcbM*, *wcbK* and *wcbJ*, were codon optimised for expression in *E. coli* and synthesised in pEX vectors (Figure 20). The pEX vectors have the pUC origin of replication with a 500-700 copy number. All of the genes, except *wcbL*, were synthesised into pEX-A with the Amp^R gene for ampicillin resistance. *wcbL* was synthesised into the pEX-K vector with the NeoR/KanR gene for kanamycin resistance.

There are four conserved BioBricks™ restriction sites. *EcoRI* and *XbaI* are situated upstream of the gene at positions 0 and 7 in all six plasmids. *SpeI* and *PstI* are situated four nucleotides downstream of the gene. The maps indicate an additional *EcoRI* restriction site in all six plasmids that could complicate excision of the gene from the vector (Figure 20). However, the proximity of the erroneous *EcoRI* sites to the correct *SpeI* and *PstI* sites meant that no problems occurred during assembly of the GDP-6dBep biosynthesis plasmid. When cut with *EcoRI* and either *SpeI* or *PstI* the additional fragment was only 25 or 35 base pairs in length and therefore easy to screen out.

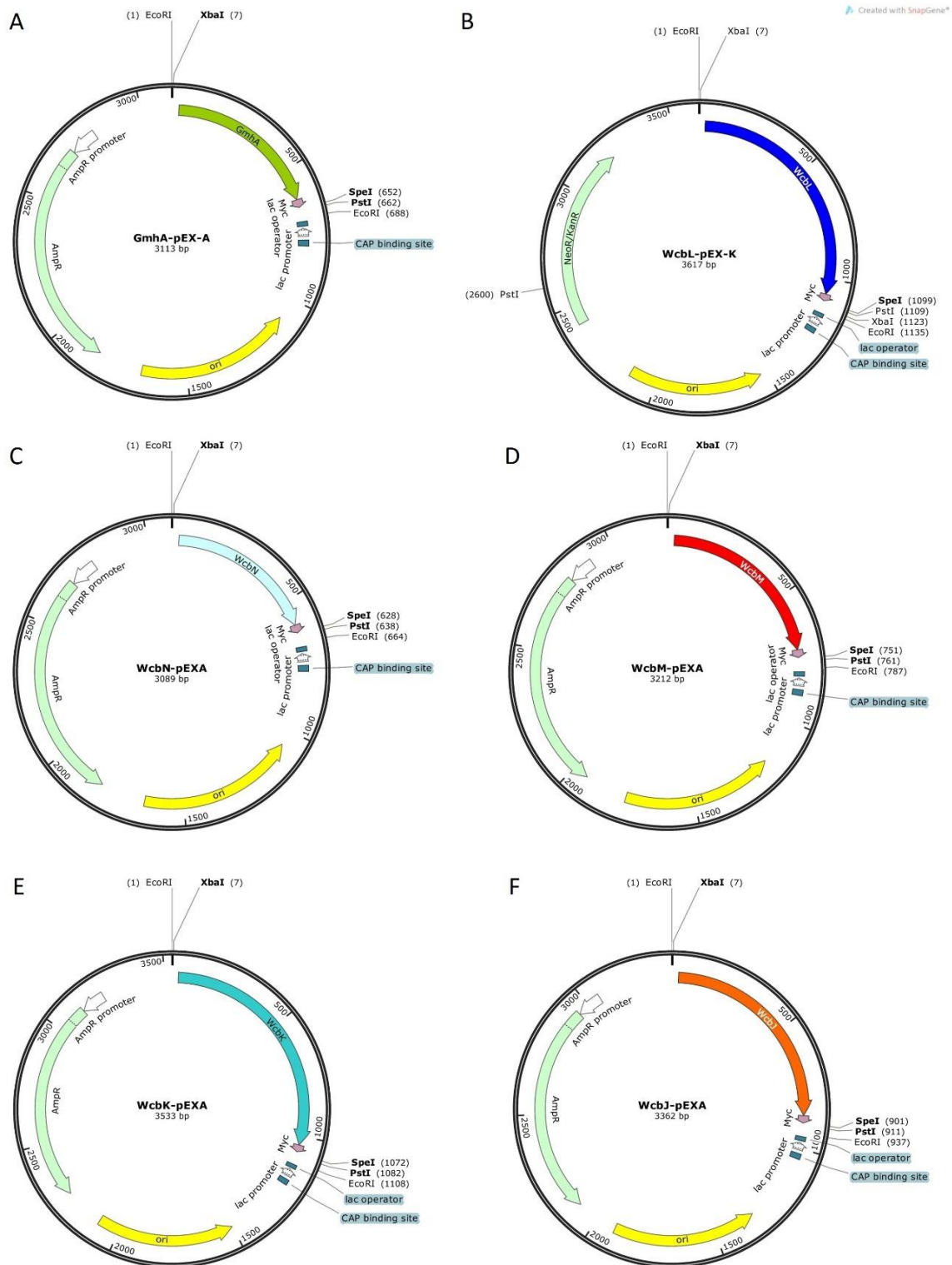


Figure 20: pEX plasmids containing the six CPS biosynthetic genes. (A) *gmhA*; (B) *wcbL*; (C) *wcbN*; (D) *wcbM*; (E) *wcbK* and (F) *wcbJ*. The position of the c-Myc tag and BioBricks™ restriction sites, *EcoRI*, *XbaI*, *SpeI* and *PstI* are indicated. Amp^R gene and NeoR/KanR genes on the vectors confer ampicillin and kanamycin resistance respectively. *ori* = pUC origin of replication. Made with SnapGene v. 3. 1.

The *WcbL_pEX-K* plasmid also has additional *XbaI* and *PstI* sites (Figure 20B). Digestion of the plasmid with *XbaI* and *PstI* would result in the correct fragment (1,102 bp) and three additional fragments (14 bp, 1,477 bp and 1,024 bp). This would have greatly reduced the cloning efficiency of the correct fragment into the destination vector. To avoid this, *wcbL* was inserted upstream of *gmhA-wcbN*, which required digestion with *EcoRI* and *SpeI*, avoiding use of *XbaI* and *PstI* (Figure 21).

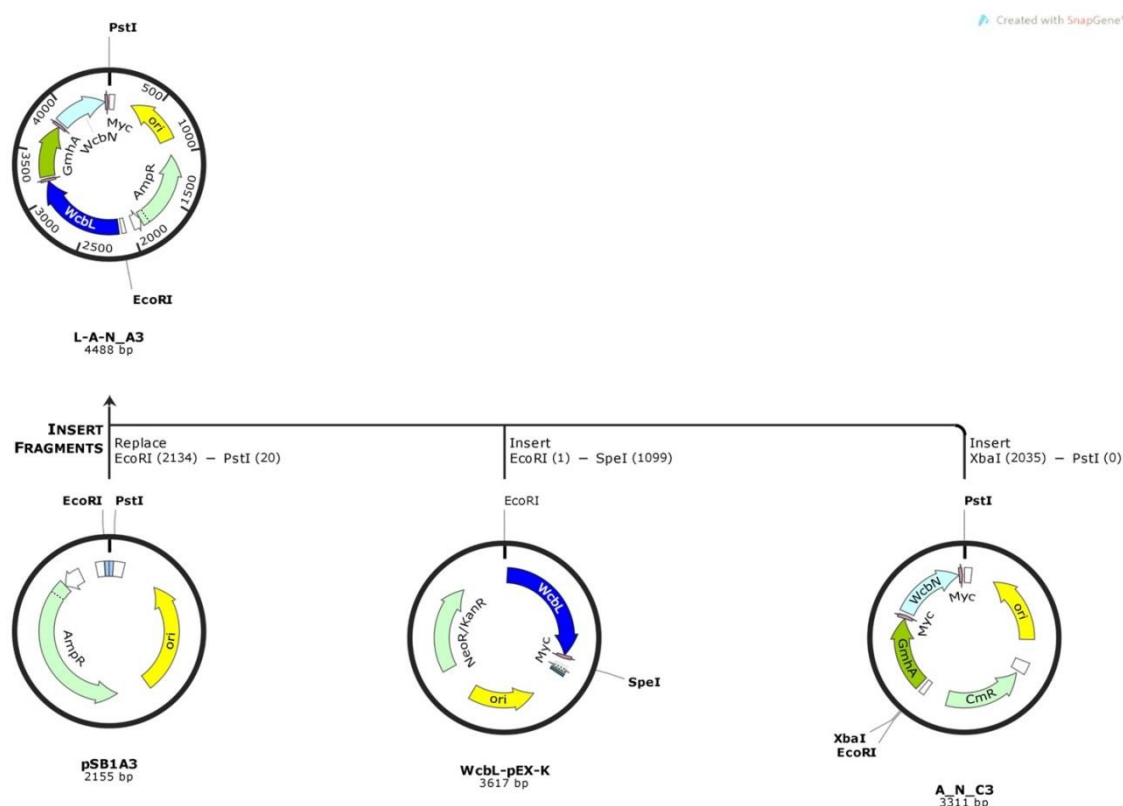


Figure 21: Addition of *wcbL* upstream of *gmhA-wcbN* in pSB1A3. The BioBricks™ 3A assembly method was used to create the new plasmid for construction of the GDP-6ddHep biosynthesis operon. *gmhA-wcbN_pSB1C3* was digested with *XbaI* and *PstI*, *wcbL_pEX-K* with *EcoRI* and *SpeI* and the destination vector pSB1A3 with *EcoRI* and *PstI*. Following ligation, a new plasmid, *wcbL-gmhA-wcbN_pSB1A3*, was created.

2.3.1.2 Assembly of GDP-6ddHep biosynthesis plasmid

After each ligation, the newly created insert becomes a transferable part flanked by the conserved BioBricks™ restriction sites. For example, in Figure 21, *gmhA-wcbN* is excised as a single fragment for transfer into the destination vector. With

the addition of *wcbL*, a new transferable fragment is generated (*wcbL-gmhA-wcbN*) (Figure 21). The full schematic for the sequential assembly of the GDP-6dbHep biosynthetic pathway can be found in Appendix 6. After each cloning step, colonies of transformed *E. coli* DH5 α were screened for the new plasmid construct by RD and PCR. The RD was used to verify formation of a new fragment of the correct size with no internal restriction sites. It also showed whether the vector backbone was the expected size. On average, each successful step had a cloning efficiency of 1 in3 (Figure 22).

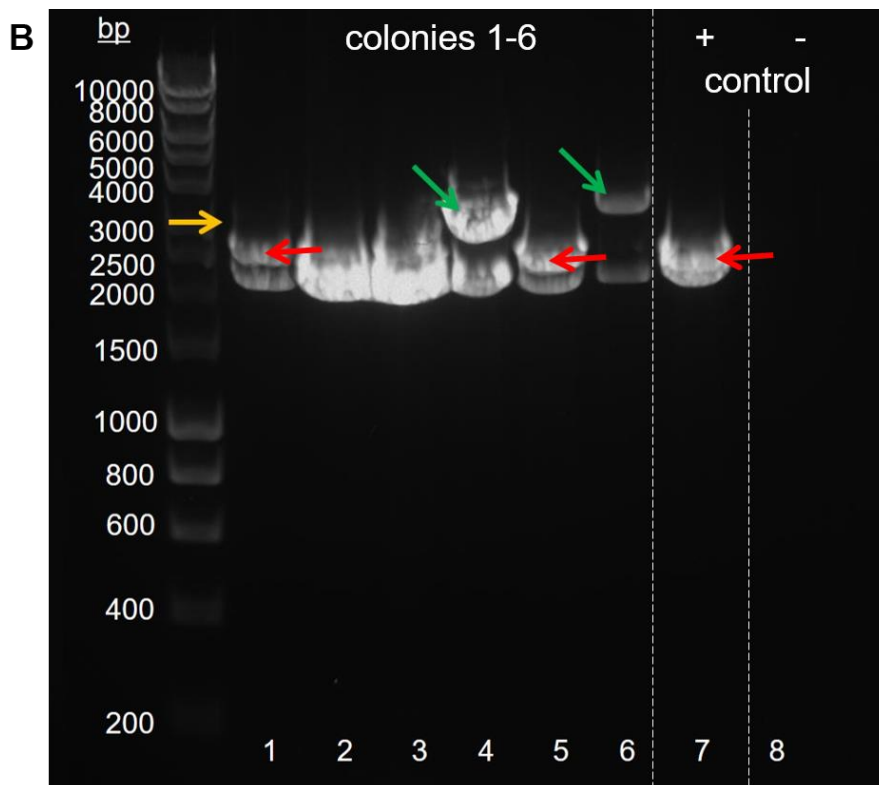


Figure 22: Confirmation of assembly of *wcbM-wcbL-gmhA-wcbN* in pSB1C3. To verify construction of *wcbM-wcbL-gmhA-wcbN_pSB1C3*, plasmids were purified from colonies of transformed *E. coli* following BioBricks™ 3A assembly cloning (Appendix 6). Purified plasmids were digested with *EcoRI* and *PstI*, then run on an agarose gel for visualisation. Lanes 1-6 are screened colonies. Orange arrow indicates expected size of *wcbM-wcbL-gmhA-wcbN* (3,112 bp). Green arrows indicate bands of correct size (colonies 4 and 6). Lane 7 shows a positive control using *wcbL-gmhA-wcbN_pSB1A3* to indicate the RD had worked. Red arrow indicates *wcbL-gmhA-wcbN* (2,368 bp) and is also seen in colonies 1 and 5. Bands with no arrow indicator are vector DNA. Lane 8 is negative control to check for contamination of the RD.

Plasmids with bands of the correct size were further verified by PCR as it has a higher specificity and can confirm fragments are assembled in the correct orientation. From Figure 22, the plasmids from lanes 4 and 6 were analysed by PCR (Figure 23).

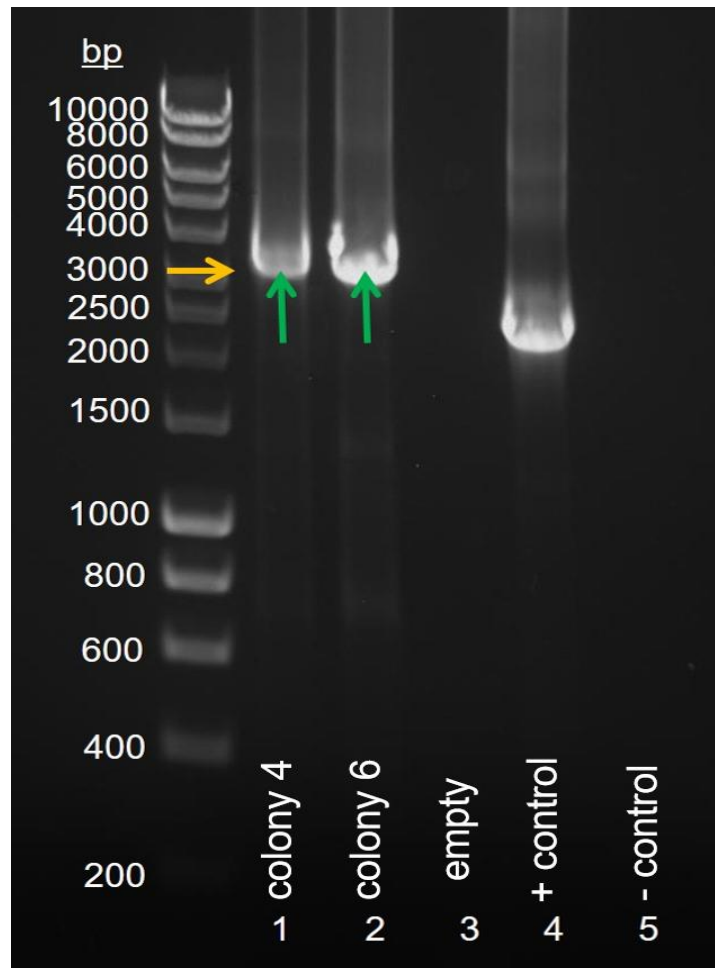


Figure 23: PCR of *wcbM-wcbL-gmhA-wcbN*_pSB1C3. Plasmids were purified from colonies 4 and 6 (RD in Figure 22). PCR using the forward primer for *wcbM* and the reverse primer for *wcbN*. Lanes 1 & 2: *wcbM-wcbL-gmhA-wcbN*_pSB1C3; Lane 3: empty lane; Lane 4: *wcbL-gmhA-wcbN*_pSB1A3 using *wcbL* F primer and *wcbN* R primer; Lane 5: no template negative control. (Primer sequences in Appendix 1).

Each cloning step was analysed. Figure 24 shows progressive construction of the GDP-6dBep plasmid. A schematic (generated by SnapGene v. 3.1.) of the full progression is shown in Appendix 6.

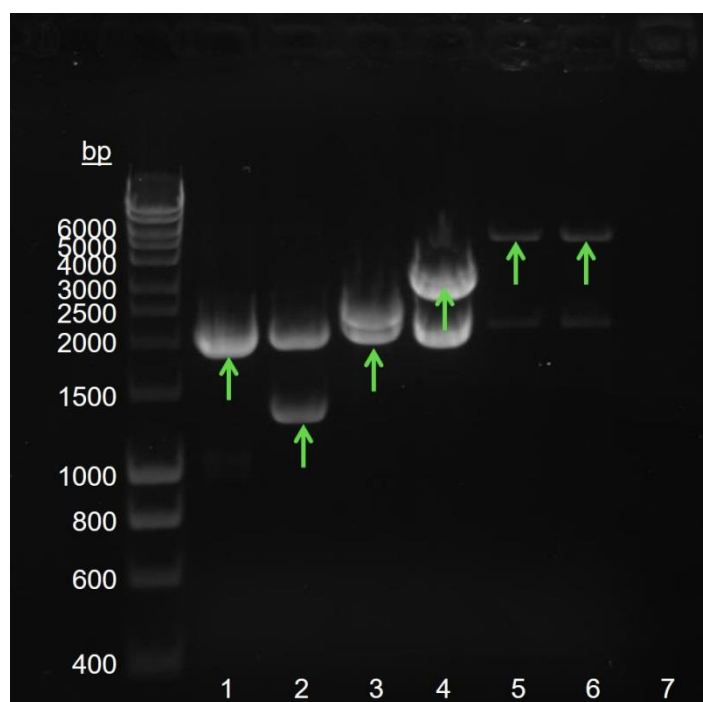


Figure 24: Progressive construction of the plasmid for GDP-6dbHep biosynthetic operon. Agarose gel showing RD with *EcoRI* and *PstI* of each stage of BioBricks™ 3A assembly. Lane 1: *wcbJ* plus *wcbK* (expected size: 1,965 bp) in pSB1C3 (2,029 bp); Lane 2: *gmhA* plus *wcbN* (expected size: 1,276 bp) in pSB1C3 (2,029 bp); Lane 3: *wcbL* plus *gmhA-wcbN* (expected size: 2,368 bp) in pSB1A3 (2,114 bp); Lane 4: *wcbM* plus *wcbL-gmhA-wcbN* (expected size: 3,112 bp) in pSB1C3 (2,029 bp); Lanes 5&6: *wcbJ-wcbK* plus *wcbM-wcbL-gmhA-wcbN* (expected size: 5,057 bp) in pSB1K3 (2,163 bp). Green arrow indicates band of expected size and bands with no arrow are vector DNA.

The expected sizes of the intermediate fragments generated during the construction of the GDP-6dbHep biosynthesis plasmid are shown in Table 5 and correspond with the fragments in the RD in Figure 24.

Table 5: Length (bp) of genes and fragments used for CPS biosynthetic operon assembly

Gene/fragment/regulator	Length (bp)	Vector	Vector length (bp)
<i>wcbJ</i>	900	pEX-A	2,450
<i>wcbK</i>	1,075	pEX-A	2,450
<i>wcbJ-wcbK</i>	1,965	pSB1C3	2,029
<i>gmhA</i>	651	pEX-A	2,450
<i>wcbN</i>	631	pEX-A	2,450
<i>gmhA-wcbN</i>	1,276	pSB1C3	2,029
<i>wcbL</i>	1,098	pEX-K	2,507
<i>wcbL-gmhA-wcbN</i>	2,368	pSB1A3	2,114
<i>wcbM</i>	750	pEX-A	2,450
<i>wcbM-wcbL-gmhA-wcbN</i>	3,112	pSB1C3	2,029
<i>wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN</i>	5,057	pSB1K3	2,163
BBa_J23118 (<i>p</i>)	35	J61002	2,948
BBa_B0015 (<i>tt</i>)	129	Unknown	Unknown

2.3.1.3 Terminator and promoter sequences

The BioBricks™ constitutive promoter BBa_J23118 (denoted ‘*p*’) and double terminator BBa_B0015 (denoted ‘*tt*’) were added to the operon (Table 6). The terminator, a 129 bp sequence, was added to *wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN* using the normal 3A assembly method described. However, as the promoter is only 35 bp long compared to the 5,257 bp *wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt* fragment, the cloning efficiency was massively reduced and it was difficult to get confirmation of successful insertion by RD. The BioBricks™ 3A assembly method was adapted by digesting *wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_pSB1C3* with *EcoRI* and *XbaI* and the promoter sequence with *EcoRI* and *SpeI*. This effectively opened up *wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_pSB1C3* and enabled the promoter to be inserted directly to construct the final GDP-6dbHep biosynthetic plasmid *p-wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_pSB1C3* (Figure 25; Figure 27).

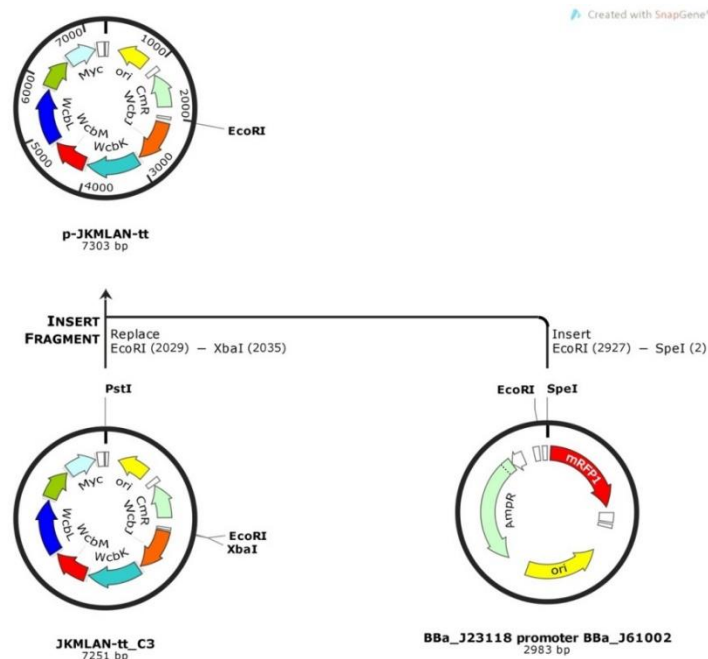


Figure 25: Insertion of BioBricks™ promoter BBa_J23118 into *wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_pSB1C3*. BBa_J23118 was excised with *EcoRI* and *SpeI* and ligated into *wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_pSB1C3* which had been linearised with *EcoRI* and *XbaI*.

The completed operon was initially verified by RD and PCR as before using the forward primer for promoter BBa_J23118 and terminator BBa_B0015 (Figure 26; primer sequences in Appendix 1).

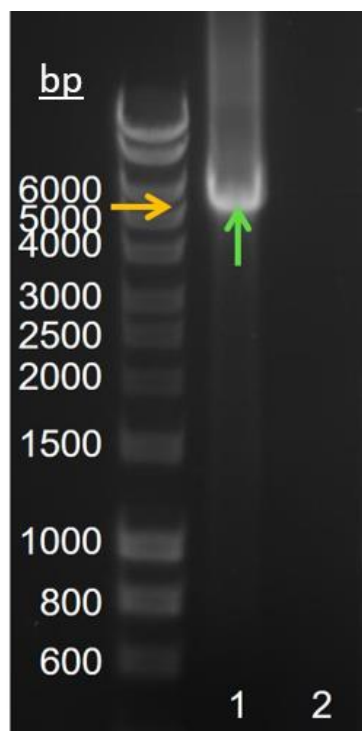


Figure 26: PCR of completed GDP-6dBep biosynthetic operon in BioBricks™ vector pSB1C3. Using a forward primer for BBa_J23118 and a reverse primer for BBa_B0015. Lane 1: GDP-6dBep biosynthesis pathway (expected size 5,257 bp, orange arrow; actual size, green arrow); Lane 2: Negative control (no template). Primer sequences in Appendix 1.

The final plasmid is 7,303 bp in length, of which 2,029 bp comprise the pSB1C3 vector backbone. The operon is under the constitutive promoter BBa_J23118. Each gene has a pBAT4 RBS and c-Myc tag. A strong double terminator, BBa_B0015 completes the operon. The restriction sites *EcoRI*, *XbaI*, *SpeI* and *PstI* are conserved, flanking the operon (Figure 27).

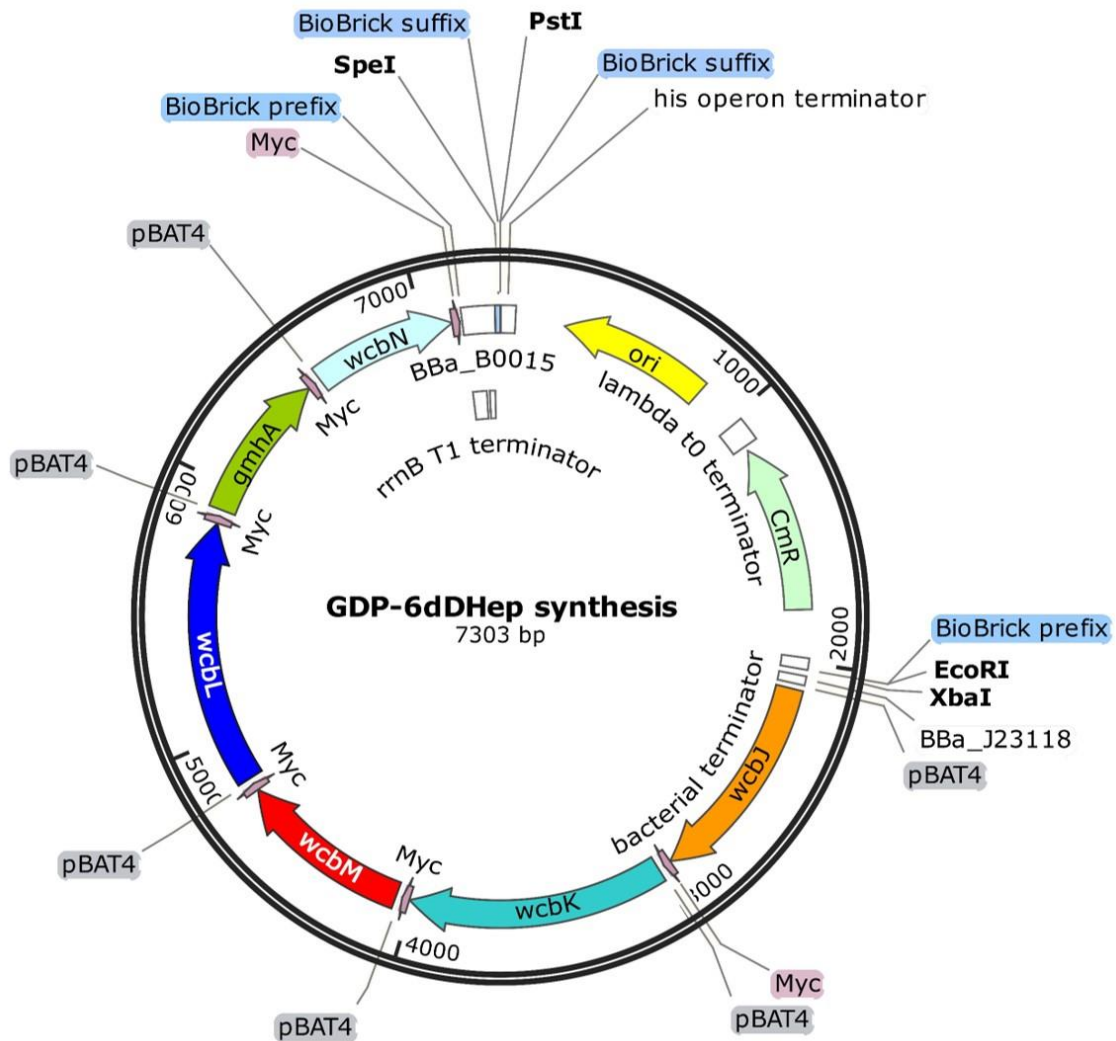


Figure 27: Final plasmid construct for GDP-6dHep biosynthesis. Vector backbone (position 0-2,029 bp) from BioBricks™ plasmid pSB1C3. The pathway consists of BioBricks™ promoter BBa_J23118 followed by genes *wcbJ*, *wcbK*, *wcbM*, *wcbL*, *gmhA* and *wcbN* and BioBricks™ terminator BBa_B0015. Each gene preceded by pBAT4 RBS and followed by c-Myc tag. Total length: 7,303 bp. Made using SnapGene v. 3.1.

2.3.2 Verification of plasmid expression

2.3.2.1 Plasmid DNA sequence

Fifteen primers were used for sequencing the plasmid, eight in the forward direction and seven in reverse. This resulted in almost the entire plasmid having two sequence reads (Figure 28; Figure 29).

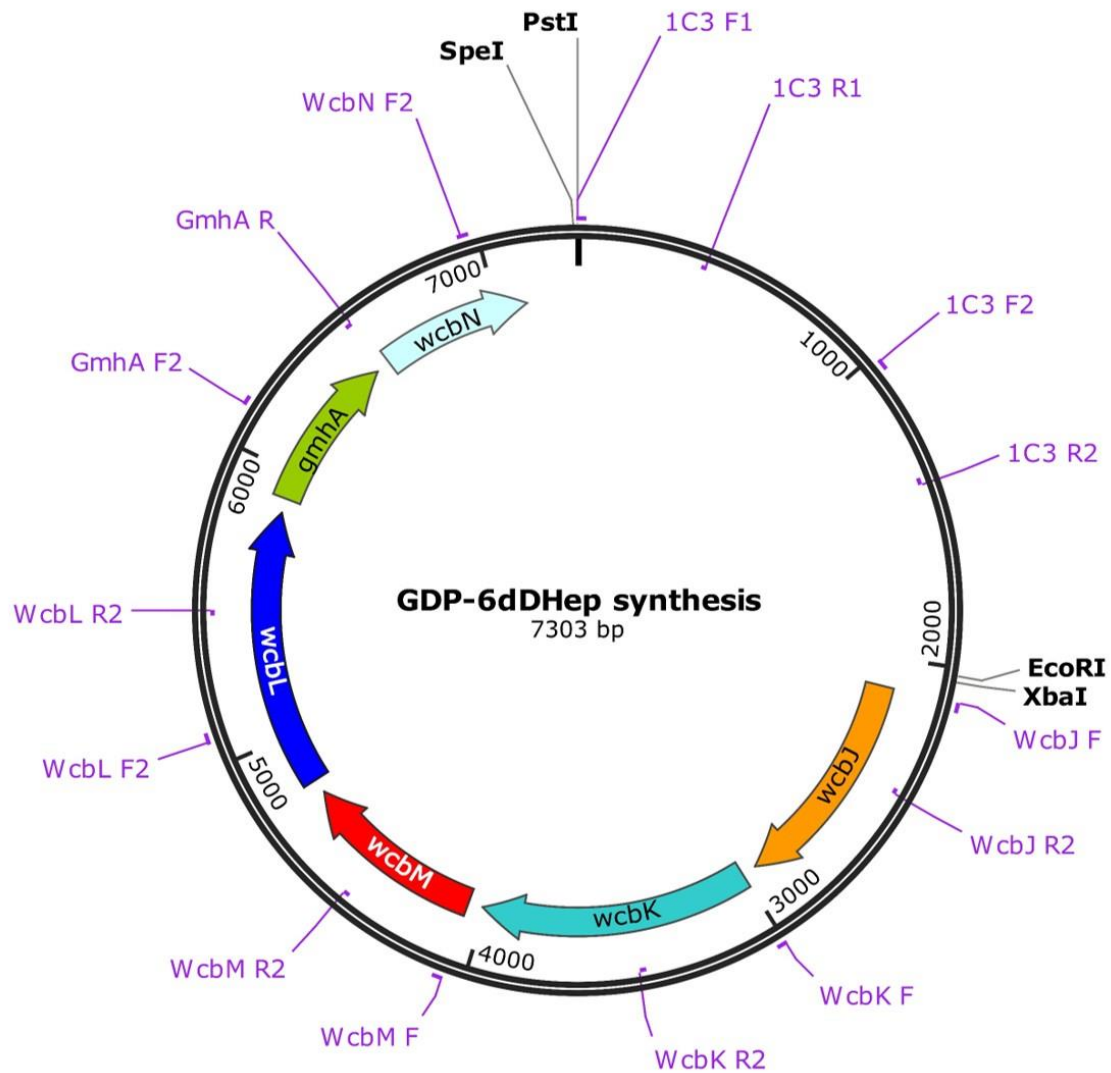


Figure 28: GDP-6dDHep biosynthesis plasmid with primer position and direction for sequencing indicated. Primer sequences in Appendix 1. Made using SnapGene v. 3.1.

There is a region of misalignment towards the 3' end of the template DNA which is confirmed in three overlapping sequence reads (Figure 29). However, the open reading frame aligns with the six genes and the downstream misalignment does not disrupt the operon.

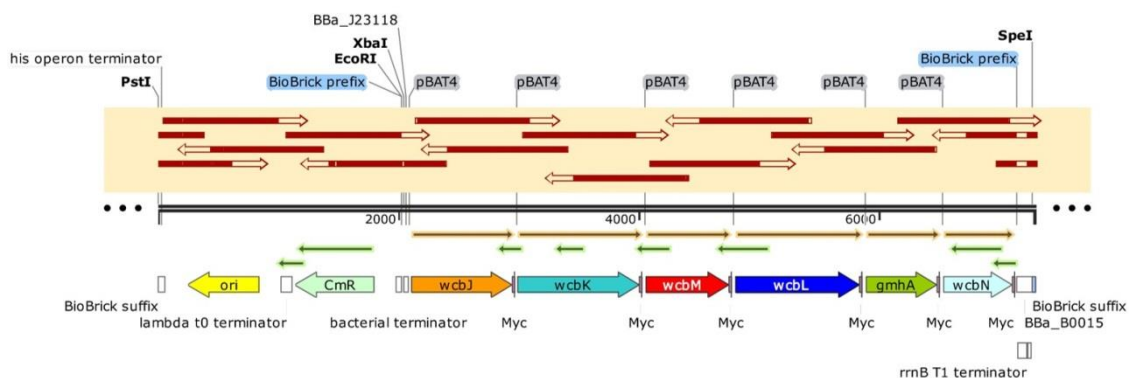


Figure 29: Sequencing reads for GDP-6dBHep biosynthesis plasmid. Grey bar represents the template (plasmid) DNA. Orange box: Arrows pointing right are 5' (forward) reads and arrows pointing left are 3' (reverse) reads. Filled arrows denote sequence consensus between the read and template DNA, open arrows denote sequence mismatches. Arrows below template DNA denote open reading frames in 5' (forward; orange) and 3' (reverse; green) directions. Also shown are the BioBricks™ restriction sites. Made using SnapGene v. 3.1.

The region of misalignment begins with an 88 bp deletion from position 7,155 to 7,242. This removes the entirety of the rrnB T1 terminator complex from the BioBricks™ sequence Bba_B0015 (Figure 30). This may be a genuine deletion or could be an artefact caused by the repetitiveness of the DNA sequence and looped secondary structure of the terminator. A 54 bp insertion occurs 47 bp downstream of the deletion between positions 7,289 and 7,290. The insertion disrupts the BioBricks™ suffix sequence. A nucleotide BLAST search of the 54 bp insert and the template sequence (including the insertion) from position 7,149 to 7,303 reveals 100 % identity with several plasmid and cloning components. This suggests the sequence is a non-coding fragment of a vector backbone. It is likely there was an error in the Bba_B0015 BioBricks™ part used in the construction of the GDP-6dBHep biosynthesis plasmid. However, as the T7Te segment of the double terminator is left intact and there is no effect on the coding sequences of the plasmid, these two errors are very unlikely to affect plasmid expression.

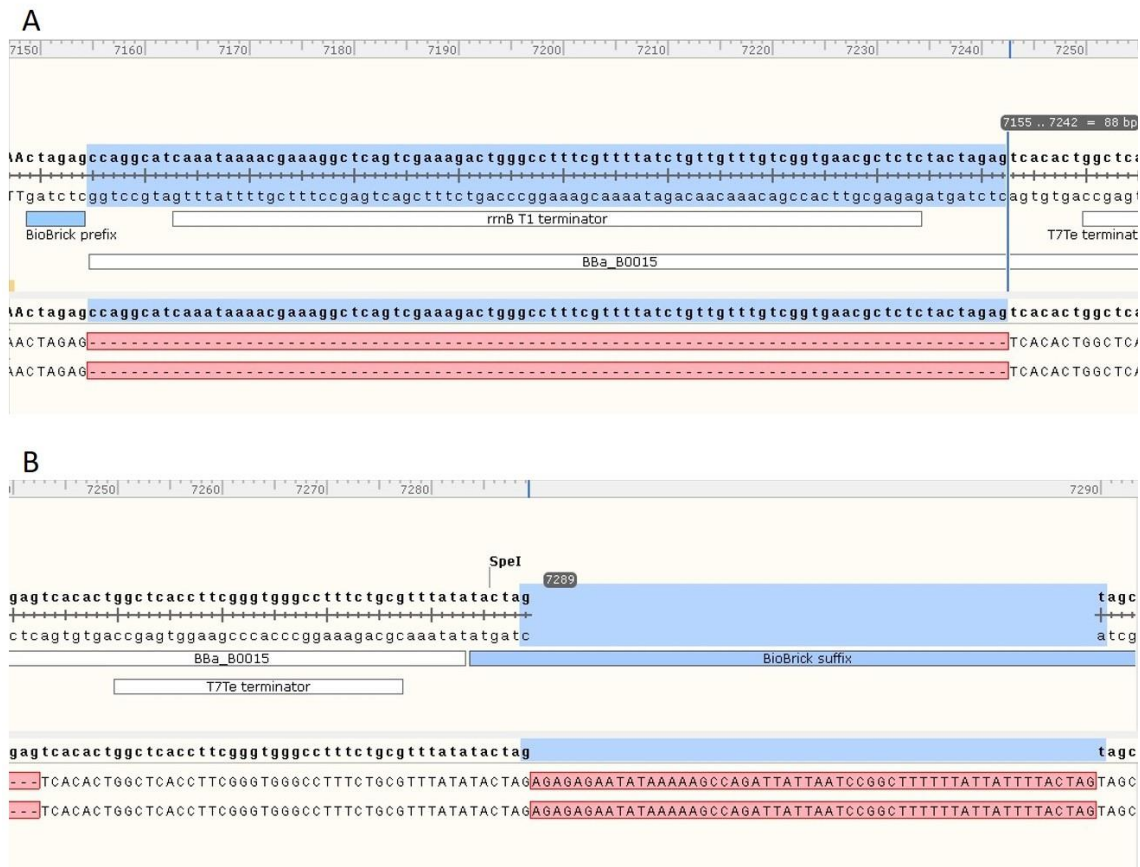


Figure 30: Sequencing data showing an 88 bp deletion and 54 bp insertion in the GDP-6dBep biosynthesis plasmid. (A) The deletion, shown in two sequence reads (red), occurs from position 7,155 to 7,242 (blue) immediately after the BioBricks™ prefix. Part of the BioBricks™ BBa_B0015 double terminator sequence (rrnB T1) is erased. (B) The insertion, shown in two sequence reads (red), occurs from between positions 7,289 and 7,290 (blue) disrupting the BioBricks™ suffix. Made using SnapGene v. 3.1.

One other error was identified by the sequencing data; a base change from A to G at position 172. This is in a non-coding region and will have no effect on the function or expression of the plasmid.

2.3.2.2 RNA expression (transcription)

To confirm the genes were being transcribed from the plasmid, total mRNA was extracted from the transformed *E. coli* and analysed by RT-PCR. *E. coli* DH5 α expressing empty pSB1C3 vector was used as a negative control. GAPDH, a constitutively expressed housekeeping gene, was also analysed as a positive control (data not shown here). Removing DNA to obtain a pure RNA extraction

and prevent false positives required three applications of the decontamination protocol (Ambion TURBO DNA-free™ Kit).

Agarose gels of the RT-PCR verify transcription of all six biosynthetic genes (Figure 31). For *gmhA* and *wcbL*, anomalous bands can be seen in negative control *E. coli* empty vector cDNA. However, the bands are of incorrect size (~500 bp too large for *gmhA* and ~400 bp too small for *wcbL*) and are probably an artefact of non-specific binding of the primers (Figure 31A; Figure 31B). For *wcbM*, bands of the expected size appear in the *wcbM* mRNA, empty vector mRNA and no reverse transcriptase control (Figure 31D). However, the bands are incredibly faint in comparison to the *wcbM* cDNA and positive control plasmid DNA, suggesting they are a product of a small amount of contaminating DNA. The strength of the cDNA band indicates the gene is being transcribed.

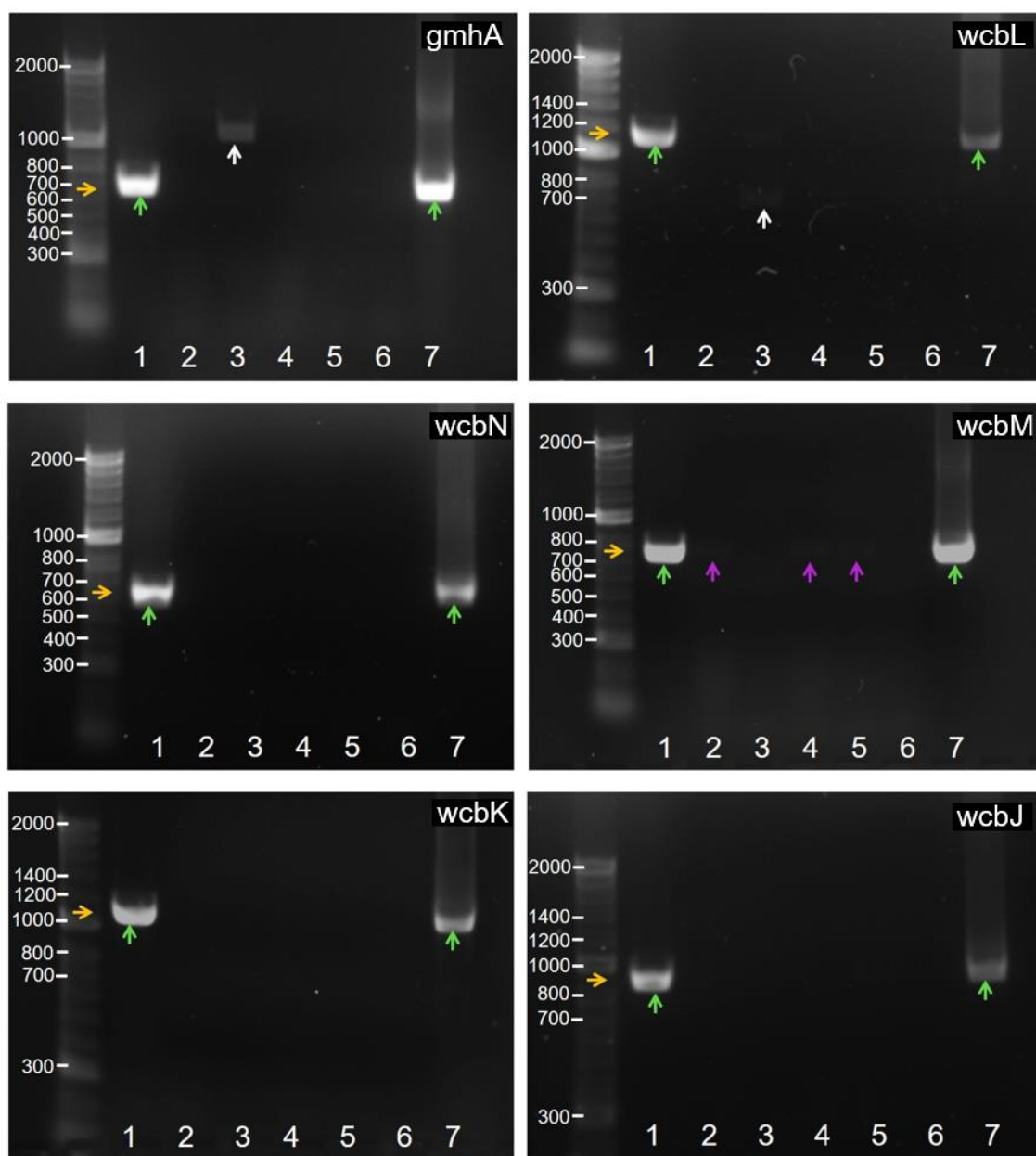


Figure 31: Reverse-transcriptase PCR of the six genes on the GDP-6dHep biosynthesis plasmid. *E. coli* DH5 α transformed with either the GDP-6dHep biosynthetic operon or the empty pSB1C3 vector were grown overnight in LB medium plus 50 μ g/mL chloramphenicol. The RNA was extracted using a Qiagen RNeasy[®] Protect Mini RNA extraction kit. A New England BioLabs[®] 2-step OneTaq[®] RT-PCR kit was used firstly to create cDNA and to then PCR for each gene. The lanes are as follows: (1) biosynthetic gene cDNA (2) biosynthetic gene mRNA (3) empty vector pSB1C3 cDNA (4) empty vector pSB1C3 mRNA (5) negative control, no reverse transcriptase (6) negative control, no nucleotide (7) positive control plasmid DNA of biosynthetic gene. The expected size (orange arrow) and observed size (green arrow) are indicated for each gene. Anomalous bands of incorrect size (white arrow) and contaminating bands of expected size (purple arrows) are indicated where appropriate.

2.3.2.3 Protein expression (translation)

To verify translation, the proteins were extracted by anti-c-Myc chromatography. As all six proteins had the same tag and were similar in size (Figure 32) it was impossible to separate them to confirm expression of each one. SDS-PAGE (Figure 32A) and Western blot (Figure 32B) indicated possible protein expression. Three bands from the SDS-PAGE at approximately 38 kDa, 31 kDa and 25 kDa were also present on the Western blot and could represent expression of WcbL, WcbK, WcbJ and WcbM. However, the Western blot was over-exposed and the c-Myc-tagged protein standard (Thermo Fisher) did not appear in Lane 1. These data were not sufficient to confirm expression of the six *B. pseudomallei* biosynthetic proteins from the plasmid in *E. coli*.

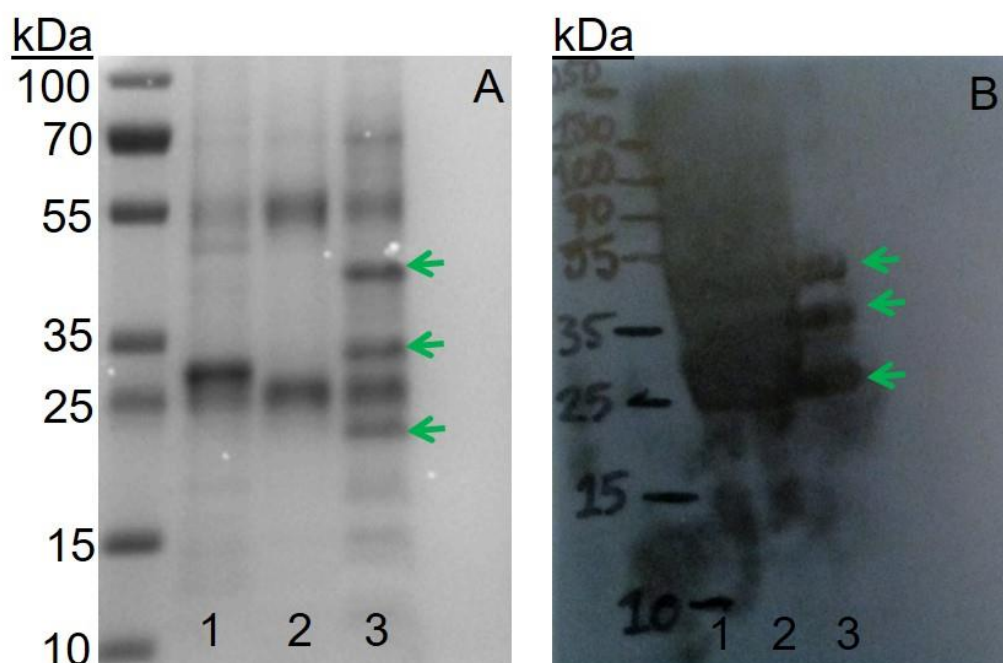


Figure 32: SDS-PAGE and Western blot of c-Myc-tagged CPS biosynthetic proteins. Lane 1: c-Myc-tagged standard control protein (Thermo Fisher); Lane 2: proteins purified from *E. coli* expressing empty pSB1C3 vector; Lane 3: proteins purified from *E. coli* expressing GDP-6dbHep biosynthesis plasmid. Purification was performed by anti-c-Myc chromatography. (A) SDS-PAGE of purified proteins. (B) Anti-c-Myc Western blot of gel (A). Green arrows indicate bands seen in the fraction purified from *E. coli* expressing the CPS biosynthetic operon (Lane 3) but are absent from the control groups (Lanes 1 and 2). These indicate potential CPS biosynthetic proteins.

After anti-c-Myc purification, 2D separation (BN-PAGE followed by SDS-PAGE) was conducted in an attempt to further separate the proteins (Figure 33). Seven spots were present on the gel, three of which lay in the expected size range (20 kDa to 38 kDa).

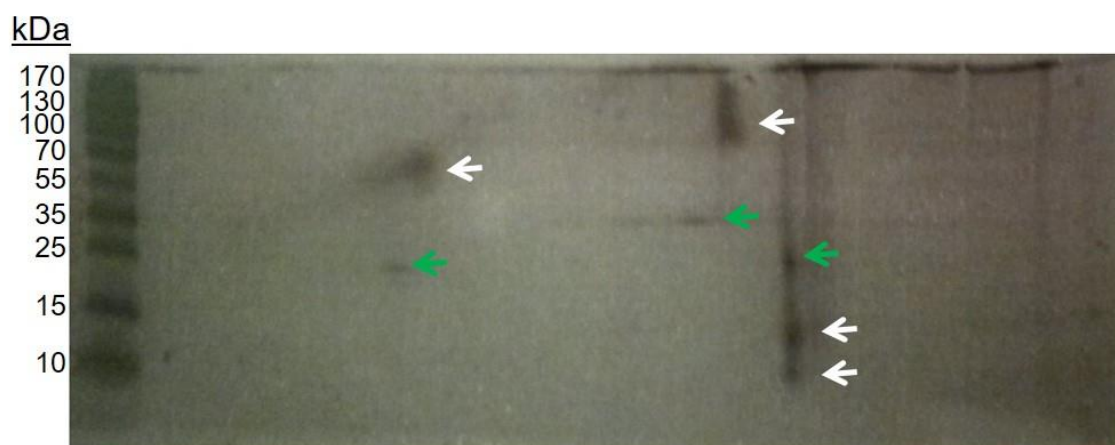


Figure 33: 2D-PAGE of CPS biosynthetic proteins. Purified by anti-c-Myc chromatography from *E. coli* DH5 α grown in 5 mL LB overnight at 37 °C. Green arrows indicate spots that are potentially the correct size, white arrows indicate spots outside the expected size range for the proteins. (GmhA: 20.8 kDa; WcbL: 38.0 kDa; WcbN: 20.4 kDa; WcbM: 24.9 kDa; WcbK: 38.1 kDa; WcbJ: 30.8 kDa)

The 2D-separation was repeated with more protein loaded onto the gel. The gel was used for a Western blot as antibody detection is more conclusive than separation on a gel alone. Four spots were detected by the anti-c-Myc antibody, two of which were in the expected size range (Figure 34.) However, the pattern of spots was not consistent with the previous 2D-PAGE (Figure 33).

No conclusive evidence for expression of *B. pseudomallei* proteins GmhA, WcbL, WcbN, WcbM, WcbK or WcbJ from the GDP-6dBep plasmid in *E. coli* could be obtained. However, the plasmid sequence was correct, transcription of the corresponding genes was confirmed and later experiments demonstrated protein function in the production of GDP-6dBep.

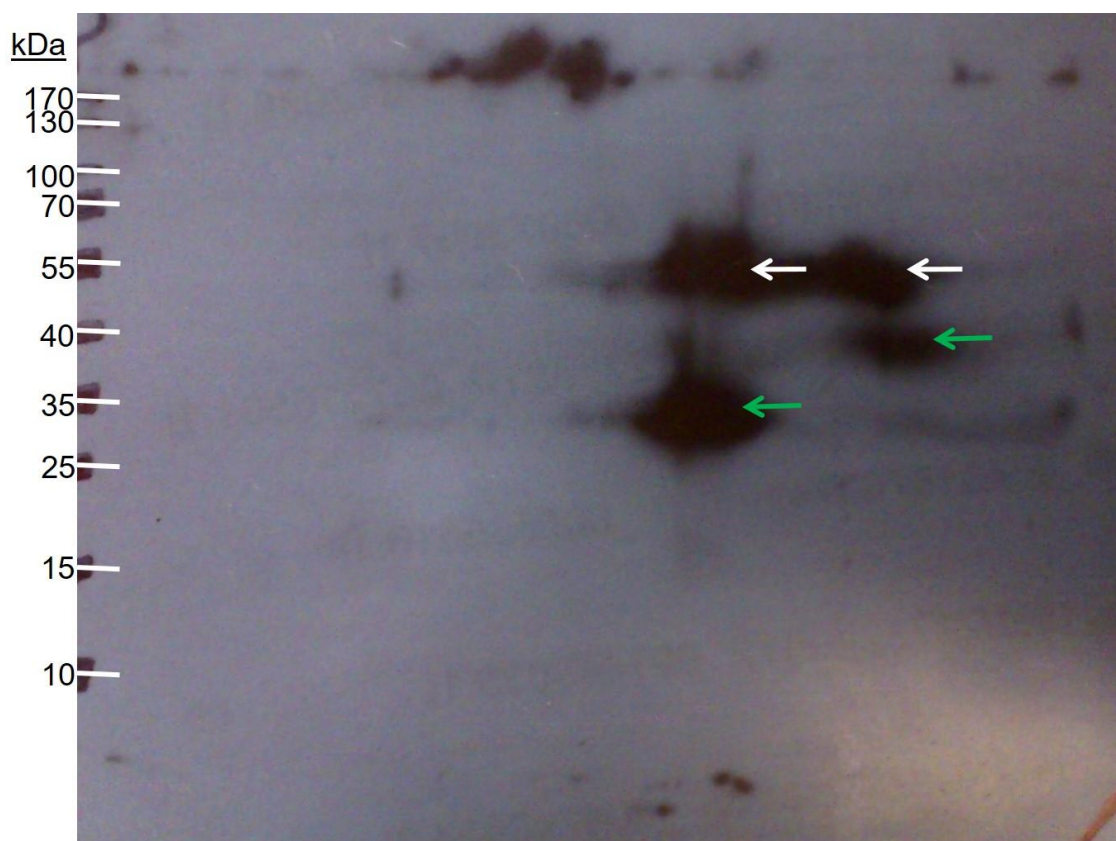


Figure 34: Anti-c-Myc Western blot of 2D-PAGE of CPS biosynthetic proteins. Extracted from *E. coli* DH5 α transformed with GDP-6dBHep biosynthesis plasmid. The proteins were extracted by anti-c-Myc chromatography (Pierce) and underwent a BN- and SDS-PAGE separation. Green arrows indicate spots that are potentially the correct size, white arrows indicate spots outside the expected size range for the proteins. (GmhA: 20.8 kDa; WcbL: 38.0 kDa; WcbN: 20.4 kDa; WcbM: 24.9 kDa; WcbK: 38.1 kDa; WcbJ: 30.8 kDa)

2.3.2.4 *E. coli* growth curve

E. coli DH5 α transformed with the GDP-6dBHep biosynthesis plasmid had an extended lag time (~390 min) compared to *E. coli* transformed with the empty pSB1C3 vector (~180 min) in LB broth cultures at 37 °C (Figure 35). Both cultures achieved similar cell densities of 7.6 OD₆₀₀ for empty pSB1C3 vector and 7.5 OD₆₀₀ for GDP-6dBHep plasmid after 1,440 min (24 h) (data not shown).

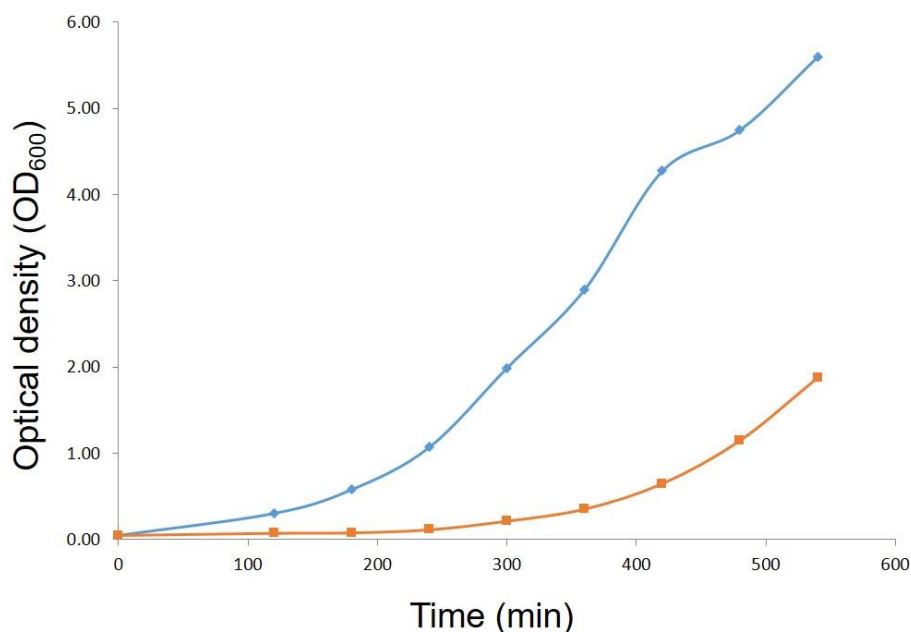


Figure 35: Effect of GDP-6dBep biosynthesis plasmid expression on growth of *E. coli*. Growth curves for *E. coli* DH5 α transformed with empty pSB1C3 vector (blue diamond) or the GDP-6dBep biosynthesis plasmid (red square). 50 mL LB broths containing 50 μ g/mL chl were inoculated to OD₆₀₀ 0.05 and incubated at 37 °C plus aeration. Bacterial growth measured by optical density (OD₆₀₀) over time (min). n = 2 biological replicates.

2.3.3 Total metabolite extraction

Total metabolite extracts from *E. coli* transformed with the GDP-6dBep biosynthesis plasmid and the empty pSB1C3 vector were analysed by Q-TOF. Over 5,000 peaks were identified in the combined positive and negative scans. Of these, 205 peaks from the negative scan and 582 from the positive scan with the biggest difference in abundance between GDP-6dBep plasmid and empty pSB1C3 vector-transformed *E. coli* were identified (analyses by Dr. Hannah Florance and Dr Venura Perera). However, a search for the most abundant compounds in the *E. coli* Metabolome Database (<http://ecmdb.ca/>) and Mass Bank (<http://www.massbank.jp/en/database.html>) failed to identify any known metabolites. The search included a detailed look at the intermediates of the pentose phosphate pathway, as S7P is sunk in to GDP-6dBep production, but no changes to this pathway were identified. Overall the data show there is an

alteration in the metabolome of *E. coli* expressing GDP-6dBep, however it has not been possible to pinpoint any specific changes.

The mass/charge ratio of GDP-6dBep (618.0855 m/z) was seen in the negative scans of the total metabolite extracts. Relative abundance was analysed by determining the peak area from each of the seven samples of *E. coli* transformed with the GDP-6dBep biosynthesis plasmid and the empty pSB1C3 vector (Table 6). The expected retention time was 11.98 min, the observed retention times were 10.41 min and 10.39 min for GDP-6dBep plasmid and pSB1C3 vector-transformed *E. coli* respectively. No peak at 618.0855 m/z was identified in 3 out of 7 of the empty vector *E. coli* and the relative abundance was insignificant in the remaining 4 samples (average 43.73) compared to the average 71,629 relative abundance in all seven samples from the GDP-6dBep *E. coli*.

Table 6: Retention time and relative abundance of GDP-6dBep (618.0855 m/z) determined by Q-TOF LC-MS. Total metabolite extracts were obtained from *E. coli* transformed with the GDP-6dBep biosynthesis plasmid or the pSB1C3 empty vector.

Sample no.	GDP-6dBep plasmid		pSB1C3 empty vector	
	Retention time (min)	Peak area	Retention time (min)	Peak area
1	10.02	41852.64	10.85	30.40
2	10.28	120751.39	10.22	57.75
3	10.34	74805.77	10.22	85.20
4	10.41	48629.87	10.28	132.75
5	10.54	56285.85	-	0.00
6	10.60	83664.66	-	0.00
7	10.66	75413.44	-	0.00
Mean	10.41	71629.09	10.39	43.73

The low abundance of 618.0885 m/z observed in the control empty vector *E. coli* is probably background noise and not GDP-6dBep, as it is not a native sugar nucleotide to *E. coli*. However, as the Q-TOF analysis does not fragment the

compounds, further analysis is required to determine whether the GDP-6dBHep biosynthesis plasmid is expressing GDP-6dBHep.

2.3.4 GDP-6dBHep purification

Sugar nucleotides were purified from metabolite extracts. Anion exchange chromatography shows a large peak at absorbance 254 nm (Figure 36). The peak was collected (fractions 1.A.3., 1.A.4. and 1.A.5.) for further purification.

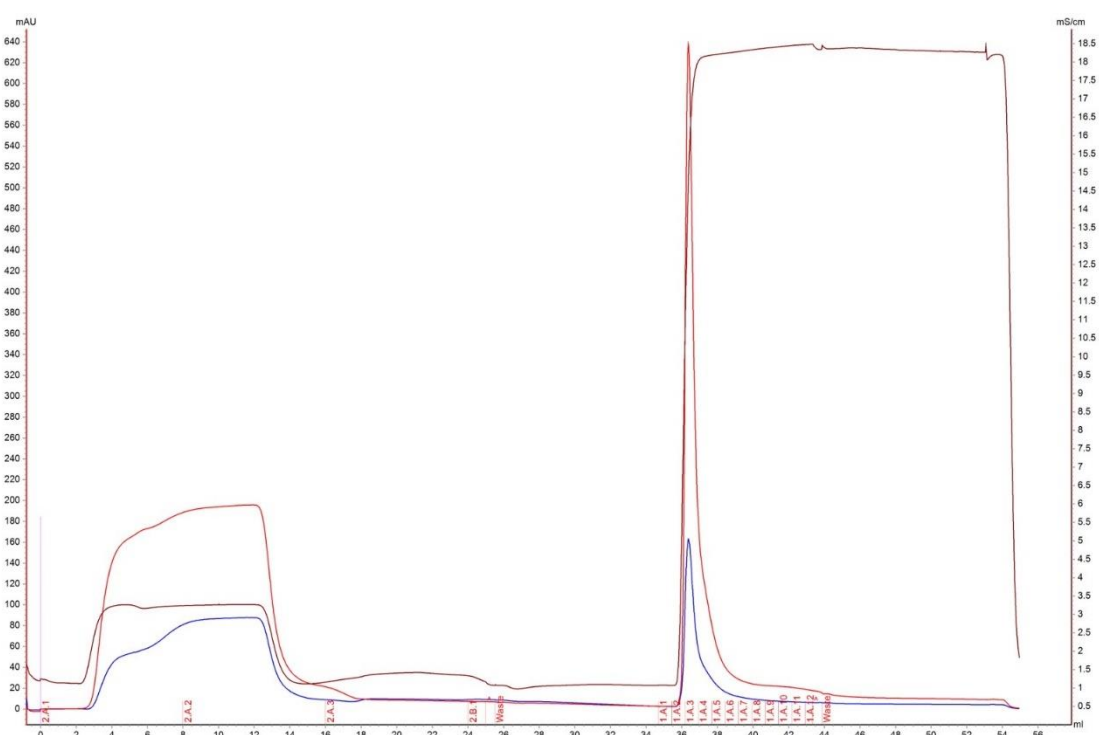


Figure 36: Anion exchange chromatography for the purification of GDP-6dBHep after solid phase extraction. Red = UV absorbance at 254 nm (mAU), blue= UV absorbance at 280 nm (mAU), brown = conductivity (mS cm⁻¹), blue The sugar nucleotide is eluted into fractions 1.A.1.-1.A.12. Only fractions 1.A.3., 1.A.4. and 1.A.5. were kept for further purification.

The fractions collected from anion exchange were vacuum centrifuged to remove excess ammonium bicarbonate and resuspended in 20 mM TEAA (pH 6.0). They were further purified by HPLC (Figure 37).

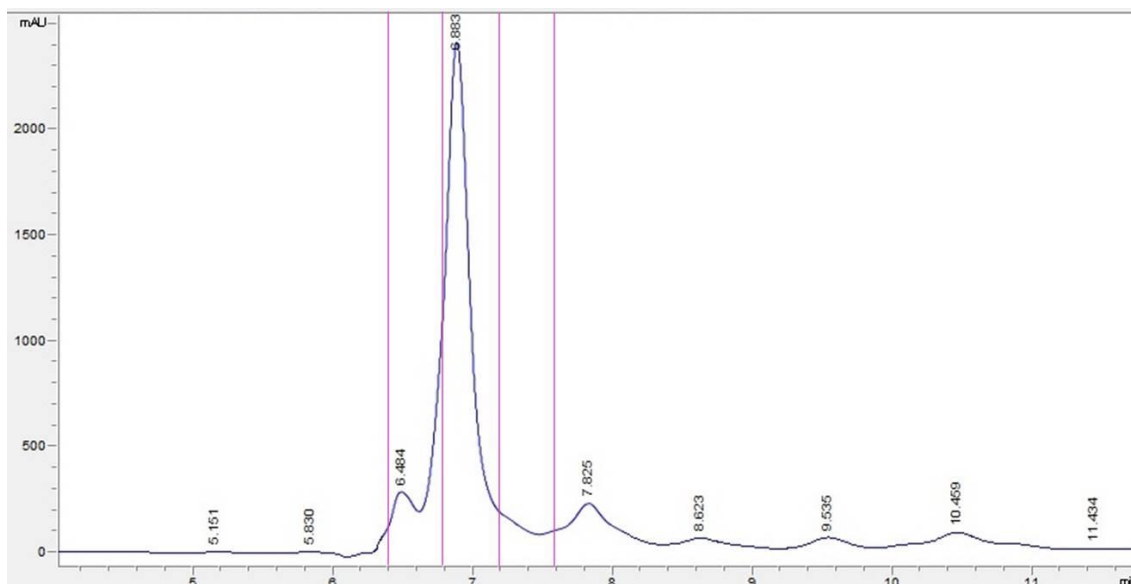


Figure 37: HPLC of GDP-6dbHep after solid phase extraction and anion exchange chromatography. Three fractions were collected between 6.4 min – 6.8 min, 6.8 min – 7.2 min and 7.2 min – 7.6 min (pink lines). The peak eluted during the isocratic phase in 20 mM TEAA (pH 6.0) buffer at a flow rate of 0.2 mL min⁻¹.

Samples of extracted from *E. coli* transformed with empty pSB1C3 vector were purified concurrently with the GDP-6dbHep samples for MS analysis as negative controls (data not shown).

2.3.5 GDP-6dbHep structure verification

2.3.5.1 LC-MS Q-TOF

The purified samples of GDP-6dbHep were analysed by LC-MS Q-TOF. The accuracy score for GDP-6dbHep on the Agilent Technologies 6520 Accurate Q-TOF mass spectrometer was calculated by Dr. Debbie Salmon. Find by Formula (FBF) predicts the chemical formula by calculating the difference in m/z ratio between the expected m/z (618.0850) and observed m/z (618.0845) of the parent ion and isotopes (Figure 38A). The predicted formula was C₁₇H₂₇N₅O₁₆P₂ with an accuracy score of 97.37 % and 1.9 ppm difference between expected and observed m/z values. (Accuracy score = overall match score weighted by different identification techniques used by the Mass Hunter software. Results are

specific to the instrument and software used.) Retention time in negative mode was 8.45 min. The predicted fragments included: 442.0149, 424.0057 and 344.0407 m/z (Figure 38B).



Figure 38: Accuracy score for GDP-6dBep from Q-TOF scan in negative mode. (A) Find By Formula spectrum of for C₁₇H₂₇N₅O₁₆P₂. Red boxes show isotopes; red lines denote observed isotopes. (B) expected fragments from GDP-6dBep.

Samples eluted from the HPLC were analysed by LC-MS Q-TOF to check for purity. Figure 39 shows there are four contaminants with m/z 524.7175, 560.2289, 595.7413 and 611.2197. The signal for GDP-6dBep (618.0855) shows a double peak, which is indicative of ion suppression. A search of the *E. coli* Metabolome Database (<http://ecmdb.ca/>) and Mass Bank (<http://www.massbank.jp/en/database.html>) failed to identify any known compounds at the m/z ratios of the contaminants.

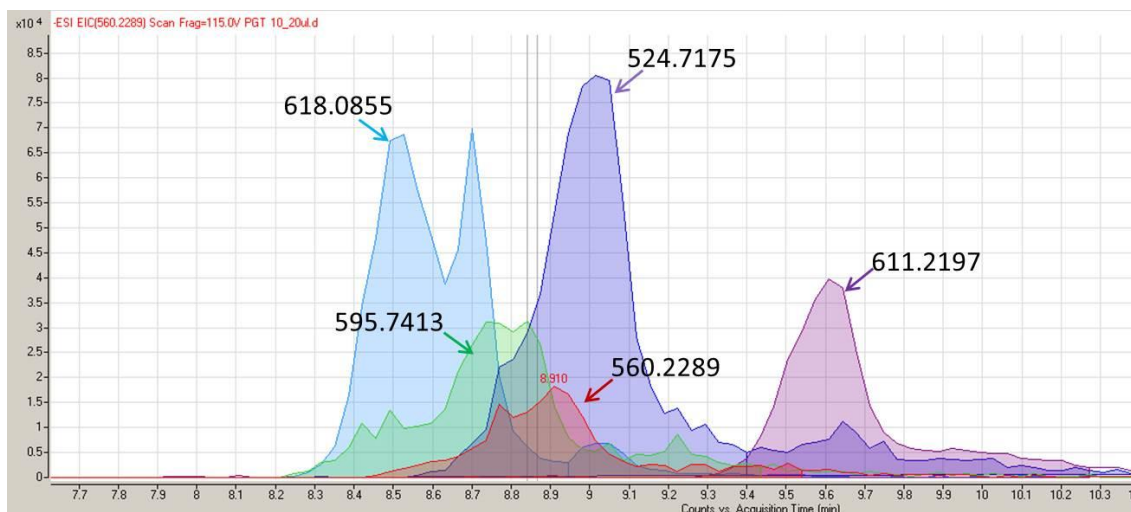


Figure 39: Q-TOF scan of GDP-6dbHep sample to assess purity. Light blue peak = 618.0855 (GDP-6dbHep), other peaks indicate contaminating compounds.

2.3.5.2 LC-MS QQQ

For a more robust verification of GDP-6dbHep purification, MRM was performed on QQQ LC-MS. Fragments scanned for were 78.9594, 158.9263, 177.0760, 344.0407, 424.0057 and 442.0150 m/z. The intact molecule, 618.0855 m/z, was also scanned for. The purified GDP-6dbHep was compared with an extract from *E. coli* expressing the empty pSB1C3 vector, which had undergone the same purification protocol. Figure 40 shows that each of the product ions scanned for were derived from the 618.0855 m/z precursor ion at retention time 8.2 min. A smaller shoulder peak, characteristic of sugar nucleotides, can be seen before each of the peaks apart from 177.0760 m/z. The signal for 177.0760 m/z had a larger amount of background noise than the other peaks which may have hidden the shoulder. For comparison, the same fragments were scanned for in the negative control pSB1C3 extract controls Figure 41.

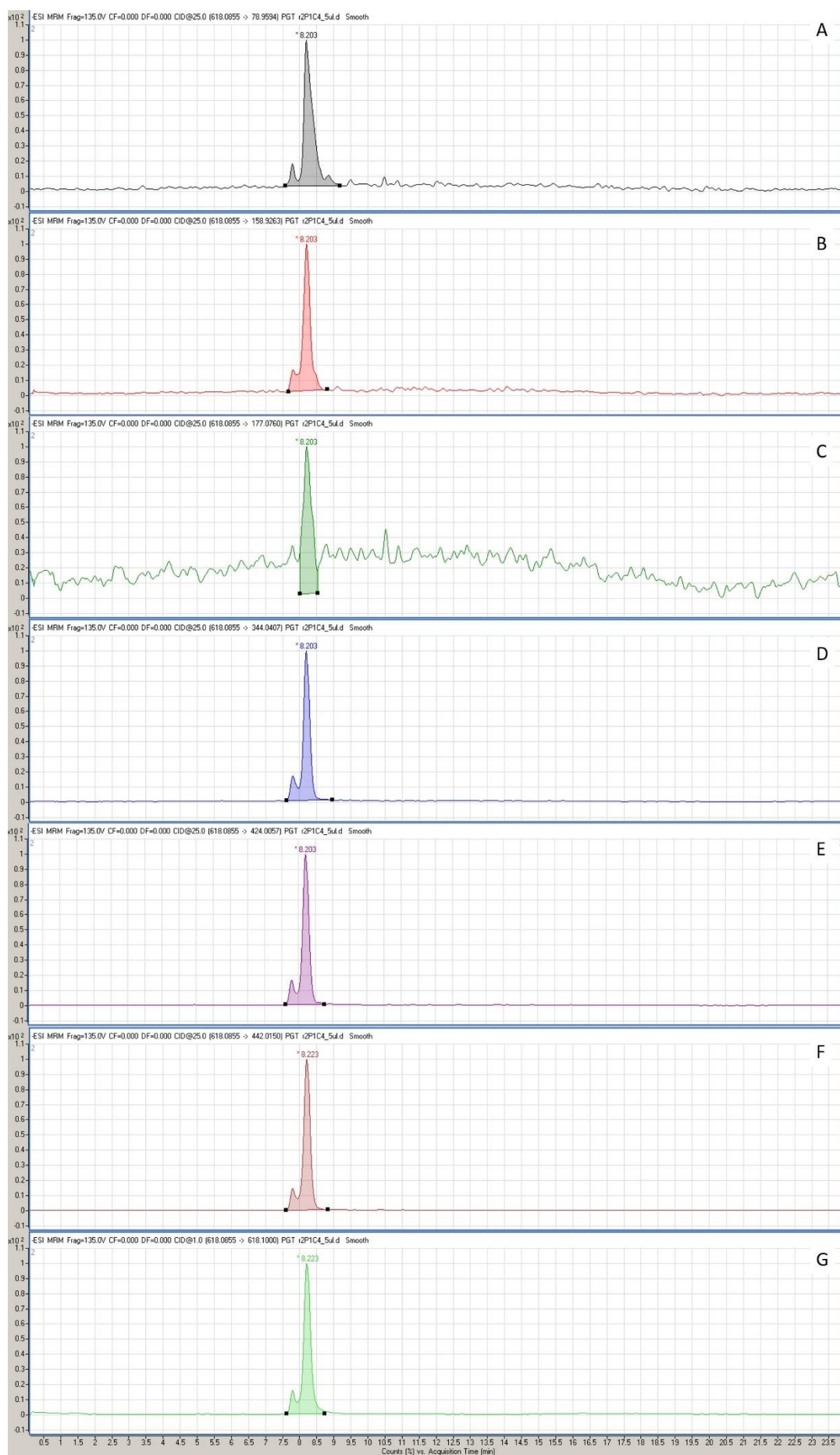


Figure 40: MRM on QQQ LC-MS of purified GDP-6dHep. Scanning for fragments (A) 78.9594; (B) 158.9263; (C) 177.0760; (D) 344.0407; (E) 424.0057; (F) 442.0150 and (G) 618.0855 m/z.

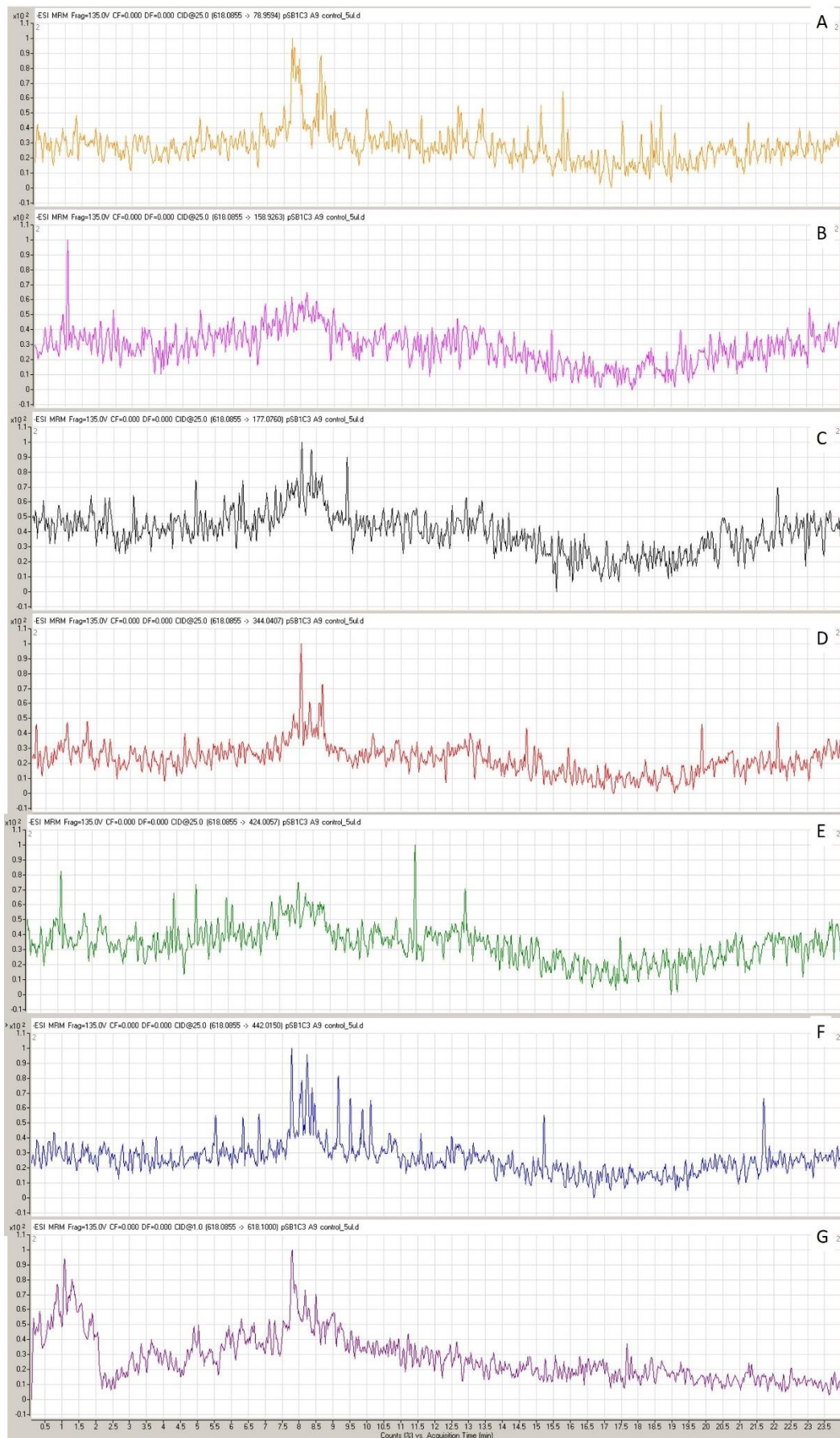


Figure 41: MRM on QQQ LC-MS of purified extract from *E. coli* expressing empty pSB1C3 vector. Scanning for fragments (A) 78.9594; (B) 158.9263; (C) 177.0760; (D) 344.0407; (E) 424.0057; (F) 442.0150 and (G) 618.0855 m/z.

There were no clear peaks present for the selected fragments, indicating no GDP-6dbHep was present in the pSB1C3 empty vector control samples as expected.

Figure 42 shows the spectra of the product ions overlaid with the precursor ion to confirm all the products seen in the GDP-6dbHep were detected at the same retention time and that no corresponding peaks were seen in the negative control sample.

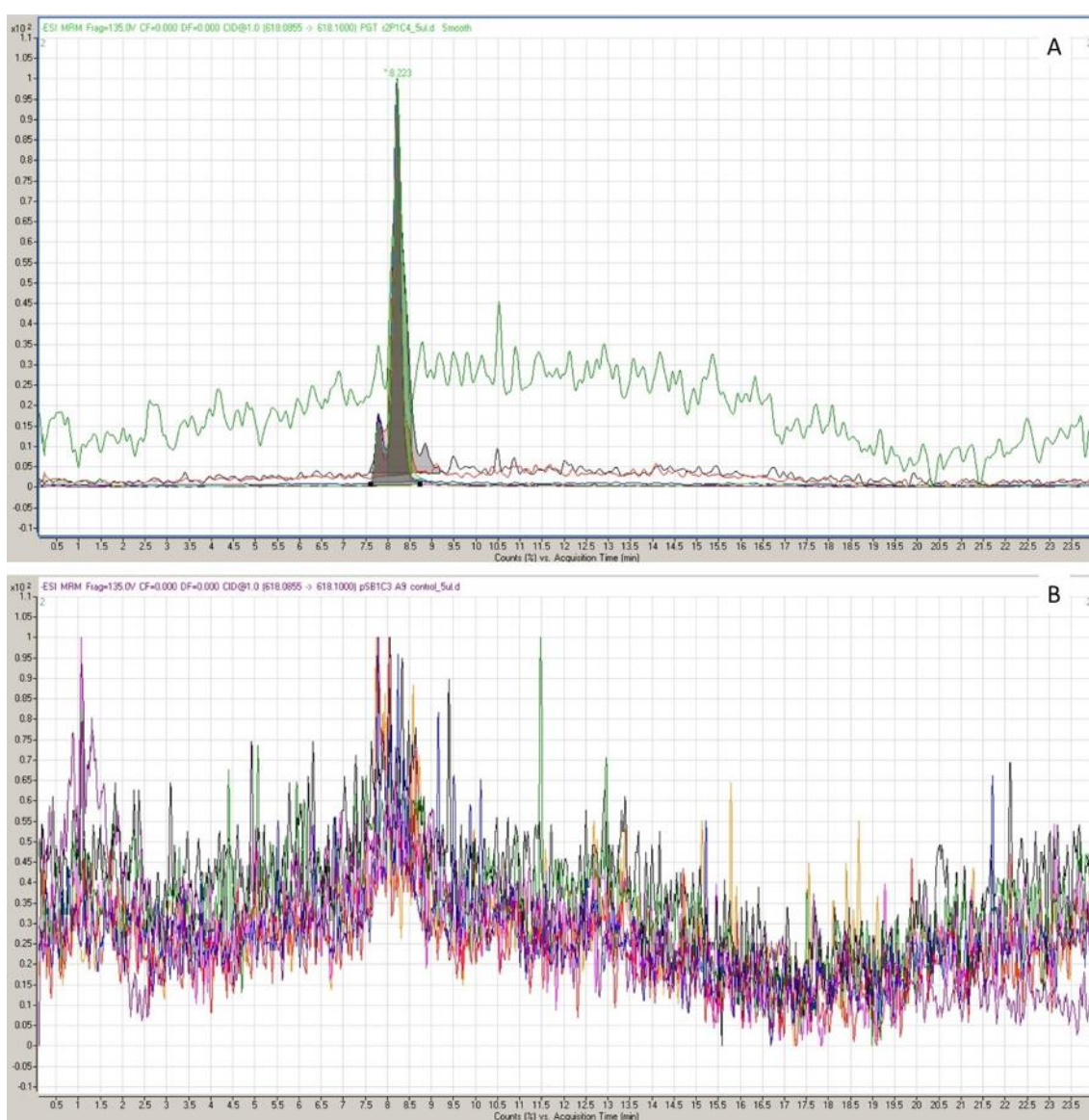


Figure 42: Overlay of MRM spectra for purified GDP-6dbHep and pSB1C3 empty vector. Fragments 78.9594, 158.9263, 177.0760, 344.0407, 424.0057, 442.0150 and 618.0855 m/z were scanned for in (A) GDP-6dbHep and (B) empty pSB1C3 vector metabolite extracts.

An MRM analysis for GDP-mannose, a sugar nucleotide that should naturally be present in *E. coli*, was conducted for the negative control pSB1C3 empty vector extract (Figure 43) Product fragments 163.0622, 442.0150 and 604.0694 m/z were scanned for from precursor ion 604.0694 m/z.

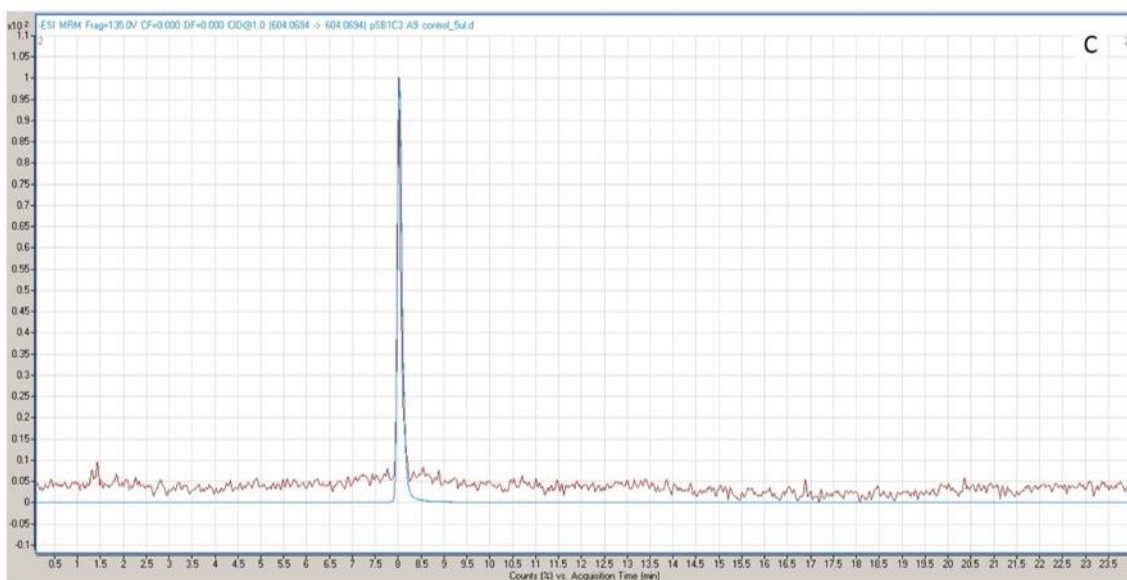


Figure 43: Overlay of MRM for GDP-mannose in pSB1C3 empty vector extract. Fragments scanned for were 604.0694, 442.0150 and 163.0622 m/z from 604.0694 m/z precursor ion.

These data show that natively expressed GDP-mannose was present in *E. coli* transformed with empty pSB1C3 vector whereas GDP-6dHep was absent. This demonstrates the GDP-6dHep operon is being expressed by the *E. coli* transformed with the plasmid containing the six biosynthetic genes.

2.3.5.3 GC-MS

To further verify the GDP-6dHep structure, the samples and standards were tested by GC-MS by Dr. Debbie Salmon. Analysis was conducted by Prof Nick Smirnov. The sample list was kept separate from the GC-MS operators to ensure the test was blind. The TCA hydrolysis was unsuccessful, as abundance in these samples was lower than expected. However, the derivatisation worked well.

Table 6 shows that 6-deoxy-*manno*-heptose was found in trace amounts in the GDP-6dDHep crude and purified samples from expression in *E. coli* and was present in the *in vitro* assay. It was also present in the 6-deoxy-*manno*-heptose standard. It was not found in any other standards or samples as expected.

Table 7: Retention time of samples and standards analysed by GC-MS. Indicating whether 6-deoxy-*manno*-heptose was present, detected in trace amounts or absent (x) and the retention time of the sugar and anomer.

Name	Sample/ Standard	6-deoxy- <i>manno</i> - heptose present?	Retention Time α and β sugar anomers
GDP-6dDHep (<i>E. coli</i> pure)	Sample	Trace	12.59 min & 12.71 min
GDP-6dDHep (<i>in vitro</i>)	Sample	Present	12.59 min & 12.71 min
GDP-4K6Hep	Sample	X	X
GDP-6dDHep (<i>E. coli</i> crude)	Sample	Trace	12.59 min & 12.71 min
GDP-mannose	Standard	X	11.86 min & 12.02 min
Mannose	Standard	X	11.86 min & 12.02 min
Sorbitol	Standard	X	12.26
Mannitol	Standard	X	12.21
6-deoxy- <i>manno</i> - heptose	Standard	Present	12.59 min & 12.71 min
M7P	Standard	X	12.25 min
S7P	Standard	X	12.25 min
M1P	Standard	X	12.25 min
Blank	Standard	X	X

The spectra for GC-MS samples showed a characteristic double peak of the α and β sugar anomers (retention times approximately 12.587 min and 12.726 min) (Figure 44). Fragment 219.1230 m/z is a generic product that is routinely scanned for in sugar analysis. It was found in high abundance in the 6-deoxy-*manno*-heptose standard and GDP-6dDHep produced in the *in vitro* assay (1.0×10^6 and 2.5×10^4 AU (arbitrary units) respectively). In the GDP-6dDHep crude and purified extracts from *E. coli* abundance of the 219.1230 m/z fragment was lower (5.0×10^1 and 4.0×10^2 AU respectively).

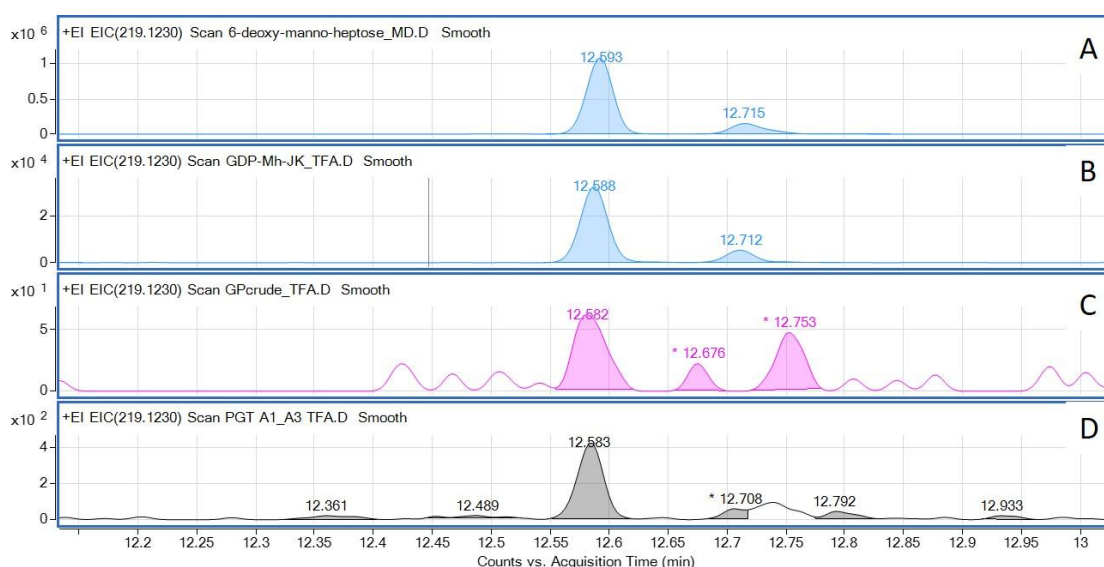


Figure 44: Spectra of 219.1230 fragment scan in GC-MS samples and standards. Samples tested were (A) 6-deoxy-*manno*-heptose standard; (B) GDP-6dHep from *in vitro* assay; (C) GDP-6dHep from crude extract of *E. coli* expressing the biosynthetic plasmid and (D) purified GDP-6dHep from extract of *E. coli* expressing the biosynthetic plasmid as detailed in Table 7.

A second fragment, 333.1738 m/z, was also scanned for. It was more abundant in the 6-deoxy-*manno*-heptose standard than in other sugars in the database (mostly hexoses; data not shown), making it diagnostic for 6-deoxy-*manno*-heptose. The spectra for this fragment showed a double-peak, representing sugar anomers (retention times approximately 12.588 min and 12.710 min) (Figure 45). Abundance of the 333.1738 m/z fragment in the 6-deoxy-*manno*-heptose standard and GDP-6dHep obtained from the *in vitro* assay was high (5.5×10^5 and 2.0×10^4 AU respectively). Abundance of the fragment was lower in the crude and purified GDP-6dHep extracts from *E. coli* (6.0×10^1 and 2.5×10^2 respectively). These data corresponded well with the spectra and relative abundancies for the 219.1230 m/z fragment (Figure 44).

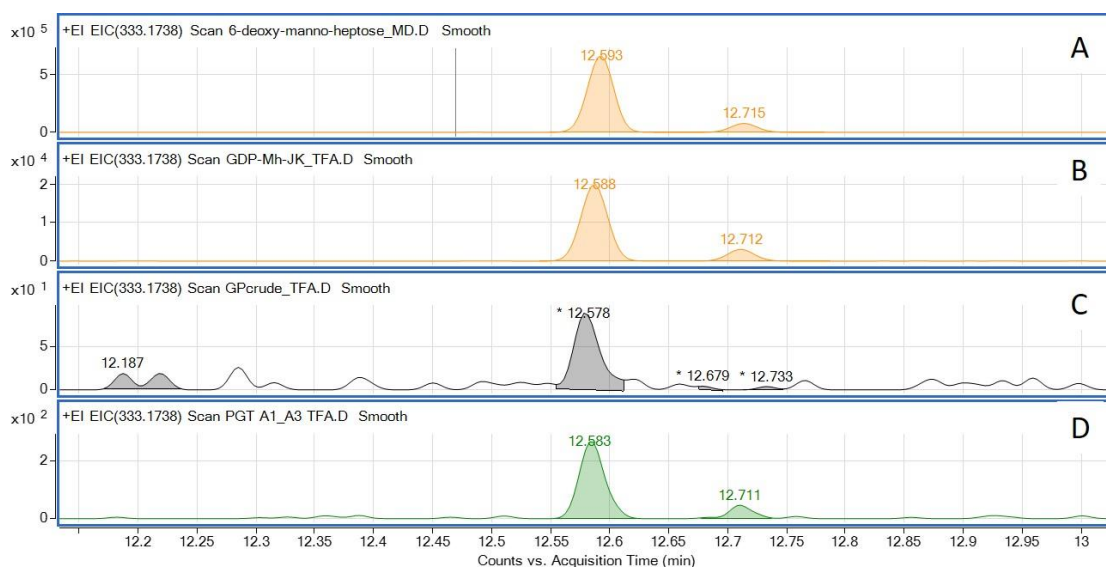


Figure 45: Spectra of 333.1738 fragment scan in GC-MS samples and standards. Samples tested were (A) 6-deoxy-*manno*-heptose standard; (B) GDP-6ddHep from *in vitro* assay; (C) GDP-6ddHep from crude extract of *E. coli* expressing the biosynthetic plasmid and (D) purified GDP-6ddHep from extract of *E. coli* expressing the biosynthetic plasmid as detailed in Table 7.

The extracted ion chromatograms for GDP-6ddHep samples (purified from *E. coli* and synthesised *in vitro*) and for the 6-deoxy-*manno*-heptose standard can be seen in Figure 46. These chromatograms demonstrate the similarities between the test samples and control standards. The GDP-6ddHep purified from *E. coli* and synthesised *in vitro* have match scores of 54 % and 52.9 % respectively when compared with the 6-deoxy-*manno*-heptose standard. These are typical values for good matches of sugars analysed on this GC-MS instrument with sugars from the database (in-house data not shown).

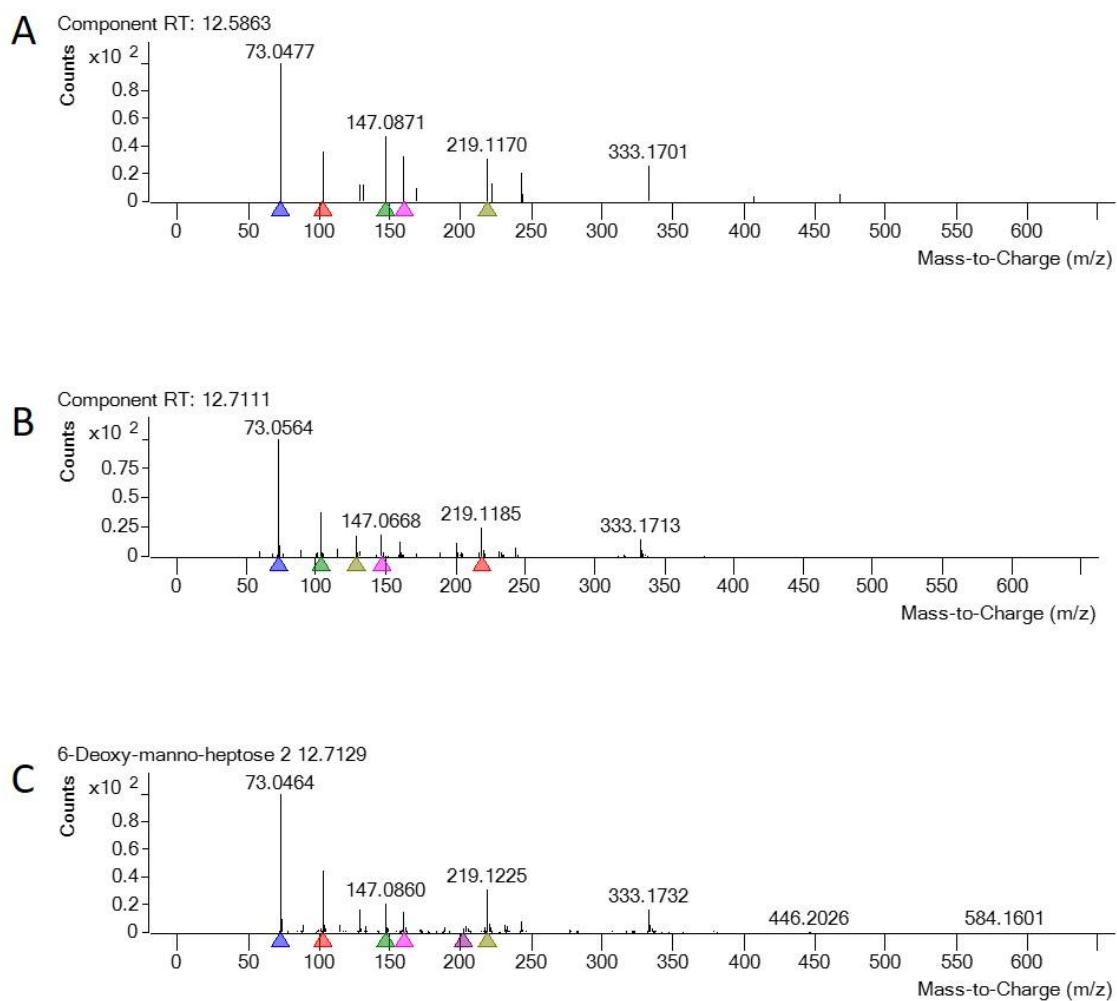


Figure 46: Extracted ion chromatograms for GC-MS of GDP-6dBep. (A) GDP-6dBep purified from *E. coli* transformed with the biosynthetic pathway and (B) GDP-6dBep synthesised *in vitro*, matched with the extracted ions from (C) 6-deoxy-manno-heptose standard.

Figure 47 demonstrates the fragments seen in the extracted ion spectra (Figure 46) have the same retention time as the component 6-deoxy-manno-heptose. The spectra for GDP-6dBep extracted from *E. coli* expressing the biosynthetic pathway is less clear as it is in smaller abundance, therefore gives lower signal, which is more error-prone (Figure 47B).

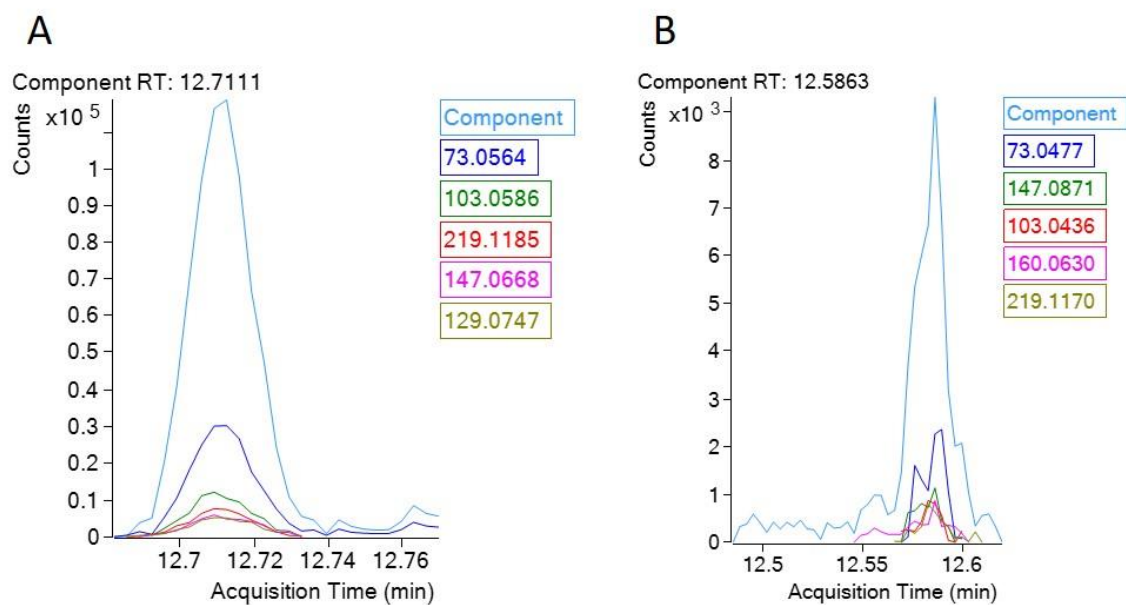


Figure 47: Overlays of fragments scanned for in GC-MS of GDP-6dDHep. Fragments were 73.0564, 103.0586, 219.1185, 147.0668 and 129.0747 m/z and the 'component' 6-deoxy-mannoheptose for (A) GDP-6dDHep synthesised *in vitro* and (B) GDP-6dDHep extracted from *E. coli* expressing the GDP-6dDHep plasmid.

2.3.5.4 NMR

^1H and ^{31}P NMR analyses were carried out by Dr. Martin Rejzek at the John Innes Centre on a purified sample of GDP-6dDHep. Structures of the sample and control compounds for comparison are shown in Figure 48.

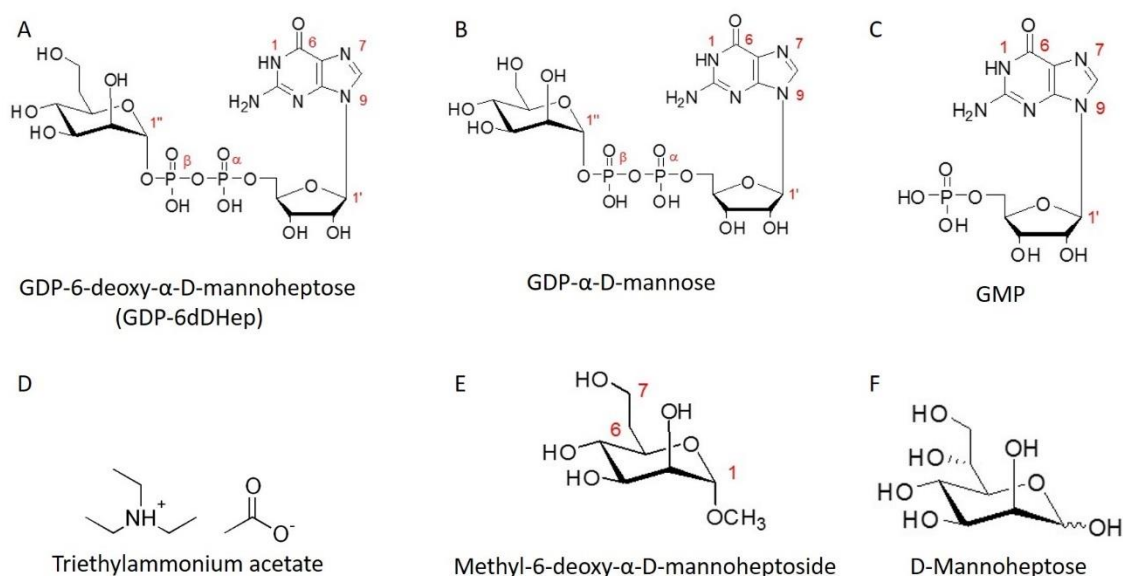


Figure 48: structures of GDP-6dDHeP and five standards analysed by ^1H and ^{31}P NMR. (A) GDP-6-deoxy- α -D-manno-heptose (GDP-6dDHeP); (B) GDP- α -D-mannose; (C) guanosine monophosphate (GMP); (D) Triethylammonium acetate; (E) methyl-6-deoxy- α -D-manno-heptoside; (F) D-manno-heptose.

The purified GDP-6dDHeP sample was compared with four standards (GDP- α -D-mannose (GDP-mannose), guanosine monophosphate (GMP), methyl-6-deoxy- α -D-manno-heptoside and D-manno-heptose) to elucidate its structure (Figure 48). The TEAA control was used to assess for contamination of the sample, as TEAA was the main component in the buffer used to purify GDP-6dDHeP by HPLC.

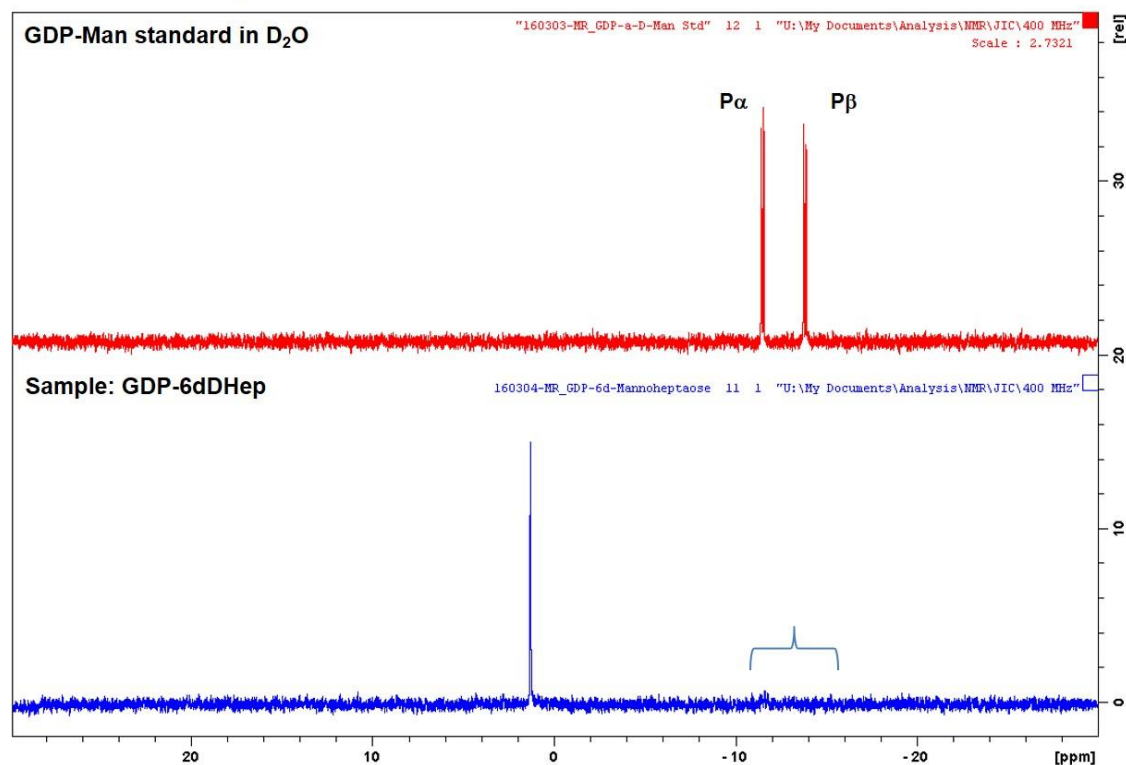
³¹P NMR: 400 MHz, D₂O

Figure 49: Spectra for ³¹P NMR of GDP-mannose standard (red) and GDP-6dHep (blue). Parenthesis indicates absence of peaks diagnostic for the two phosphate groups (P_α and P_β).

³¹P NMR compared the GDP-6dHep sample with a GDP-mannose standard. Two peaks (P_α and P_β) are clearly identifiable in the GDP-mannose control, representative of the phosphate groups in GDP that link to the ribose sugar (Figure 48; Figure 49). These are absent from the GDP-6dHep sample. There is an additional peak in the sample that is absent from the control that may represent a contaminating inorganic phosphate or sugar-1-phosphate (Figure 49).

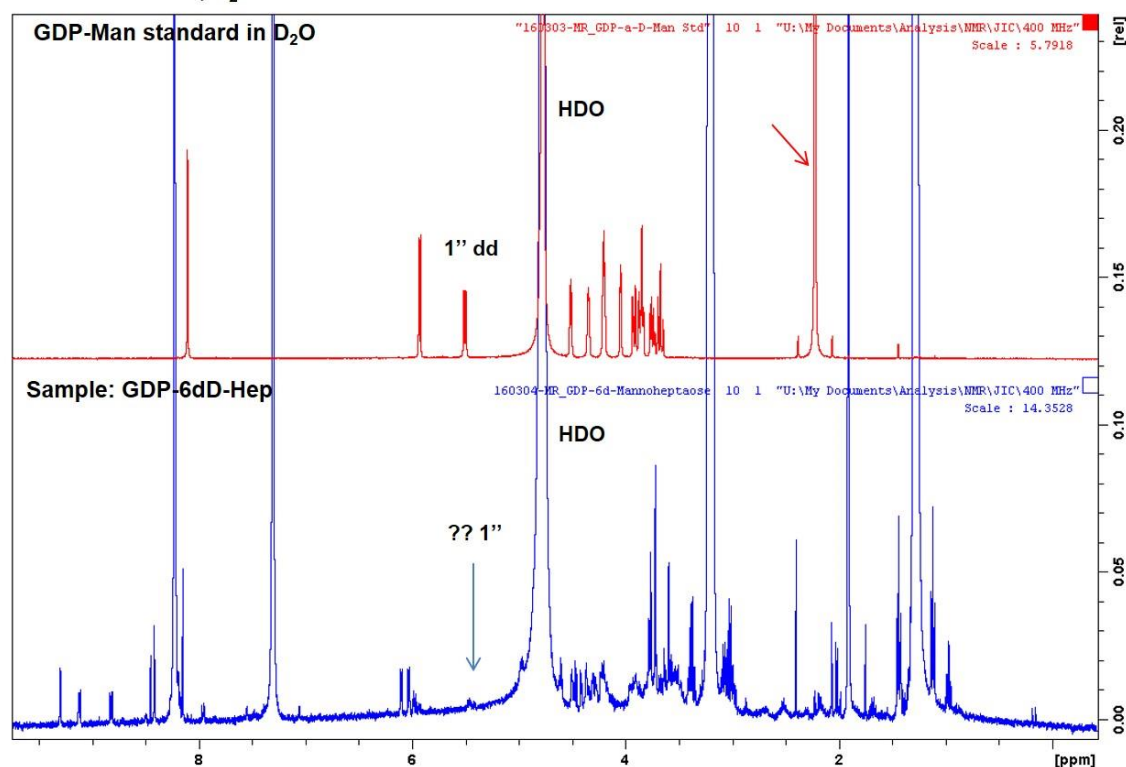
^1H NMR: 400 MHz, D_2O 

Figure 50: Spectra for ^1H NMR of the GDP-mannose standard (red) and GDP-6dD-Hep (blue). The doublet of doublets signal for carbon 1 is seen in the GDP-mannose standard as expected (1'' dd). A potential signal is indicated in the GDP-6dD-Hep spectra (?? 1''; blue arrow). Red arrow denotes peak for the internal acetone standard. HDO = heavy water solvent.

Spectra for ^1H NMR of the GDP-6dD-Hep sample and GDP-mannose standard did not match as expected. There were several unrecognised peaks in the sample suggesting contamination of the GDP-6dD-Hep sample. Furthermore, the identifying peak for doublet of doublet (dd) signal at carbon 1 was negligible, indicating abundance of GDP-6dD-Hep in the sample was extremely low (Figure 48; Figure 50).

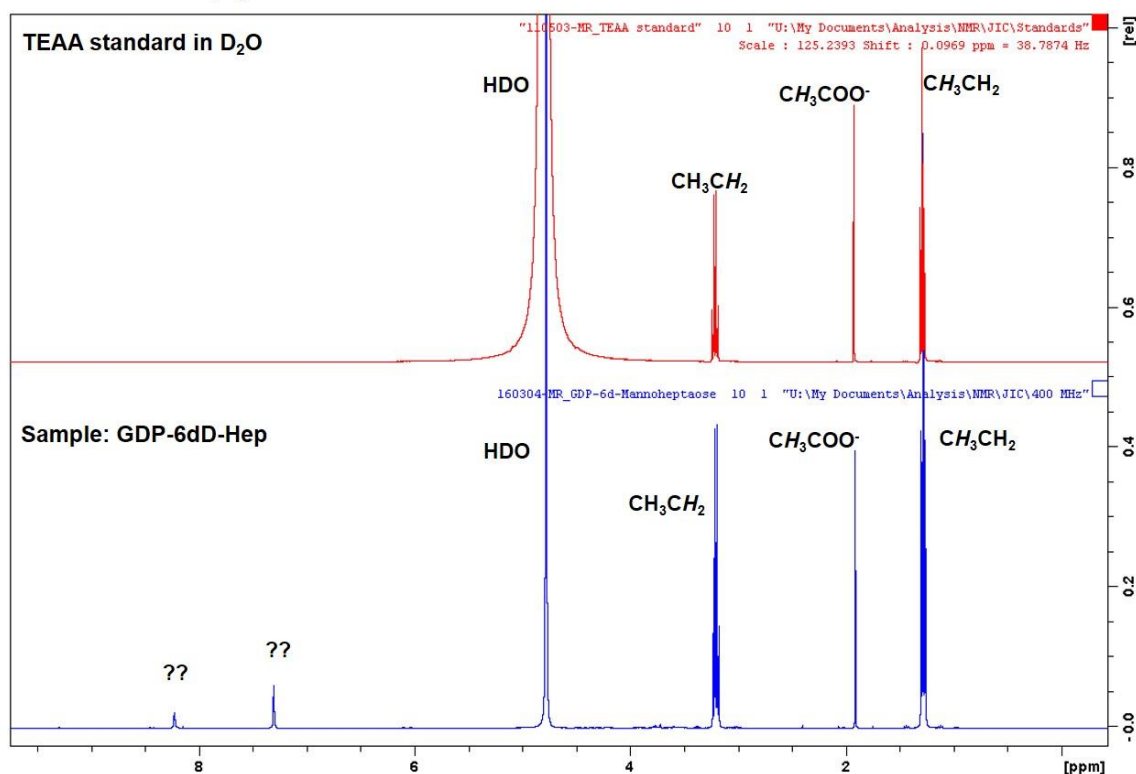
^1H NMR: 400 MHz, D_2O 

Figure 51: Spectra for ^1H NMR of TEAA in D_2O (red) and GDP-6dD-Hep (blue). Peaks for different hydrogen atoms (italicised) in carbohydrates (CH_3CH_2 ; CH_3COO^- and CH_3CH_2) are indicated. There are two peaks of unknown origin in the GDP-6dD-Hep sample (??). HDO = heavy water solvent.

Corresponding peaks for CH_3CH_2 and CH_3COO^- are seen in the ^1H NMR spectra comparing the GDP-6dD-Hep sample with TEAA spiked D_2O . This indicates the sample is heavily contaminated with TEAA from the HPLC purification step. There are two small, unknown peaks (labelled "??") in the GDP-6dD-Hep sample but they are not diagnostic of a sugar-nucleotide (Figure 48; Figure 51).

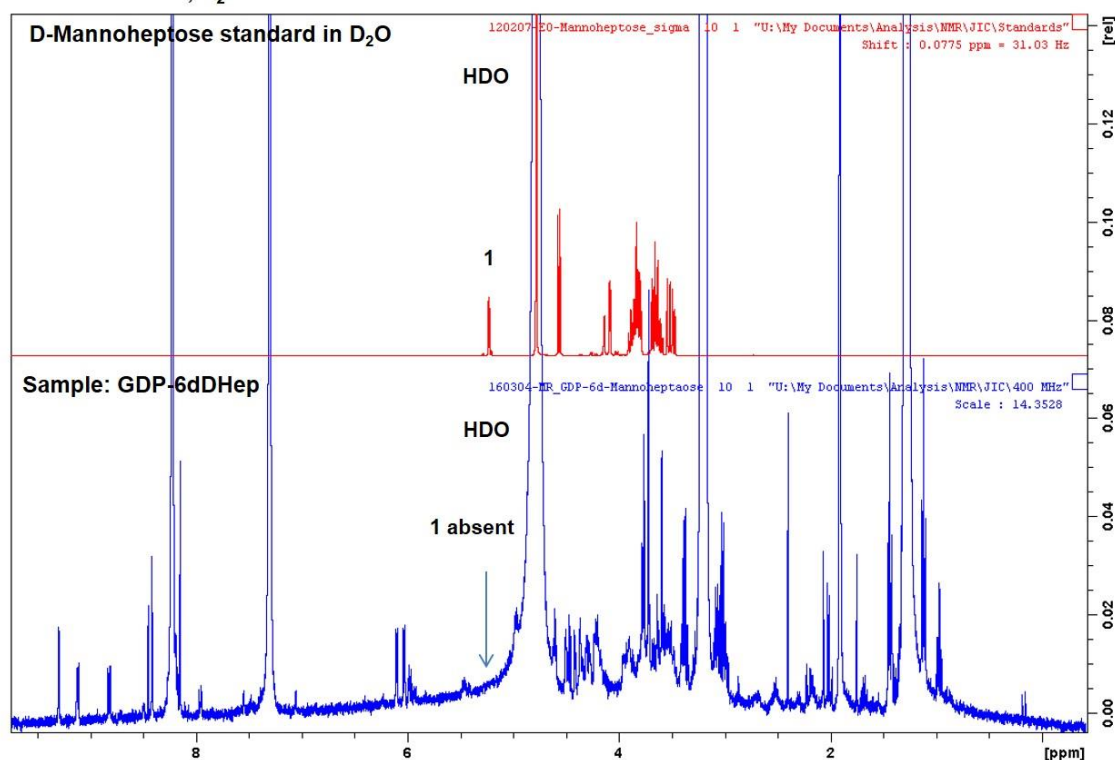
^1H NMR: 400 MHz, D_2O 

Figure 52: Spectra for ^1H NMR of *D-manno*-heptose standard (red) and GDP-6dHep (blue). The signal for carbon 1 is present in the *D-manno*-heptose standard as expected, but absent from the GDP-6dHep sample. HDO = heavy water solvent.

The spectra for ^1H NMR comparing *D-manno*-heptose with GDP-6dHep may indicate presence of a *manno*-heptose derivative in the sample, though with the amount of noise in the GDP-6dHep spectrum it is impossible to decipher. However, the peak for anomeric carbon 1 indicative of a reducing sugar is not present in the GDP-6dHep sample (Figure 48; Figure 52).

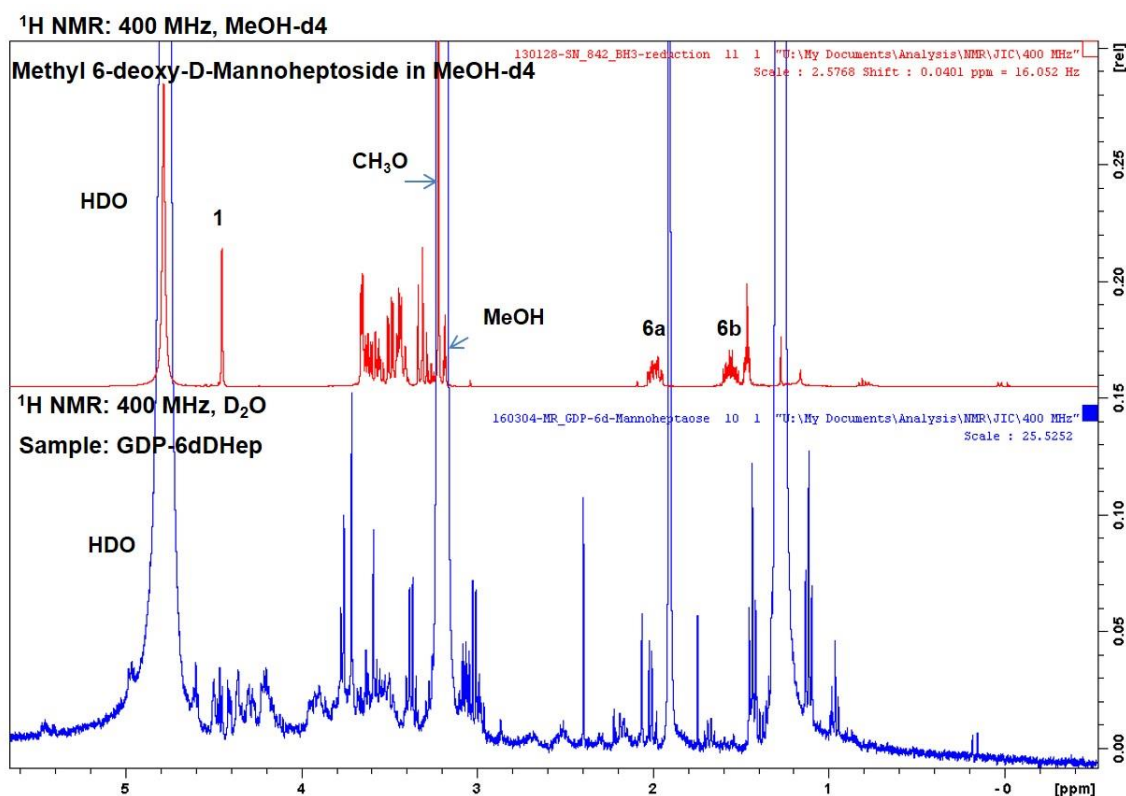


Figure 53: Spectra for ¹H NMR of methyl-6-deoxy-D-manno-heptoside standard (red) and GDP-6dHep (blue). Peaks for carbon 1 and carbon 6 are present in the methyl-6-deoxy-D-manno-heptoside standard as expected, but absent from GDP-6dHep. HDO = heavy water solvent; MeOH = methanol; CH₃O = methoxide.

Spectra for ¹H NMR comparing methyl-6-deoxy-D-manno-heptoside with GDP-6dHep show the peak for carbon 1 and two peaks for carbon 6 are present in the standard but absent from the GDP-6dHep sample (Figure 48; Figure 53). As with Figure 53, noise from contamination may be masking the presence of a low abundance of GDP-6dHep but it is impossible to decipher.

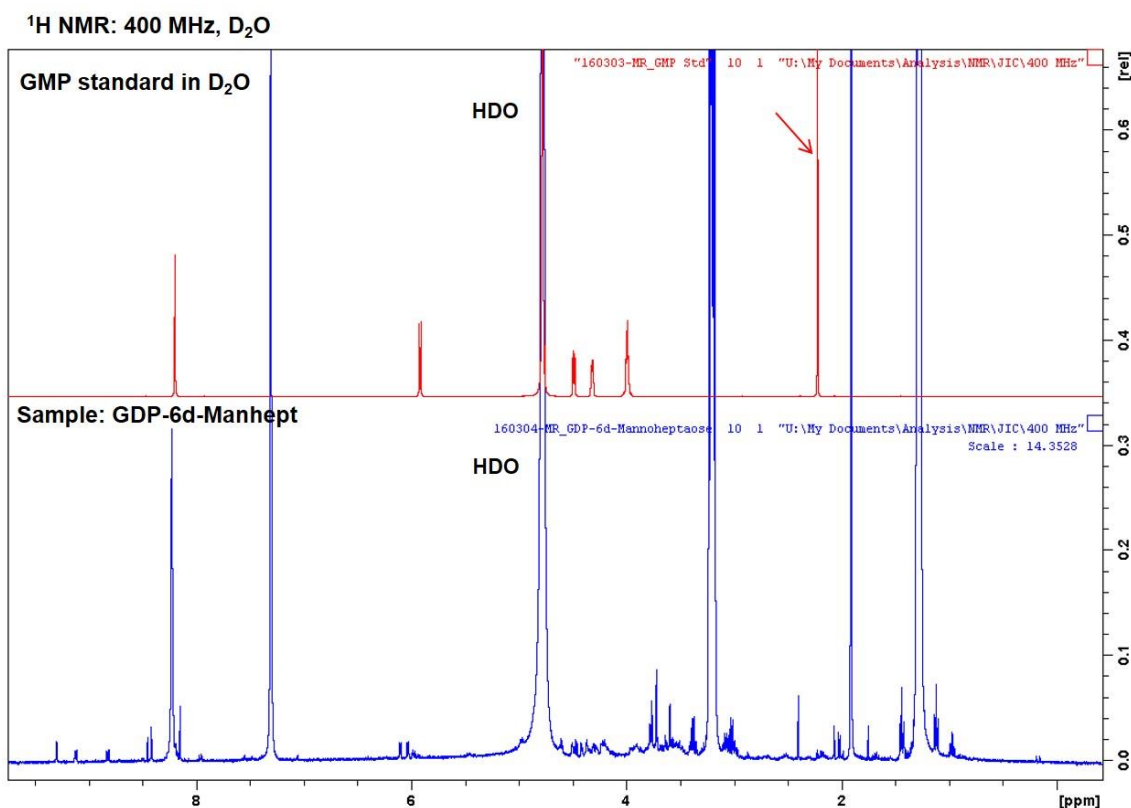


Figure 54: Spectra for ¹H NMR of *D-manno*-heptose standard (red) and GDP-6dHep (blue) indicating the expected peak for carbon 1. Red arrow denotes internal acetone standard. HDO = heavy water solvent.

Spectra for ¹H NMR of the GDP-6dHep sample and GMP standard did not match, indicating absence of GMP from the sample (Figure 48; Figure 53).

Overall, the NMR spectra for GDP-6dHep did not correspond with those of the standard controls. This may be due to a combination of low abundance of the GDP-6dHep (which was clearly identified by QQQ LC-MS Figure 48 and GC-MS Figure 48) and large amounts of TEAA contamination distorting the spectra. There is insufficient data to characterise the structure of GDP-6dHep by NMR in this study.

2.4 Discussion

The genes for expression of CPS in *B. pseudomallei* are found on a single coding locus on chromosome 1 (Sim *et al.*, 2008; Cuccui *et al.*, 2012). Previous attempts to clone and express the entire operon in *E. coli* were unsuccessful (data not shown). In this project, only the essential genes were selected from the coding locus, with the intention of exploiting native mechanisms in *E. coli* to complete CPS expression (see Chapter 5). The first set of essential genes identified were the six biosynthetic genes for production of the activated nucleotide sugar precursor, GDP-6d_DHep. A plasmid containing the operon was constructed and successfully expressed in *E. coli* (section 2.3.4). However, levels of GDP-6d_DHep production were extremely low and although the correct structure was verified by LC-MS and GC-MS, there were insufficient amounts purified for NMR analysis

2.4.1 Construction of GDP-6d_DHep biosynthetic operon

To construct the biosynthetic operon, the BioBricks™ 3A assembly method was chosen (Shetty *et al.*, 2008). It is a modular approach with standardised components that would be beneficial to future development of a platform for bespoke polysaccharide expression. Furthermore, it eliminates the requirement of a PCR step which is time-consuming and can introduce point mutations. For the BioBricks™ 3A assembly method, each gene was synthesised with four conserved restriction sites; *EcoRI* and *XbaI* flanking upstream and *SpeI* and *PstI* flanking downstream. To prevent incorrect digestion, each gene had to be checked for internal restriction sites. Whilst none of the six genes contained any of the four sites, the cloning vectors each had an additional *EcoRI* site and the *wcbL* pEX-K plasmid contained an *XbaI* site (Figure 20). Whilst careful planning enabled the cloning strategy to continue without error, these are four common

restriction sites and avoiding them in either the synthetic gene or cloning vector could be challenging for larger operons or high-throughput testing of different components.

In addition to the restriction sites, the genes were synthesised with a pBAT4 RBS and c-Myc tag for protein purification (Figure 18). pBAT4 is derived from the Shine-Dalgarno sequence of the RBS for the T7 bacteriophage coat protein g10. It is a strong RBS and has previously been shown to enhance non-native gene transcription in *E. coli* and cell-free operons (Olins, 1988; Olins & Rangwala, 1989; Karig *et al.*, 2012). However, to precede all the genes with the same RBS presumes they all require the same rate of transcription. To investigate, the RBS sequences should be identified in the native *B. pseudomallei* operon and then analysed using several computational tools to determine their relative strengths. (Reeve *et al.* (2014) analysed the RBS calculator (<https://salislab.net/software/>), UTR designer (https://sbi.postech.ac.kr/utr_designer/) and RBS designer (http://ssbio.cau.ac.kr/web/?page_id=195) and found they were all excellent at predicting and characterising upstream regulatory sequences.) This could inform design of the operon to include, if not native then native-strength, RBS sequences for each individual gene. In the interest of standardisation and modularity, each gene was given a c-Myc protein tag. This was unfortunate planning as GmhA and WcbN have molecular weights of 20.8 and 20.4 kDa respectively and WcbL and WcbJ have molecular weights of 38.0 and 38.1 kDa respectively. Due to their closeness in size, they could not be separated by SDS-PAGE and their expression verified by Western blot. This is an excellent example of the importance of careful planning in synthetic biology.

The synthetic genes were also codon-optimised for expression in *E. coli*. A recent large-scale study of 6,348 genes expressed in *E. coli* found codon optimisation

increased production of mRNA and protein compared to native codon usage (Boël *et al.*, 2016). Whilst the genomic island containing the CPS coding locus does not have high GC-content like typical of the *B. pseudomallei* genome, codon optimisation may still be beneficial for optimising non-native operon expression. Once synthesised, the operon was assembled sequentially using restriction digestion and ligation. The method worked extremely well, and a third of transformed colonies contained the correct new plasmid (Figure 22). When the project began in 2012 the cost of having the entire 5,257 bp operon was prohibitively expensive. Current technologies allow operons to be synthesised for less than 5p per base. This demonstrates how quickly the field of synthetic biology is progressing. However, the sequential assembly of an operon enables each new combination of components to be analysed and this can be beneficial for troubleshooting.

2.4.1.1 The problem of WcbM (part 1)

Butty *et al.* (2009) synthesised GDP-6dBep *in vitro* using orthologous enzymes purified from *C. jejuni* and *Y. pseudotuberculosis*. They separated the pathway into two reactions, the first performed by GmhA, WcbL and WcbN orthologues, the second by WcbM, WcbJ and WcbK orthologues. No explicit reason for this was given, so it was hypothesised that the WcbM reaction (a guanosyltransferase; Figure 10) was in some way problematic or inhibitory and this may materialise as cytotoxicity in *E. coli*. Rather than assembling the pathway as it is encoded in *B. pseudomallei* (*wcbJ>wcbK>wcbL>gmhA>wcbM>wcbN*; Figure 9) the second construct made was *wcbL-gmhA-wcbN* and then *wcbM* was added to create *wcbM-wcbL-gmhA-wcbN* (Figure 22). The reason for this was to test for cytotoxic effects of WcbM before constructing the remainder of the

operon. There were no cytotoxic effects observed, however if there had been this approach would have enabled the problem to be addressed more easily than if the operon had been synthesised in its entirety. Unfortunately, this resulted in the non-native ordering of the genes expressed on the operon: *wcbJ>wcbK>wcbM>wcbL>gmhA>wcbN*.

2.4.1.2 Operon optimisation to improve GDP-6dbHep yield

The re-ordering of genes on the operon to investigate potential WcbM cytotoxicity could have caused the low GDP-6dbHep expression. There is evidence to suggest that co-linearity (genes on a locus ordered the same as enzymes in a pathway) is important for optimal expression (Nishizaki *et al.*, 2006; Kovács *et al.*, 2009; Lim *et al.*, 2011). The distance from the start of the operon and the overall length of the operon can affect the strength of gene expression. It is hypothesised that time spent transcribing mRNA lasts longer for genes at the start of the operon and that time also increases with the length of the operon (increases 40 % per 1,000 bp) (Lim *et al.*, 2011). Therefore, by altering the natural order of the genes on the operon, the rates of expression for each gene changed and resulted in low production of GDP-6dbHep. In the synthetic operon *wcbM* is the third gene in the sequence, whereas on the natural operon it is fifth. Therefore, it is likely to be more strongly expressed in the synthetic operon. Similarly, *wcbL* is moved from third in the natural operon to fourth in the synthetic operon meaning it is likely to be more weakly expressed. In Chapter 3, there is evidence that WcbM activity is inhibitory to the WcbL reaction and this could be the cause of the low GDP-6dbHep expression *in vivo*.

To optimise the synthetic operon, it would be wise to first analyse the natural *B. pseudomallei* operon. Rates of transcription and translation along the operon

could be determined using several methods; transcriptomics such as RNA sequencing to quantify level of transcription (Conway *et al.*, 2014; Smanski *et al.*, 2015), functional assays to determine levels of protein expression (Smanski *et al.*, 2014) or by inserting a fluorescent protein at each step of the pathway and measuring its relative fluorescence to determine transcription rates (Lim *et al.*, 2011). There are usually many iterations of an operon to be tested before an optimal combination is found. For example, van Summeren-Wesenhagen *et al.* (2015) tested 81 operon variants simply by altering the spacer distance between the Shine-Dalgarno sequence and start codon to optimise production of *p*-coumaryl alcohol in *E. coli*. They increased yield from 22 mg L⁻¹ to between 48 and 52 mg L⁻¹.

A strategy similar to the approach taken by this project was explored by Smanski *et al.* (2014). They chose 16 essential genes for nitrogen fixation from a possible 20 genes spread over 7 operons in *Klebsiella oxytoca* to create a single optimised synthetic operon. By varying regulatory sequences, gene order and gene orientation they analysed a total of 506 different combinations to search for the most effective operon. They found that the optimal synthetic operon only restored 57 % functionality in mutant *K. oxytoca*, suggesting that millennia of natural evolution may provide the best solution for gene organisation and regulation.

2.4.1.3 Alternative reasons for low GDP-6_DHep yield

Whilst sub-optimal operon expression is the most likely reason for low GDP-6_DHep yield, there are alternative explanations that cannot be ignored. It can be hypothesised that, due to the large amount of capsule produced and expressed across the entirety of the cell surface, a large amount of precursor material is required (Figure 8). Therefore, strong RBSs, a strong promoter (BBa_J23118)

and a high copy number vector (pSB1C3) were used to express the operon. However, GDP-6dbHep requires the diversion of S7P from the pentose phosphate pathway and there could be significant implications for the host cell metabolism (Figure 10). The fact that *E. coli* expressing the biosynthetic plasmid had a significantly extended lag time suggests there was a detrimental effect on the host cell metabolism (Figure 35). To investigate, the whole metabolome was analysed by mass spectrometry and although it confirmed there were differences between the metabolomes of transformed and wild-type *E. coli* this could not be attributed to specific metabolites (section 2.3.3).

A second reason for poor GDP-6dbHep yield could be lack of glycosyltransferase activity. If high levels of GDP-6dbHep accumulate it could become cytotoxic due to lack of glycosyltransferases to assemble the CPS polymer.

2.4.2 GDP-6dbHep purification

GDP-6dbHep was purified using an adapted method by Rabinä *et al.* (2009). The nucleotide sugar was detected and verified by LC-MS and GC-MS. Unfortunately, there were four contaminants remaining in the sample which purified with GDP-6dbHep as they are probably metabolites with similar molecular weights, structures and properties (e.g. charge) (Figure 39). Furthermore, the sample was also contaminated with TEAA from the HPLC buffer which could not be removed despite repeated vacuum centrifugations (Figure 51). Though quantities of GDP-6dbHep were undoubtedly low, the inability to detect it by NMR was hampered by high background noise (Figure 52; Figure 53) probably caused by the four contaminants. The only way to solve these issues is to improve the expression levels in *E. coli* by optimising the operon and running repeat extractions and pooling the samples to further concentrate GDP-6dbHep.

Chapter 3

Synthesis of GDP-6-deoxy- α -D-manno-heptose in a microfluidics system

3.1 Introduction

All the genes for CPS expression in *B. pseudomallei* are clustered on the same locus (Cuccui *et al.*, 2012). This project proposes to clone only the essential genes for CPS expression into *E. coli*. Chapter 2 described the successful biosynthesis of nucleotide sugar precursor (GDP-6dDHep) in *E. coli* expressing *gmhA*, *wcbL*, *wcbN*, *wcbM*, *wcbK* and *wcbJ*. Whilst data showed the correct structure was being produced, yields were extremely low as the operon had not been optimised. The purified GDP-6dDHep also contained some contaminants that could not be removed easily. An *in vitro* assay could solve these problems and produce sufficient GDP-6dDHep for contribution to glycosyltransferase characterisation studies.

Following nucleotide sugar precursor biosynthesis, the next step in CPS production is assembly of the polymer by three putative glycosyltransferases, *wcbB*, *wcbE* and *wcbH*. None of these enzymes have been purified or characterised previously. As glycosyltransferase purification is challenging, a method for low-volume assays is desirable to minimise the quantity of enzyme required. (See Chapter 5 for discussion about glycosyltransferase purification). The development of microfluidics has drastically reduced the volume of liquids required for several applications (Baxani *et al.*, 2016). A system designed by Leptihn *et al.* (2013) suspends small volumes of aqueous solution within a lipid layer. The droplets are formed by introducing an oil into a flowing aqueous solution. The flow rate of the oil and aqueous solution determines droplet properties such as size, shape and lipid envelope thickness. Droplet volumes can be as low as 1 μL , greatly reducing the amount of enzyme required to carry out reactions. (For example, in Chapter 4, the smallest reaction volume used is 15 μL .) Two aqueous solutions, containing enzymes and substrates separately

for example, can be mixed together before the oil is introduced. As the oil is introduced into a flowing solution, the conditions within the resultant droplet are turbulent; ideal for mixing substrates and enzymes. The lipid envelope also provides a surface at which membrane-associated enzymes can locate, perhaps increasing their stability. (This may be particularly important for the *B. pseudomallei* glycosyltransferases which have the GT-B fold often associated with proximity to membrane surfaces (Breton *et al.*, 2006; Albesa-Jové *et al.*, 2014)) Finally, by joining two or more droplets via a transporter protein, reactions can be compartmentalised. This could be useful when characterising pathways or for reactions with a toxic or inhibitory substrate or product.

In this chapter, the biosynthetic pathway for GDP-6dDHep is performed *in vitro* in the microfluidics system. Whilst orthologous enzymes from *C. jejuni* and *Y. pseudotuberculosis* were used previously (Butty *et al.*, 2009), this is the first time all six *B. pseudomallei* enzymes have been tested together.

3.1.1 Aims

The aim of this chapter was to use a microfluidics system to synthesise GDP-6dDHep *in vitro* using the six biosynthetic enzymes from *B. pseudomallei*. The purpose of this aim was to identify a low-volume method for conducting assays that in the future will reduce the quantity of precious glycosyltransferase required for characterisation studies. It also provides an alternative method for the synthesis of GDP-6dDHep in a cell-free system, reducing contamination and improving yield.

3.2 Materials and methods

All reagents were purchased from Sigma-Aldrich unless otherwise stated.

3.2.1 Large scale protein purification

Plasmids for the purification of His-tagged GmhA, WcbL, WcbN, WcbM, WcbJ, WcbK and WcbI were kindly provided by Dr. Nicholas Harmer, University of Exeter (Harmer, 2010; Vivoli *et al.*, 2015). The genes from *B. pseudomallei* strain K96243 were cloned into pNIC28-Bsa4, using methods described by Savitsky *et al.* (2010). These plasmids were transformed into *E. coli* as previously described (section 2.2.3.2).

3.2.1.1 Isopropyl β -D-1-thiogalactopyranoside induction

2.5 mL aliquots from overnight cultures of *E. coli* were used to inoculate 500 mL LB plus 100 $\mu\text{g mL}^{-1}$ of appropriate antibiotic. The LB cultures were incubated at 37 °C plus aeration until OD₆₀₀ 0.4 – 0.6 was achieved, then induced with 200 μM isopropyl β -D-1-thiogalactopyranoside (IPTG). They were then incubated at 20 °C plus aeration for 24 hr.

3.2.1.2 ZYM-5052 auto-induction

2.5 mL aliquots from overnight cultures of *E. coli* were used to inoculate 500 mL ZYM-5052 media (Studier, 2005). Cultures were grown at 37 °C plus aeration for two hours then moved to 20 °C for 24 hr.

3.2.1.3 Harvesting cells

Cells were harvested by centrifugation (24,000 $\times g$, 30 min, 4 °C). The supernatant was discarded and the pellet resuspended in 30 mL buffer (20 mM Tris-HCl pH

8.0, 100 mM NaCl, unless otherwise indicated). The resuspended pellet could be stored at -20 °C or used immediately.

3.2.1.4 Cell lysis

Protease inhibitors (either 1 mg mL⁻¹ phenylmethylsulfonyl fluoride (PMSF), 1 mg mL⁻¹ benzamidine or two Pierce™ Protease Inhibitor tablets (Thermo Fisher)) were added to the resuspended pellet, which was incubated at room temperature for 20 min with gentle rocking. The resuspended pellet was then put on ice for cell lysis by sonication (six cycles of 20 s at amplitude 15 µm, 20 s rest). The lysate was centrifuged (24, 000 xg, 30 min, 4 °C), the supernatant was collected and re-centrifuged (24, 000 xg, 10 min, 4 °C). Finally, the supernatant was removed and filtered at 20 µm.

3.2.2 Purification of proteins for *B. pseudomallei* CPS biosynthesis

All columns, column matrices and systems used in the large-scale purification of proteins were from GE Healthcare and other reagents from Sigma Aldrich unless otherwise indicated.

3.2.2.1 Nickel affinity and size exclusion chromatography

After extracting proteins WcbJ, WcbK, WcbL, WcbM, WcbN and WcbI from large scale cultures, a 20 µL aliquot was kept for SDS-PAGE and the remaining volume purified on an automated ÄKTExpress system.

Nickel affinity chromatography was performed using a 1 mL His-Trap FF column, equilibrated and washed with buffer (20 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole, pH 8.0) and eluted in 20 mM Tris-HCl, 500 mM NaCl, 250 mM imidazole, pH 8.0. The eluate was loaded onto a Superdex 200 16/600

size exclusion chromatography column and eluted isocratically in 10 mM HEPES pH 7.0, 500 mM NaCl. For WcbK, the size exclusion buffer was 10 mM HEPES pH 7.0. The eluate was collected in 2 mL fractions.

3.2.2.2 Gravity flow nickel affinity chromatography

Following extraction from large scale culture, GmhA was purified using a gravity flow nickel affinity chromatography only.

An empty column was filled with 2 mL Ni Sepharose High Performance bead slurry (GE Healthcare) and equilibrated with buffer A (20 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole pH 8.0). The clarified culture supernatant was applied to the column, the flow through collected and reapplied to the column. The column was washed with 10 mL buffer A, then the protein was eluted into 10 mL buffer B (20 mM Tris-HCl, 500 mM NaCl, 250 mM imidazole pH 8.0). The eluate was collected in 1 mL fractions.

3.2.2.3 Quality control, concentration and storage of proteins

Collected fractions were analysed by SDS-PAGE (section 2.2.4.7). The purest samples were pooled and concentrated by centrifugation at 3,000 $\times g$ at 4 °C in Vivaspin Protein Concentrator Spin columns. Glycerol was added to 20 % (v/v) final concentration and the protein concentration was obtained by absorbance at 280 nm using a Nanodrop ND-2000c (Thermo Fisher). Purified proteins were divided into 250 μL aliquots and stored at -80 °C until use.

3.2.3 Analysis of *B. pseudomallei* CPS biosynthetic pathway *in vitro*

All reactions were either set up manually or automated on an epMotion® 5070 (Eppendorf) robot. Results were obtained using an Infinite M200 microplate reader (Tecan).

3.2.3.1 Pi ColorLock™ Gold colourimetric kit

This assay was used to examine the combined activity of the first three enzymes (GmhA, WcbL, WcbN) in the *B. pseudomallei* CPS biosynthetic pathway. The third reaction, catalysed by WcbN, releases an inorganic phosphate (P_i) molecule that can be detected by absorbance when bound to a dye (Figure 55).

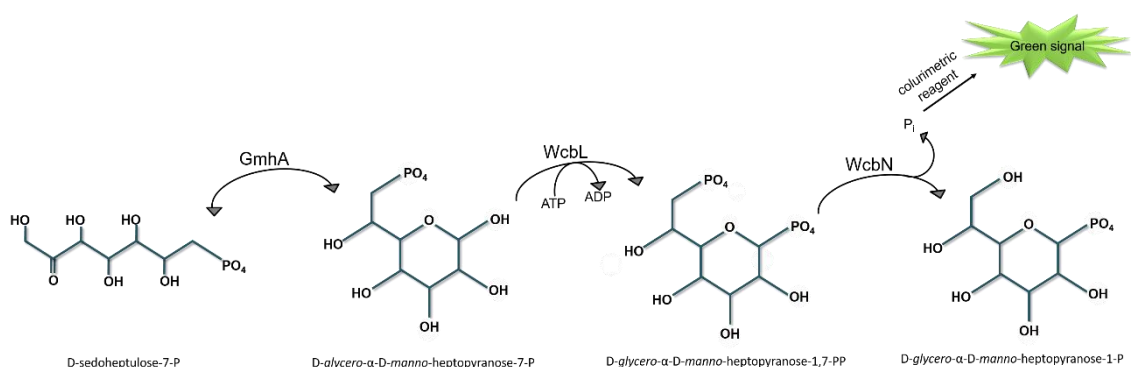


Figure 55: Detection of GmhA, WcbL and WcbN activity. Reactions of the first three enzymes from the *B. pseudomallei* CPS biosynthetic pathway WcbN activity releases inorganic phosphate (P_i) which can be measured with a colourimetric assay.

The 200 μ L reactions were carried out in triplicate in transparent flat-bottomed 96-well microtitre plates (Greiner). To start the reaction, 100 μ L enzyme mix was added to 100 μ L substrate mix. In a standard reaction, the final concentrations were; 2 μ M GmhA, 1 μ M WcbL, 1 μ M WcbN, 0.5 mM S7P and 1 mM ATP. The buffer was 20 mM KCl, 20 mM MgCl₂, 40 mM HEPES pH 8.0. The reaction was run at 37 °C for 10 min. The reaction was stopped with Pi ColorLock™ Gold (Innova Biosciences) added as directed (50 μ L Gold mix and 20 μ L stabiliser). A

P_i standard (0 – 50 μM) was run concurrently for comparison. Finally, the plates were read at 620 nm absorbance.

3.2.3.2 Adapted Pi ColorLock™ Gold colourimetric kit

The Pi ColorLock™ Gold kit was adapted to include the fourth reaction in the pathway, catalysed by WcbM. This reaction releases inorganic pyrophosphate (PP_i), which normally cannot react with the dye. However, inorganic pyrophosphatase (PPIase) breaks this down into two P_i units, which can be detected (Figure 56). A purified recombinant PPIase from *Thermus thermophilus* HB27 was kindly provided by Will Finnigan, University of Exeter. The reactions were performed as previously described with the addition of 1 μM WcbM, 0.6 mM GTP and 135 pM PPIase (final concentrations).

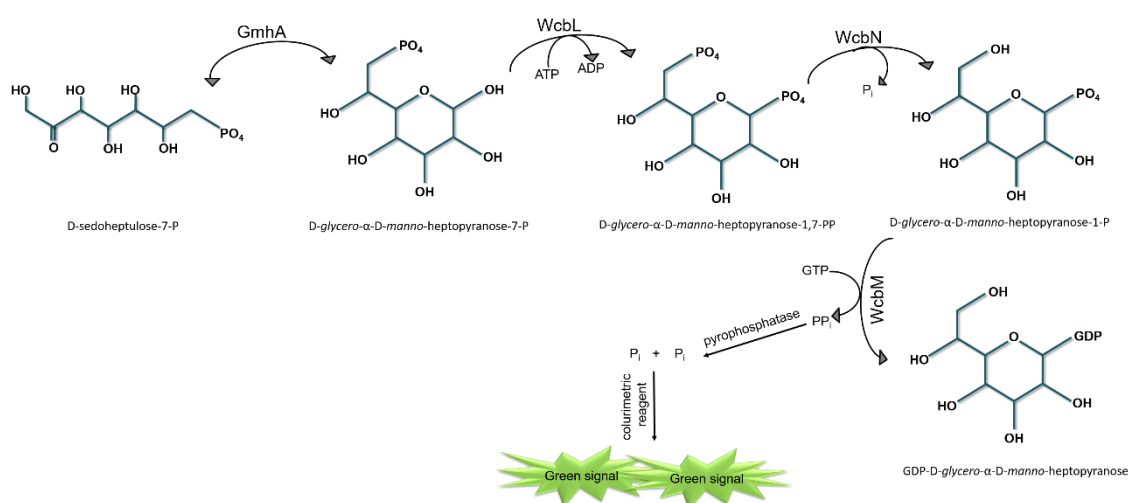


Figure 56: Detection of GmhA, WcbL, WcbN and WcbM activity. Reactions by the first four enzymes from the *B. pseudomallei* CPS biosynthetic pathway. A pyrophosphatase cleaves pyrophosphate released by the WcbM reaction to two inorganic phosphates (P_i). Activity is measured by detecting release of P_i with a colourimetric assay.

3.2.3.3 Kinetic NADPH oxidation assay

This assay was used to examine the activity of the whole *B. pseudomallei* CPS heptose biosynthetic pathway. The final reaction, catalysed by WcbJ, converts

NADP⁺ to NADPH, which can be monitored in real time by a change in absorbance at 340 nm (Figure 57).

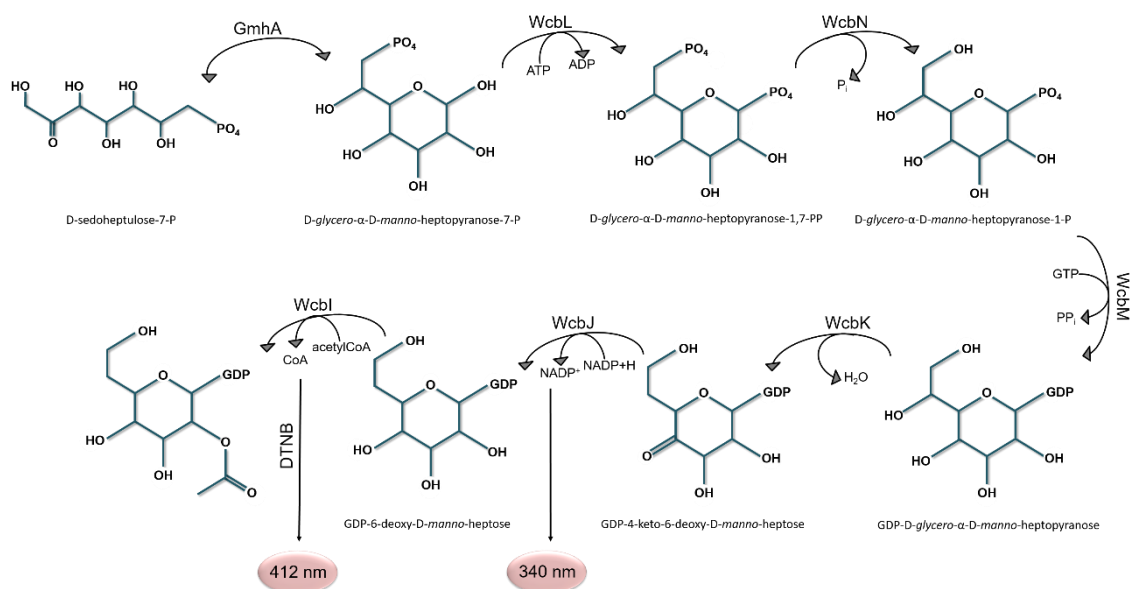


Figure 57: Detection of activity from the *B. pseudomallei* CPS biosynthetic pathway. Reduction of NADP⁺ to NADPH can be detected at 340 nm to measure WcbJ activity. To measure acetylation by WcbI, released CoA is bound to DTNB and detected at 412 nm.

Reactions were set up in 96-well clear bottom black microtitre plates (Corning). The final volume was 200 μ L and the reaction was started by adding 100 μ L enzyme mix to 100 μ L substrate mix. Immediately after starting the reaction the plate was placed in the reader, pre-warmed to 37 $^{\circ}$ C, and the reaction continuously monitored at 340 nm absorbance in approximately 200 s intervals for 1 hr. Final concentrations were; 1 μ M for each enzyme (except 2 μ M GmhA and 135 μ M PPIase), 0.5 mM S7P, 1 mM ATP, 0.6 mM GTP and 0.2 mM NADP⁺. The buffer was 20 mM KCl, 20 mM MgCl₂, 40 mM HEPES pH 8.0.

3.2.3.4 DTNB assay for WcbI activity

Reactions for the entire pathway were set up as previously described, with the addition of 1 μ M WcbI and 1 mM acetyl coenzyme A (acetyl-CoA). The reactions

were run for 10 min at 37 °C, then stopped by adding denaturation buffer to a final concentration 6 M guanidinium chloride (Gu-HCl), 100 mM disodium phosphate, (pH 8.0). DTNB (5,5'-dithio-bis-[2-nitrobenzoic acid]) in 100 mM phosphate buffer, pH 8.0 was added to a final concentration of 10 mM. The plate was read at 412 nm (Figure 57). (DNTB is a colourimetric reagent, which binds CoA released from acetyl-CoA during the WcbI reaction).

3.2.4 Synthesis of GDP-6ddHep in a microfluidics system

3.2.4.1 Microfluidics assay

Reagents were separated into enzymes and substrates for reactions 1 and 2 in two different syringes. Reaction 1 contained: GmhA, WcbL and WcbN with S7P and ATP. Reaction 2 contained WcbM, WcbK, WcbJ and WcbI with GTP, NADP+H and acetyl-CoA. The syringes were mounted onto a vice that applied pressure to generate a specified flow rate. The tube lines from the syringes (0.5 mm diameter) were joined by an X-junction to mix the two and start the reaction. A third line contained mineral oil and this was applied at a specified flow rate to envelop the reaction mix. Droplet size was regulated by manipulating the flow rate of the substrate/enzyme lines and the mineral oil line. The tubing was coiled round a 50 mL falcon tube which was suspended in a 37 °C water bath. The pumps could be stopped to allow the reaction to occur in the water bath for a specified duration or continually flowed through the tubing with the 37 °C incubation time determined by the flow rate. Samples were collected in 1 droplet to 100 droplet volumes (roughly 1 µL to 100 µL). They were analysed by LC-MS for the synthesis of GDP-6ddHep.

3.2.4.2 Preparation of samples for targeted LC-MS (QQQ) analysis

To remove excess mineral oil, the larger samples (10 and 100 droplets) were gently centrifuged (3,000 xg , 5 min, 4 °C) to fully separate the liquid and lipid layers. The mineral oil was removed by pipette and discarded. The volume of each fraction was measured and recorded. To samples 10 μL in volume or less, 50 μL 80 % methanol (kept at -80 °C) was added. Those greater than 10 μL in volume were made up to 100 μL with cold methanol.

The targeted MS/MS (QQQ) was run as previously described (section 2.2.8.2) with additional blank runs included to reduce damage to the column and mass spectrometer by the mineral oil.

3.3 Results

3.3.1 Purification of the biosynthetic enzymes

All six biosynthetic proteins and the putative acetyltransferase Wcbl were purified by nickel affinity and size exclusion chromatography on the Äktaxpress, except GmhA, which was purified by gravity flow nickel affinity chromatography only (Figure 58). The protocol for GmhA purification differed to the other biosynthetic enzymes because it precipitated in the Äktaxpress system (data not shown). However, gravity flow nickel affinity chromatography yielded a high concentration of pure GmhA so further purification steps were not necessary (Figure 58). Wcbl required elution in 10 mM HEPES, pH 7.0 with no NaCl to prevent precipitation.

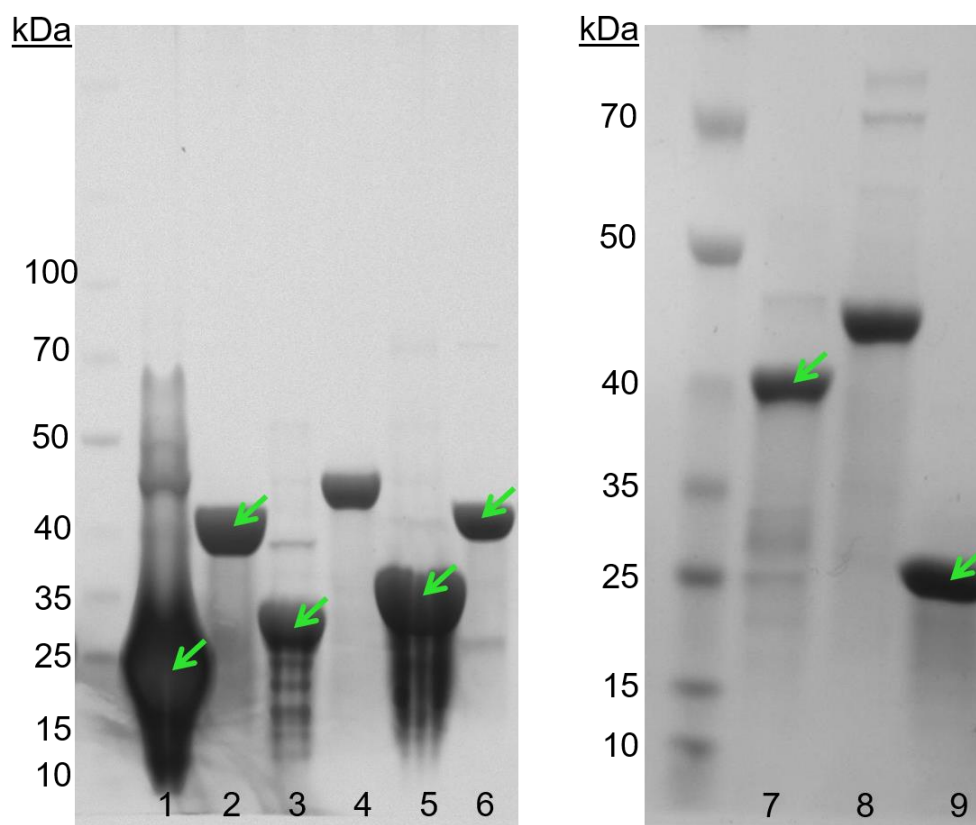


Figure 58: SDS-PAGE of *B. pseudomallei* CPS biosynthetic proteins purified from *E. coli* DH5 α . (1) GmhA, 20.8 kDa; (2) WcbK, 38.1 kDa; (3) WcbM, 24.9 kDa; (4) incorrect protein; (5) WcbJ, 30.8 kDa; (6) WcbL, 38.0 kDa; (7) Wcbl, 36.1 kDa; (8) incorrect protein; (9) WcbN, 20.4 kDa. (1) from gravity flow nickel chromatography, (2-9) from Äkta pure nickel affinity and size exclusion chromatography. Green arrows indicate proteins of correct size.

3.3.2 GDP-6dbHep biosynthetic pathway *in vitro*

3.3.2.1 GmhA, WcbL and WcbN activity

To ensure the biosynthetic pathway was active *in vitro*, the first three enzymes were initially checked for activity. To do this, S7P and ATP were mixed with GmhA, WcbL and WcbN and incubated at 37 °C. The reactions catalysed by GmhA and WcbL would have to proceed to provide the substrate *D-glycero- α -D-manno*-heptopyranose-1,7-pyrophosphate for WcbN. A phosphate group is liberated by the WcbN reaction and its production was measured using the colourimetric Pi ColorLock™ Gold assay. The substrates GTP and NADPH were included to check for potential substrate inhibition in future experiments (Table 8).

Table 8: Enzymes and substrates for GmhA, WcbL, WcbN activity assay. GTP and NADPH are included to check for potential substrate inhibition for future experiments.

Enzymes	Substrates	Phosphate (μM)	Phosphate $\mu\text{M min}^{-1}$
GmhA, WcbL, WcbN	S7P, ATP	77.00	7.70
GmhA, WcbL	S7P, ATP	25.55	2.56
GmhA, WcbL, WcbN	S7P, ATP, GTP	71.16	7.12
GmhA, WcbL, WcbN	S7P, ATP, NADPH	68.35	6.84
GmhA, WcbL, WcbN	S7P, ATP, GTP, NADPH	72.78	7.28

Table 8 shows most activity in the first reaction with GmhA, WcbL, WcbN, S7P and ATP at 7.70 $\mu\text{M min}^{-1}$ phosphate produced. There was slight inhibition of the reaction upon addition of GTP and/or NADPH (lowest rate 6.84 $\mu\text{M min}^{-1}$ phosphate for NADPH only). Phosphate production was lowest in the assay with no WcbN as expected (2.56 $\mu\text{M min}^{-1}$). The amount of phosphate detected was probably due to contaminating free phosphate in the buffer and degradation of the S7P and ATP substrates.

3.3.2.2 Addition of WcbM

WcbM is the next enzyme in the pathway and releases a pyrophosphate (PPi), which can be broken down into two phosphates by a pyrophosphatase. The two liberated inorganic phosphates can be detected with the Pi ColorLock™ Gold assay.

Table 9: Enzymes and substrates for GmhA, WcbL, WcbN, WcbM activity assay. GTP and NADPH are included to check for potential substrate inhibition for future experiments.

Ppiase	Enzymes	Substrates	Phosphate (μM)	Phosphate $\mu\text{M min}^{-1}$
Ppiase	GmhA, WcbL	S7P, ATP, GTP	13.59	1.36
Ppiase	GmhA, WcbL, WcbN	S7P, ATP	51.01	5.10
Ppiase	GmhA, WcbL, WcbN, WcbM	S7P, ATP, GTP	21.24	2.12
x	GmhA, WcbL, WcbN, WcbM	S7P, ATP, GTP	17.85	1.79
Ppiase	GmhA, WcbL, WcbN, WcbM	S7P, ATP	48.84	4.88
Ppiase	GmhA, WcbL, WcbM	S7P, ATP, GTP	15.36	1.54
Ppiase	GmhA, WcbM	S7P, ATP, GTP	15.25	1.52
Ppiase	None	S7P, ATP	15.11	1.51

Table 9 shows the highest reaction rate is achieved by GmhA, WcbL and WcbN with S7P and ATP at $5.10 \mu\text{M min}^{-1}$ phosphate produced. When WcbM is added, the rate drops to $4.88 \mu\text{M min}^{-1}$ phosphate. As WcbM liberates PPi and there is a pyrophosphatase present to break it down into two phosphate molecules, the expected rate would be approximately $10.00 \mu\text{M min}^{-1}$ phosphate, (double the rate observed by GmhA, WcbL and WcbN). Addition of GTP reduced the rate further to $2.12 \mu\text{M min}^{-1}$, suggesting GTP is causing inhibition of the pathway. Removal of the pyrophosphatase decreases the rate to $1.79 \mu\text{M min}^{-1}$, showing that although it is not sufficient to recover activity in the pathway, it does improve the rate of phosphate production a little. All the controls where no reaction was expected had background noise between 1.36 and $1.54 \mu\text{M min}^{-1}$ phosphate.

Table 10: Enzymes and substrates for separate GmhA, WcbL and WcbN then WcbM reactions. Also tested the effect of separating the two reactions and on doubling the concentrations of GmhA, WcbL, WcbN and WcbK. All reactions contained PPIase.

Reaction 1		Reaction 2		Phosphate	
Enzymes	Substrates	Enzymes	Substrates	(μM)	($\mu\text{M min}^{-1}$)
GmhA, WcbL, WcbN	S7P, ATP	WcbM	GTP	68.86	6.89
GmhA, WcbL, WcbN	S7P, ATP	-	-	71.64	7.16
GmhA, WcbL, WcbN, WcbM	S7P, ATP, GTP	-	-	36.08	3.61
GmhA, WcbL, WcbN, WcbM, WcbK	S7P, ATP, GTP	-	-	46.21	4.62
2XGmhA, WcbL, WcbN, WcbM	S7P, ATP, GTP	-	-	37.22	3.72
GmhA, 2XWcbL, WcbN, WcbM	S7P, ATP, GTP	-	-	38.58	3.86
GmhA, WcbL, 2XWcbN, WcbM	S7P, ATP, GTP	-	-	35.67	3.57
GmhA, WcbL, WcbN, WcbM, 2XWcbK	S7P, ATP, GTP	-	-	42.91	4.29

Next, the separation of WcbM from the rest of the pathway and the doubling of the concentrations of GmhA, WcbL, WcbM and WcbK were tested. All reactions contained PPIase. Despite not being able to monitor the WcbK reaction, it was included to see if removing the product from the WcbM reaction (GDP-D-glycero-a-D-manno-heptopyranose) would alleviate the inhibitory effects.

Table 10 shows that the first three reactions have the quickest rate at 7.16 $\mu\text{M min}^{-1}$ phosphate. The next highest rate was the separated WcbM reaction at 6.89 $\mu\text{M min}^{-1}$ phosphate. This was half the rate expected, though it is greatly improved from the combined reaction (3.61 $\mu\text{M min}^{-1}$ phosphate). Addition of WcbK slightly improved the reaction (to 4.62 $\mu\text{M min}^{-1}$ phosphate), although doubling the concentration of either GmhA, WcbN or WcbL had no effect.

Finally, different combinations of enzyme concentration and separation of reactions were tested. All reactions including pyrophosphatase (Table 11).

Table 11: Enzymes and substrates for testing the effect of different enzyme concentrations and separation of reactions. All reactions contained PPIase.

Reaction 1		Reaction 2		Phosphate	
Enzymes	Substrates	Enzymes	Substrates	(μM)	($\mu\text{M min}^{-1}$)
GmhA, WcbL	S7P, ATP	WcbN, WcbM	GTP	72.37	7.24
GmhA	S7P	WcbL, WcbN, WcbM	ATP, GTP	35.82	3.58
GmhA, WcbL, WcbN, 0.5xWcbM, WcbK,WcbJ	S7P, ATP, GTP, NADPH	-	-	24.03	2.40
GmhA, WcbL, WcbN, WcbM, WcbK,WcbJ	S7P, ATP, GTP, NADPH	-	-	23.38	2.34
GmhA, WcbL, WcbN	S7P, ATP	WcbM	GTP	73.44	7.34
GmhA, WcbL, WcbN, WcbM	S7P, ATP, GTP	-	-	27.41	2.74
-	S7P, ATP, GTP, NADPH	-	-	11.24	1.12

Table 11 shows that separating the WcbM reaction achieved the highest activity rate $7.34 \mu\text{M min}^{-1}$ phosphate (compared to $2.74 \mu\text{M min}^{-1}$ phosphate when all reactions are run together) as expected. Addition of WcbK and WcbJ did not restore activity ($2.34 \mu\text{M min}^{-1}$ phosphate). Halving the concentration of WcbM with WcbK and WcbJ included did not improve the rate of activity ($2.40 \mu\text{M min}^{-1}$ phosphate).

Inhibition caused by the WcbM reaction does not directly affect WcbN activity, as including WcbN in the second reaction had no effect on rate ($7.24 \mu\text{M min}^{-1}$ phosphate). However, inclusion of WcbL in the second reaction caused the rate to drop to $3.58 \mu\text{M min}^{-1}$ phosphate, suggesting this is where the inhibitory effect

occurs. Reactions with no enzymes produced a background noise of $1.12 \mu\text{M min}^{-1}$ phosphate.

3.3.2.3 Addition of WcbK and WcbJ

The final two enzymes, WcbK and WcbJ were added to the reaction. WcbJ oxidises NADP+H which can be monitored at absorbance 340 nm.

Figure 59 shows that WcbJ can oxidise either NADH or NADPH. The reaction proceeds more quickly if you separate WcbM, WcbK and WcbJ from the first three steps in the pathway, confirming observations from previous experiments.

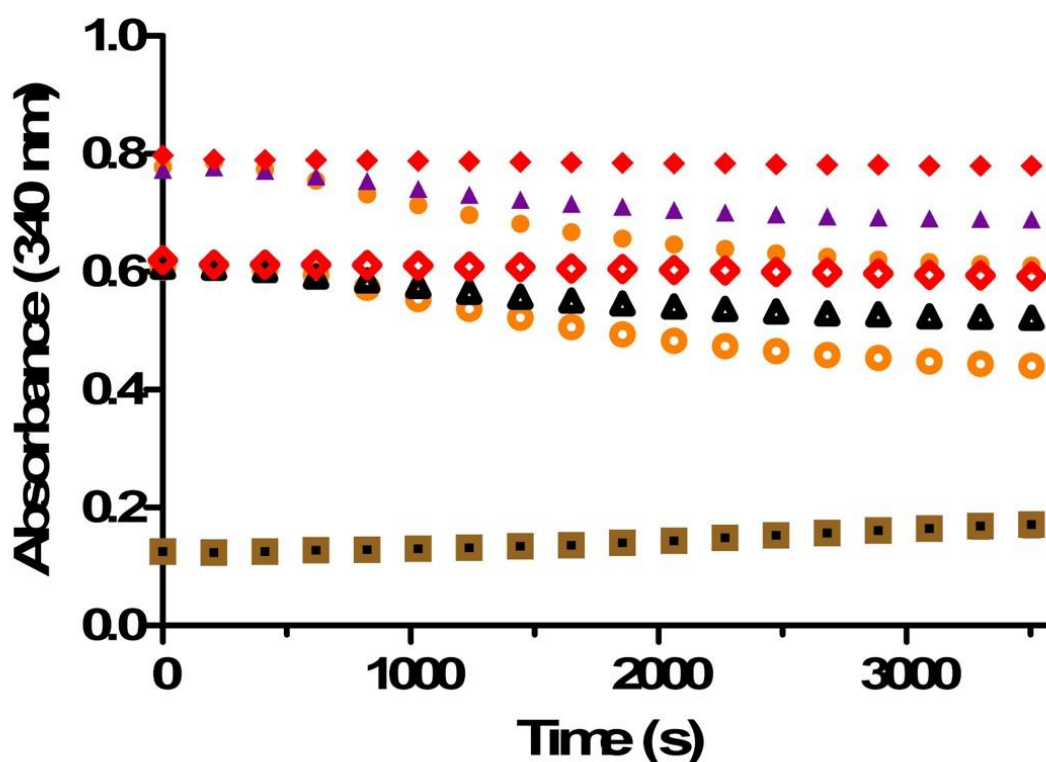


Figure 59: GDP-6dbHep biosynthetic pathway from GmhA through to WcbJ monitoring oxidation of NADH or NADPH. Closed symbols for NADPH, open symbols for NADH. (Red diamond) NADPH or NADH only control; (purple triangle) All six enzymes in one reaction; (orange circle) GmhA, WcbL and WcbN in reaction one, then WcbM, WcbK and WcbJ in reaction 2. (Brown square) all enzymes, no reagent; (Black square) GmhA, WcbL, WcbN in reaction 1, then WcbM, WcbK and WcbJ in reaction 2 with no substrates.

Overall these data show that the six biosynthetic proteins are active and perform the functions as annotated on the CPS coding locus (Cuccui *et al.*, 2012) and as observed in an experiment with orthologous enzymes (Butty *et al.*, 2009). The reaction pathway can be monitored *in vitro* by measuring WcbN and WcbM reactions by phosphate release and by monitoring NADPH oxidation by the WcbJ reaction.

3.3.3 GDP-6dbHep biosynthetic pathway in a microfluidics system

After verifying the assay is active *in vitro*, experiments were conducted in 1 μL to 100 μL volumes in the microfluidics system. Droplets were collected and analysed by MS-MS QQQ for presence of GDP-6dbHep, GDP-2-O-acetyl-6-deoxy- β -D-*manno*-heptopyranose and intermediates in the biosynthetic pathway. S7P and GDP-mannose were also scanned for, as positive and negative controls respectively (Table 12).

Table 12: MS-MS QQQ analysis of droplets recovered from microfluidics samples. Reactions contained six biosynthetic enzymes GmhA, WcbL, WcbN, WcbN, WcbK and WcbJ and putative acetyltransferase, WcbI. Substrates were S7P, ATP, GTP, NADP+H and acetyl-CoA. Products and intermediates screened for were S7P, *D-glycero- α -D-manno*-heptopyranose-1,7-bisphosphate (M1,7P), GDP-*D-glycero- α -D-manno*-heptopyranose, (GDP-MH), GDP-4-keto-6-deoxy-*D-manno*-heptose (GDP-KH), GDP-6dHep, GDP-2-O-acetyl-6-deoxy- β -*D-manno*-heptopyranose (GDP-2OAc-6dHep and GDP-mannose. Samples are recorded by volume except for the S7P 0.5mM and 1.0 mM standards and the blanks as indicated. RT = retention time; rA = relative abundance, derived from area under curve on MS-MS chromatogram.

Sample	S7P		M1,7PP		GDP-MH		GDP-KH		GDP-6dHep		GDP-2OAc-6dHep		GDP Mannose	
	RT	rA	RT	rA	RT	rA	RT	rA	RT	rA	RT	rA	RT	rA
100 μ L	12.36	1108	10.28	8	12.41	834	11.64	10	11.77	3055	9.26	2	10.41	2
20 μ L	12.52	82	10.33	18	12.55	59	11.75	1	11.9	2016	9.57	1	10.41	2
15 μ L	12.76	163	10.24	2	12.57	432	11.68	2	11.92	1558	9.51	1	10.12	2
50 μ L	12.34	1345	10.08	3	12.39	1349	11.41	3	11.74	847	9.24	1	10.12	3
5 μ L	12.52	35	10.3	3	12.48	199	11.59	5	11.86	559	9.44	2	10.25	1
5 μ L	12.29	22974	10.25	15	12.47	77	11.25	4	11.81	426	-	-	10.24	7
2 μ L	12.18	18685	10.19	10	12.32	38	11.52	4	11.7	303	9.71	2	10.25	4
10 μ L	12.52	277	10.3	20	12.48	535	11.48	5	11.83	254	9.42	2	10.03	2
1 μ L	12.67	2236	10.39	25	12.81	9	11.86	2	11.94	92	9.06	2	10.01	2
1 μ L	12.78	2519	10.46	16	12.61	22	11.55	5	11.94	86	9.53	3	10.34	3
S7P 0.5 mM	12.24	5439	-	-	12.55	2	11.34	1	11.63	1	-	-	10.08	4
S7P 1.0 mM	12.15	11831	-	-	12.33	1	11.43	1	11.71	1	-	-	10.19	3
Blank	12.8	6	10.4	15	12.49	1	11.52	3	11.62	1	-	-	10.15	3
Blank	12.84	5	10.28	10	12.28	2	11.42	2	11.64	0	-	-	10.11	6

The results from MS-MS analysis of the reactions run in droplets in the microfluidics system show that the six biosynthetic enzymes were active and produced GDP-6dDHep from S7P (Table 12). GDP-6dDHep was absent in the S7P and blank samples as expected, but present in all the droplets containing the enzymes and substrates. It was produced to detectable levels in volumes as low as 1 μ L. Whilst relative abundance (rA) of GDP-6dDHep roughly correlated with increasing volume (for example, 89 rA in 1 μ L, 254 rA in 10 μ L, 3,055 in 100 μ L) there are some anomalies (2,016 rA in 20 μ L compared to 847 rA in 50 μ L). This is probably due to difficulties in accurately counting the number of droplets collected and accidental removal of aqueous solution when extracting the lipid fraction in preparation for LC-MS. There was no GDP-4-keto-6-deoxy-D-*manno*-heptose and only trace amounts of D-glycero- α -D-*manno*-heptopyranose-1,7-bisphosphate detected in the samples, suggesting these intermediates are rapidly turned over in the subsequent reaction, (WcbJ and WcbN respectively), or are unstable products. Interestingly, GDP-D-glycero- α -D-*manno*-heptopyranose, the product of the WcbM reaction, reached up to 1,349 rA in the 50 μ L reaction. This could be the product that causes inhibition of the WcbL kinase reaction. No acetylation of GDP-6dDHep was detected. Possible explanations are WcbI is not the acetyltransferase, it may require co-purification with another enzyme or the acetylation may occur only on polymerised GDP-6dDHep. Finally, the S7P and GDP-mannose controls were present and absent in the samples respectively as expected.

3.4 Discussion

The synthetic biology approach to generating new products requires the redesign of a natural process to perform a desired function within a cell. To streamline the design process, it is helpful for constituent parts to be characterised. There has been a drive towards the standardisation and open source availability of biological components to simplify the synthetic biology design process. However, characterisation of pathways for polysaccharide biosynthesis has proven challenging due to difficulties in expressing and purifying Leloir-type glycosyltransferases, which have the GT-B Rossmann-like fold and are membrane-associated (Breton *et al.*, 2006; Albesa-Jové *et al.*, 2014). In the development of a platform for the biosynthesis of bespoke polysaccharides, understanding the enzyme activity and kinetic parameters of glycosyltransferases is vital for the design of a novel system (Chapter 5). As they are challenging and costly to obtain, purified glycosyltransferases are precious and this can limit the amount of high-throughput screening that can be conducted to elucidate their functions and characterise their kinetic parameters.

For these reasons, a new approach to *in vitro* experimentation was sought. An emerging technology is the field of microfluidics, which has found application for small volume vesicles in many diverse areas (Baxani *et al.*, 2016). It was identified as a potentially useful method for conducting low-volume enzyme assays.

3.4.1 Microfluidics as a novel system for conducting enzyme assays *in vitro*

3.4.1.1 Droplet formation

The microfluidics system creates droplets or vesicles that consist of a lipid coat containing an aqueous solution, flowing through a small-diameter capillary line.

This is achieved by introducing the lipid (in this case, mineral oil) into a flowing aqueous solution. The process of droplet formation causes turbulence within the encapsulated aqueous fraction, which is ideal for the mixing of enzymes and substrates in an assay.

By adjusting the flow rate of the lipid and the aqueous solution the size, shape and frequency of droplet formation can be influenced. One unforeseen complication was the inclusion of glycerol in the aqueous solution which derived from the enzyme stocks. The viscosity of the glycerol interfered with droplet formation making the size, shape and frequency less predictable. This meant flow rates were constantly monitored and adjusted during experimentation to maintain consistency in droplet volume.

3.4.1.2 Droplet structure and stability

In contrast to the turbulent interior, the lipid coat could provide stability to glycosyltransferases, which are usually membrane-associated within a cell and rapidly denature during when removed from the membrane (as observed in Chapter 5). The system is currently being further developed to join two droplets together by forming a lipid bilayer between them but maintaining separate interiors. The lipid bilayer could be the ideal environment for glycosyltransferase stability and activity. Another modification could be the insertion of α -haemolysin into the lipid bilayer, providing a channel between the two droplets. The hypothesis is that the product of an enzyme reaction could be transported into the second droplet and be more easily quantified when collected at the end of the experiment. Another application could be the removal of a toxic product from a pathway, enabling downstream reactions to take place uninhibited.

Another issue that had not been encountered before in the context of the microfluidics system was the precipitation of proteins out of solution. Fortunately, the six biosynthetic enzymes remained stable and soluble throughout the experimentation. However, WcbI was observed precipitating out of solution, which could be an indication as to why there was no acetylation of the GDP-6dDHep.

3.4.1.3 The problem of WcbM (part 2)

A major advantage of the microfluidics system was the ability to flow two different aqueous solutions together. By adjusting the flow rates, the ratio of each solution could be altered. By maintaining one solution at a constant rate and raising the other on a gradient, the system could be used to produce enzyme activity and Michaelis-Menten curves in the future.

Initially, the idea was to use separate the enzymes from the substrates. Activation of the pathway would only occur upon mixing the two solutions. However, *in vitro* assay data showed that WcbM significantly inhibits the pathway correlating with the previous observations of Butty *et al.* (2009). Therefore the separate flows were used to separate the pathway into two sets of reactions; GmhA, WcbL and WcbN, then WcbM, WcbJ and WcbK.

Interestingly, the LC-MS data showed presence of GDP-D-glycero- α -D-manno-heptopyranose (the product of the WcbM reaction) in the droplets (Table 12). None of the other intermediates were present in detectable quantities, demonstrating that this nucleotide sugar accumulates during the reaction and is likely the source of inhibition of WcbL.

Here, a novel system for conducting small-volume *in vitro* assays was identified. It was possible to allow six enzyme activities to proceed simultaneous in a volume of as little as 1 μL and still detect the end product by mass spectrometry. It also enabled the accumulation of an intermediate product to be detected to provide a clue as to why there is inhibition in the pathway. As there was a low yield of GDP-6dbHep recovered from *E. coli* in Chapter 1, perhaps step in the pathway required attention for optimisation of the operon to improve productivity. There is much scope to further develop the microfluidics system for conducting enzyme assays, especially as it may allow the screening of minute quantities of purified glycosyltransferases and fill the knowledge gap in the published literature.

Chapter 4

Development of a novel assay for characterisation of glycosyltransferases

4.1 Introduction

Following biosynthesis of the nucleotide sugar precursor, GDP-6_DHep, the next step is to assemble the CPS polymer. As there are no polymerases or chain length determinants present on the CPS I coding locus, Cucci *et al.* (2012) proposed a group 3 assembly mechanism mediated by glycosyltransferases. There are three glycosyltransferases on the coding locus, WcbB, WcbE and WcbH that are required for capsule expression in *B. pseudomallei* (Cuccui *et al.*, 2012). They have not been purified or characterised however the proposed functions are the attachment of a mannose priming sugar to a lipid anchor, linking *manno*-heptose to the mannose primer and finally polymerisation of the *manno*-heptose residues (Cuccui *et al.*, 2012). To characterise the glycosyltransferases, they must first be purified. However, due to the association of many glycosyltransferases with cell membranes, they become unstable in extracellular conditions and are challenging to purify. The difficulty in obtaining glycosyltransferases means purified enzyme is precious. Chapter 3 discussed the potential of a microfluidics approach for enzyme assays which would greatly reduce the reaction volume and therefore the quantity of enzyme required per reaction. However, lack of pure glycosyltransferases has resulted not only in a scarcity of structural and kinetic data for these enzymes, but also of assays for characterising their kinetic parameters. Many commercially-available assays are either expensive (Wu *et al.*, 2011), require specialist equipment (such as capillary-electrophoresis MS; Holden *et al.*, 2012) or consume a lot of enzyme (Tedaldi *et al.*, 2014). As more glycosyltransferases are successfully purified it is desirable to have a cost-effective, reliable assay for characterising their activities and is essential for the full understanding of the CPS synthesis pathway in *B. pseudomallei*.

4.1.1 Discontinuous kinase assay

For Leloir-type glycosyltransferases, such as those involved in *B. pseudomallei* CPS biosynthesis, the reaction produces a free nucleotide diphosphate. During the reaction, a nucleotide sugar donor is linked to an acceptor via a glycosidic bond, liberating the nucleotide diphosphate (NDP) (Figure 60).

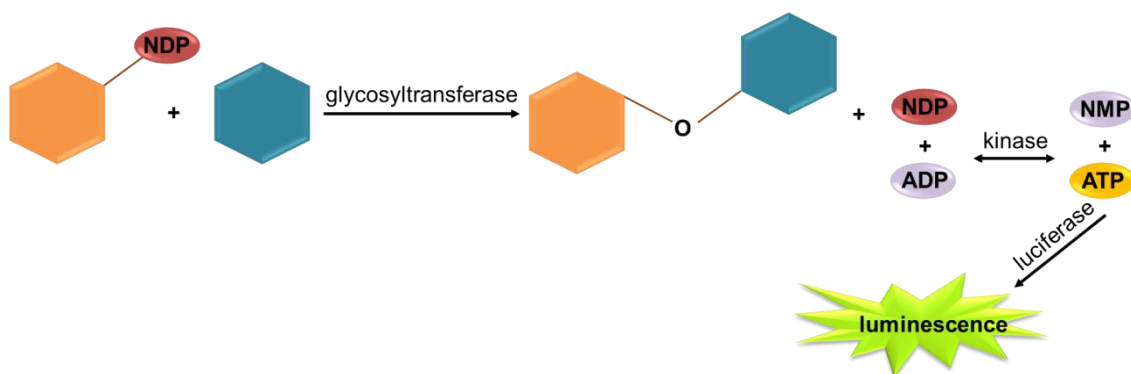


Figure 60: Discontinuous coupled kinase assay for glycosyltransferase activity. Glycosyltransferase activity with a nucleotide sugar donor can be characterised using a coupled kinase reaction. Nucleotide diphosphate liberated from the glycosyltransferase reaction can be used to generate ATP by addition of a kinase and ADP. The production of ATP is proportional to nucleotide diphosphate released by the glycosyltransferase assay and can be measured quantitatively by a luminescence kit (BacTiter-Glo™ (Promega)).

The addition of a kinase and ADP reversibly produces nucleotide monophosphate (GMP) and ATP (Figure 60). Production of ATP can be quantitatively measured using a commercially-available bioluminescence assay (Figure 61). Marketed as a test for microbial cell viability, BacTiter-Glo™ (Promega) was adapted for use in this assay. The reagent buffer contains luciferin and a recombinant luciferase that utilises ATP from the coupled-kinase assay for generating luminescence (Figure 61).

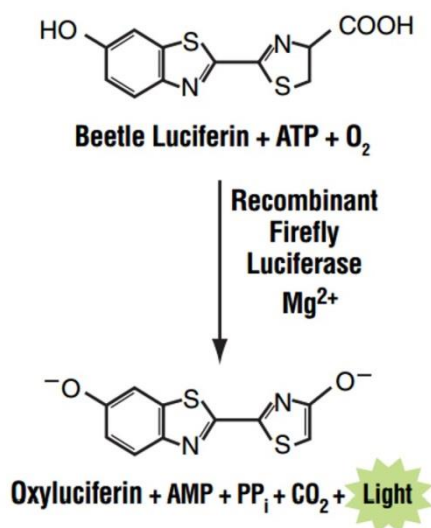


Figure 61: BacTiter-Glo™ reagent activity. Conversion of beetle luciferin in the BacTiter-Glo™ buffer plus ATP generated by the coupled-kinase assay and atmospheric oxygen to oxyluciferin, light (luminescence) and by-products AMP, PP_i and carbon dioxide (CO₂). (Reproduced from BacTiter-Glo™ Technical Bulletin TB337, Promega)

4.1.1.1 Mesophilic Cmk and Gmk from *E. coli*

Several kinases are found in *E. coli* in pathways for purine and pyrimidine metabolism and catalyse the reaction $\text{NMP} + \text{ATP} \rightleftharpoons \text{NDP} + \text{ADP}$ reversibly. To ensure the assay is applicable to a broad range of glycosyltransferase activities, different kinases are required. Table 13 shows the nucleotide substrate specificities of cytidylate, guanylate, thymidylate and uridylate kinases (Cmk, Gmk, Tmk and Umpk respectively) (Fricke *et al.*, 1995; Gentry *et al.*, 1993; Reynes *et al.*, 1996; Serina *et al.*, 1995). The Cmk enzyme can also use UMP as a substrate, though with only 0.8 % activity compared with CMP (Briozzo *et al.*, 1998). The literature available discusses these kinases with the nucleoside monophosphate as the substrate of interest, so no values for the NDP as the substrate currently exist.

Table 13: Mesophilic and thermophilic kinase parameters. Temperature tolerance, reaction specificities, molecular weight and isoelectric point of kinases purified for assay development.

Growth temperature	Kinase	Reaction(s)	Molecular weight (kDa)	Isoelectric point (pI)
Mesophilic	Cmk	CDP + ADP \rightleftharpoons CMP +ATP UDP + ADP \rightleftharpoons UMP +ATP	24.7	5.56
	Gmk	GDP + ADP \rightleftharpoons GMP +ATP	23.6	6.06
Thermophilic	Cmk	CDP + ADP \rightleftharpoons CMP +ATP UDP + ADP \rightleftharpoons UMP +ATP	22.6	6.86
	Gmk	GDP + ADP \rightleftharpoons GMP +ATP	25.3	6.88
	Tmk	TDP + ADP \rightleftharpoons dTMP +ATP	22.0	9.67
	Umpk	UDP + ADP \rightleftharpoons UMP +ATP	25.3	6.76

4.1.1.2 Thermophilic Cmk, Gmk, Tmk and Umpk from *T. thermophilus*

Homologues for the kinases are found in many eukaryote and prokaryote species. One improvement investigated was the use of thermostable kinases from *Thermus thermophilus*. The additional ADP and kinase are included with the glycosyltransferases and its reagents in the same solution (Figure 62). This would enable glycosyltransferase activity to occur at 37 °C, before raising the temperature to 65 °C, denaturing the glycosyltransferase and allowing the kinase reaction to proceed.

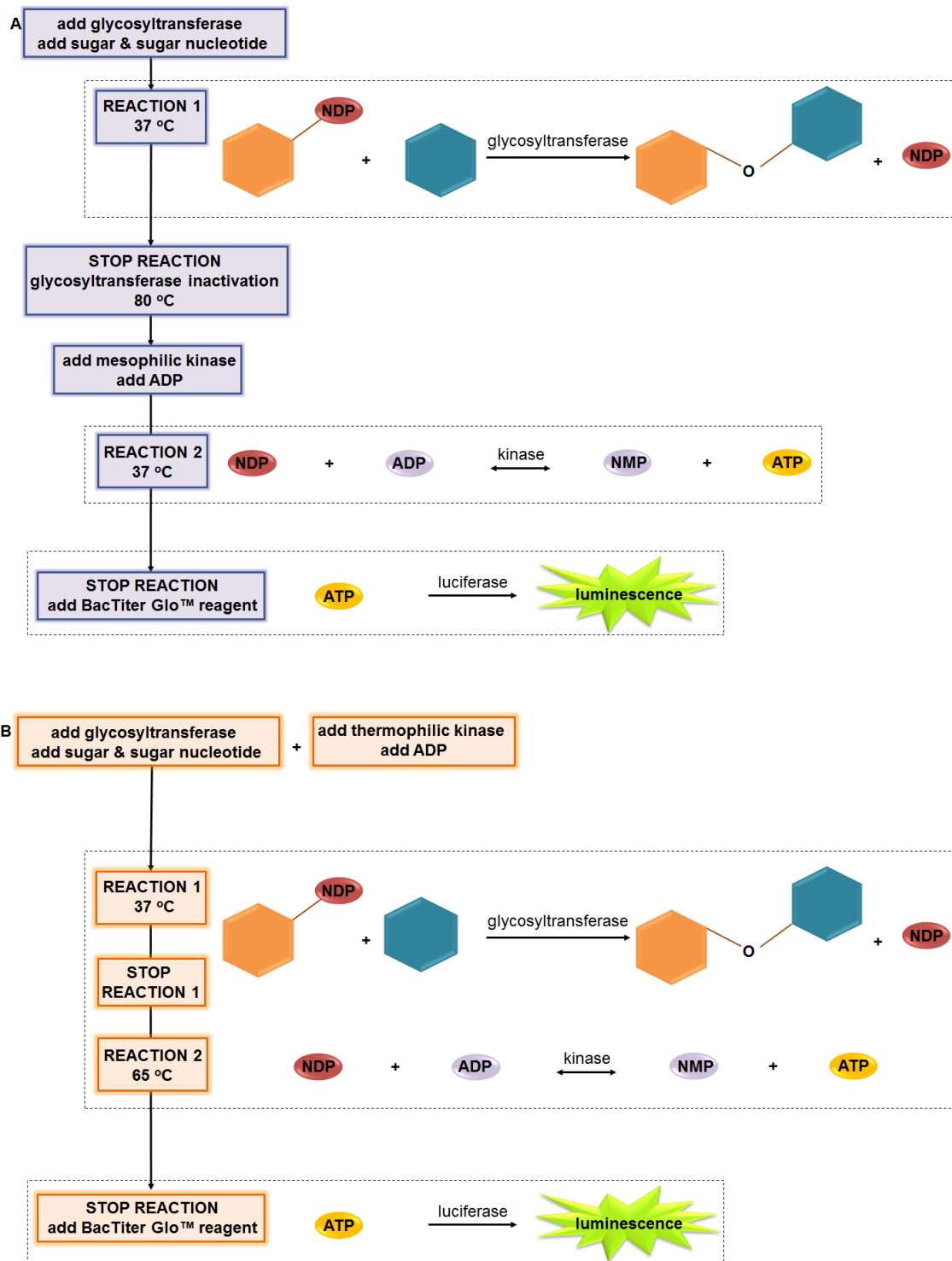


Figure 62: Glycosyltransferase assay coupled with mesophilic or thermophilic kinase. (A) The glycosyltransferase reaction is run at 37 °C. The reaction is stopped by denaturing the enzyme at 80 °C. The mesophilic kinase and ADP is added and the reaction run at 37 °C. Kinase activity is stopped by addition of the BacTiter Glo™ reagent and luminescence detected. (B) All reagents are added at the start. The reaction is run at 37 °C to allow glycosyltransferase activity to occur then the temperature ramped up to 65 °C. This denatures the glycosyltransferase but enables thermophilic kinase activity. Kinase activity is stopped by addition of the BacTiter Glo™ reagent and luminescence detected.

Four thermostable kinases, Cmk, Gmk, Tmk and Umpk, provide a full complement of coupled-enzyme reactions for characterising CDP-, GDP-, dTDP- and UDP-sugar glycosyltransferase activities (Table 1). The novel assay must be user-friendly and adaptable to a wide range of glycosyltransferases as well as cost-effective, to compete with existing assays on the market. Although differing in DNA sequence, the kinases share structural similarities and conserved regions within and between species (Bucurenci *et al.*, 1996).

4.1.2 Discontinuous phosphatase assay using an apyrase

A second approach to developing a novel glycosyltransferase assay was explored. The alternative method is based on the Glycosyltransferase Activity Kit (R&D systems; Wu *et al.*, 2011). The kit uses a human recombinant ectonucleoside triphosphate diphosphohydrolase to release an inorganic phosphate from the liberated nucleotide diphosphate. This can be measured quantitatively using a colourimetric detector such as Malachite Green (Figure 63).

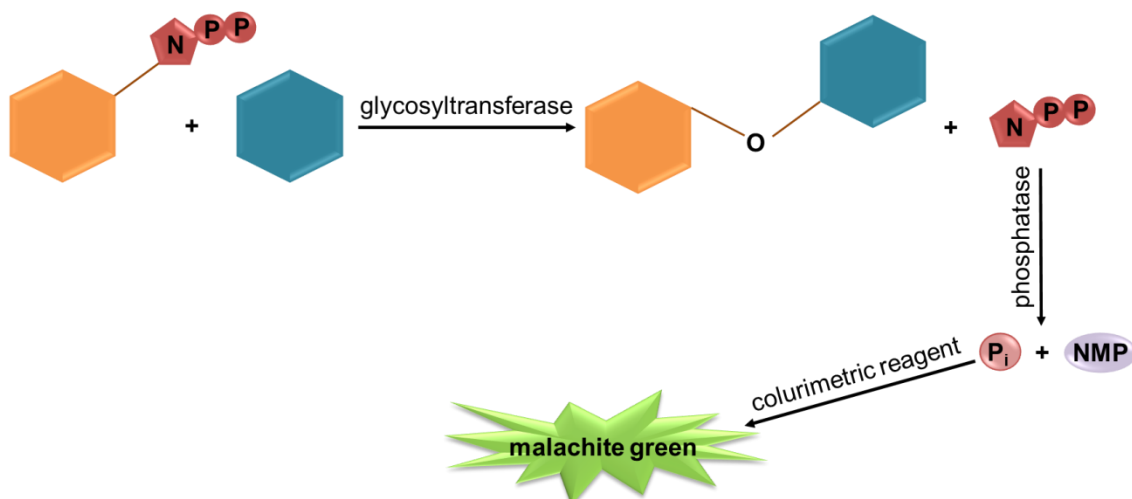


Figure 63: Discontinuous coupled phosphatase assay for glycosyltransferase activity with a nucleotide sugar donor can be characterised using a coupled phosphatase reaction. Nucleotide diphosphate liberated from the glycosyltransferase reaction can be hydrolysed to produce nucleotide monophosphate and a free inorganic phosphate (P_i). Production of P_i can be detected by malachite green reagent and quantitatively related to the glycosyltransferase activity. .

The advantage of this approach is the phosphatase is able to use any nucleotide diphosphate as a substrate, allowing a range of glycosyltransferase activities to be characterised. It also does not require an additional step before the quantitative reaction (i.e. the proposed kinase assay requires conversion of NDP release to ATP generation before the ATP can be measured; Figure 60). However, the commercial kit is prohibitively expensive and requires a large quantity of purified glycosyltransferase. The kit is expensive because the enzyme (human recombinant ectonucleoside triphosphate diphosphohydrolase) is very difficult to purify.

Apyrase from bedbug *Cimex lectularius*, is proposed here as an alternative to the human recombinant ectonucleoside triphosphate diphosphohydrolase. The main activity of apyrase is hydrolysis of ATP > ADP > AMP, however it is able to use other nucleotide diphosphates, such as UDP and GDP, as substrates (Valenzuela *et al.*, 1996; 1998).

4.1.3 β -1,4-galactosyltransferase and LgtC

As none of the *B. pseudomallei* glycosyltransferases for CPS biosynthesis have been purified, two other enzymes were used for developing the assay. The first was β -1,4-galactosyltransferase (β 14GalT) (Amado *et al.*, 1999; Ramakrishnan & Qasba, 2001; Qasba *et al.*, 2008). Found in eukaryotes, including seven homologues in humans, this was a commercially-available preparation from *Bos taurus* milk. β 14GalT catalyses the reaction: UDP-galactose + *N*-acetylglucosamine (GlcNAc) \rightleftharpoons UDP + *N*-acetyllactosamine. In mammary glands, it associates with α -lactalbumin where it performs a secondary reaction: UDP-galactose + D-glucose \rightleftharpoons UDP + lactose.

The second enzyme, LgtC, is part of the biosynthetic pathway for LPS synthesis in *N. meningitidis*. LgtC homologues are found in other pathogenic Gram-negative bacteria such as *H. influenzae*. It is an α -1,4-galactosyltransferase catalysing the addition of α -galactose from UDP-galactose to the terminal acceptor lactose on nascent LPS (Persson *et al.*, 2001; Chan *et al.*, 2012; Xu *et al.* [under review]).

4.1.1 Aims

The aim of this chapter is to develop a novel assay for the characterisation of glycosyltransferases. It must be accurate, reliable and more cost-effective than commercially-available assays. It is also desirable to achieve a low reaction volume to reduce the quantity of enzyme required. Ideally, the assay would have scope to determine the kinetic parameters for wide range of novel glycosyltransferase activities, contributing to a wider platform for a biosynthetic approach to manufacturing polysaccharides.

The purpose of these aims is to develop an economically-viable assay for characterising WcbB, WcbE and WcbH, which are required for CPS expression in *B. pseudomallei*.

4.2 Materials and Methods

All reagents were purchased from Sigma-Aldrich unless otherwise stated.

4.2.1 Preparation of kinases for assay development

4.2.1.1 Cloning genes for mesophilic kinases

Genes encoding mesophilic kinases, *cmk* and *gmk*, were cloned from *E. coli* K-12 genomic DNA using primers with flanking sequences for ligation-independent cloning (LIC) (Appendix 1). PCR was performed as previously described (section 2.2.3.4). Bands were excised from a 1 % (w/v) agarose gel and purified using a QIAquick gel extraction kit (Qiagen). The genes were sub-cloned into vector pNIC28-Bsa4 using LIC (Savitsky *et al.*, 2010). The constructs were transformed into *E. coli* DH5 α cells as described (section 2.2.3.2). Plasmids were then purified from DH5 α cells, sequences verified by Sanger sequencing, and transformed into Rosetta (DE3) cells for protein expression using the same transformation procedure.

4.2.1.2 Gene synthesis for thermostable kinases

Genes for thermostable kinases (*cmk*, *gmk*, *tmk* and *umpk*) from *Thermus thermophilus* HB8 were codon-optimised for expression in *E. coli* and synthesised in pEX-A2 plasmids (Eurofins). The sequences included the T7 promoter, lac operator, RBS and TEV cleavable N-terminal His-tag sequences from pNIC28-Bsa4 (Appendices 2-5). Once reconstituted, the plasmids were transformed into *E. coli* DH5 α and BL21 (DE3) cells for protein expression.

4.2.1.3 Purification of mesophilic kinases

The mesophilic kinases, Cmk and Gmk, were purified from large scale cultures as described in section 3.2.1 using the ZYM-5052 auto-induction media. The kinases were purified from large scale cultures by immobilised nickel affinity and size exclusion chromatography using an ÄKTA pure system (GE Healthcare). The protocol was as described in section 3.2.2.1. However, the fractions from the nickel affinity column were manually collected and injected onto the size exclusion column. This also enabled fractions from the nickel column to be assessed by SDS-PAGE. The purified proteins were quality controlled, concentrated and stored as described in section 3.2.2.3.

4.2.1.4 Purification of thermophilic kinases

Three phases were required for purification of the thermophilic kinases, Cmk, Gmk, Tmk and Umpk. After large scale culture in LB with IPTG induction, the cells were harvested and the protein extracted as previously described, resuspending the cell pellets in a buffer of 20 mM Tris-HCl pH 8.0. This enabled an initial purification by ion exchange. Either anion or cation exchange was performed using 5 mL HiTrap Q FF or 5 mL HiTrap SP FF columns (GE Healthcare) respectively on an ÄKTA pure system. For either ion exchange method, the salt concentration was increased stepwise in four equal increments from 0 mM to 100 mM NaCl, collecting 25 mL for each step. Fractions and flow through were collected from each and analysed using SDS-PAGE. The fraction or flow through containing the protein was then loaded onto a nickel column for immobilised metal affinity exchange. Once loaded, the column was washed with 20 mM Tris-HCl pH 8.0, 100 mM NaCl. To elute, imidazole was added in increments (0, 25 and 250 mM), with 20 mM Tris-HCl pH 8.0, 100 mM NaCl

remaining constant. Finally, collected fractions were analysed by SDS-PAGE, purified with size exclusion chromatography and stored as previously described (sections 3.2.2.1 and 3.2.2.3).

4.2.2 Kinase activity assay

Reactions were carried out in triplicate in 20 μL volumes in U-bottomed 96-well PCR plates (Thermo Fisher). The appropriate nucleotide (UDP, CDP, GDP or TDP) was diluted from 0.6 to 0 mM in reaction buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl_2 , 10 mM MnCl_2) supplemented with 0.1 mM ADP. The reactions were started by adding kinase to a final concentration of 0.2 mg mL^{-1} . The reactions were incubated at either 37 $^{\circ}\text{C}$ (mesophilic kinases) or 65 $^{\circ}\text{C}$ (thermostable kinases) for 10 min then cooled to 4 $^{\circ}\text{C}$. Next, a 15 μL aliquot was removed and transferred to a white 384-well plate (Corning) containing 15 μL BacTitre Glo™ (Promega) and thoroughly mixed. The plate was covered with an adhesive foil (Thermo Fisher) and kept at room temperature for 20 min. Luminescence was read using an Infinite M200 microplate reader (Tecan) (attenuation automatic; integration time 500 ms; settle time 150 ms). An ATP standard curve was run with 0 to 0.6 mM ATP in reaction buffer.

4.2.3 Characterisation of β -galactosyltransferase and LgtC

4.2.3.1 Coupled discontinuous kinase assay

Reactions were carried out in triplicate in 20 μL volumes in U-bottomed 96-well PCR plates. The donor substrate, UDP-Gal, was diluted from 6 to 0 mM in reaction buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl_2 , 10 mM MnCl_2) containing 0.1 mM ADP, 0.25 mU mL^{-1} β 14GalT and 0.2 mg mL^{-1} kinase. The reactions were started by adding GlcNAc to a final concentration of 5 mM or

15 mM. The reactions were incubated at 37 °C for 10 min (or, if indicated, 15 min) then cooled to 4 °C. BacTitre-Glo™ was added, and the resulting luminescence measured as described above.

The results were interpolated from a standard Michaelis-Menten curve to provide an output of UDP generated per minute. A range of UDP (0-0.6 mM) was added to 0.1 mM ADP and 0.2 mg/mL mesophilic Cmk which was incubated, detected with BacTiter Glo™ and measured simultaneously with the coupled assay.

4.2.3.2 Coupled discontinuous phosphatase assay

The coupled phosphatase assay was performed using the Glycosyltransferase Activity Kit (R&D systems; Wu *et al.*, 2011). Reactions were carried out in triplicate in 50 µL volumes in flat 96-well microplates (Greiner). GlcNAc, UDP-Gal and β14GalT were diluted in assay buffer from the kit. UDP-Gal was diluted from 6 to 0 mM in assay buffer containing 2 µg mL⁻¹ coupling phosphatase and either 5 mM or 15 mM GlcNAc. The reaction was started by adding β14GalT to a final concentration of 0.25 mU mL⁻¹. The plate was covered with an adhesive foil (Thermo Fisher) and incubated at 37 °C for 10 min (or, if indicated, 15 min). To stop the reaction, 30 µL Malachite Green Reagent A was added, followed by 100 µL dH₂O and 30 µL Malachite Green Reagent B. The plate was incubated at room temperature for 20 min then absorbance was measured at 620 nm using an Infinite M200 microplate reader (Tecan).

4.2.3.3 LgtC characterisation

Both the kinase and phosphatase assays were used to characterise LgtC using the methods described. LgtC was kindly gifted by Dr Vivoli (University of Exeter), prepared as described in Xu *et al.*, (under review). Instead of β14GalT and

GlcNAc, $7 \mu\text{g mL}^{-1}$ LgtC was used with 2 mM lactose and 2 mM dithiothreitol (DTT).

4.2.5 Data Analysis

All data were analysed on Prism® GraphPad (v. 5.0) software using the Michaelis-Menten equation to determine K_M and V_{max} . Where appropriate, substrate inhibition (K_i) was also examined. For the glycosyltransferases, results were interpolated from a Michaelis-Menten standard curve (0 – 0.6 mM NDP + 0.1 mM ADP) using the appropriate kinase to convert luminescence to production of free NPD. For the comparative phosphatase assay, results were interpolated from a linear standard curve of phosphate (0 – 5 μM).

4.2.5.1 Michaelis-Menten and substrate inhibition analyses

To determine the efficacy of the novel assay it was compared with a commercially-available kit. Two important kinetic parameters, V_{max} and K_M , were determined for β 14GalT and LgtC using the two assays and the results compared. The Michaelis-Menten equation was used to determine V_{max} and K_M , where v = velocity, $[S]$ = substrate concentration and K_M is the equation constant:

$$v = \frac{V_{max}[S]}{K_M + [S]}$$

The maximum rate or velocity of the reaction (V_{max}) is used to determine substrate turnover (K_{cat}) if the concentration of enzyme active sites is known. The Michaelis-Menten constant, K_M , is the substrate concentration required to reach 50 %

maximum velocity. It denotes the affinity of an enzyme for a specific substrate. Therefore, a higher K_M denotes a lower affinity.

Another important measurement is the substrate inhibition equation where v = velocity, $[S]$ = substrate concentration, K_M is the equation constant and K_i is the dissociation constant:

$$v = \frac{V_{max} [S]}{(K_M + [S])(1 + [S]/K_i)}$$

Substrate inhibition can occur when the substrate concentration is sufficient to cause two molecules to bind to the enzyme and block the reaction from occurring. The equation is used on the same data set as the Michaelis-Menten and provides an additional parameter; the dissociation constant, which is the substrate concentration required for two molecules to bind the enzyme.

4.2.6 Fluorescent antibody assay

A fluorescent antibody assay kit, Transcreener® ADP Fluorescence Polarization Assay (BellBrook Labs), was trialled as a control for the developing coupled kinase assay. The kit was used following the manufacturer's instructions. Enzyme titration of β 14GalT was attempted: briefly, the following conditions were used: 0 – 0.25 mU mL⁻¹ β 14GalT, 20 mM UDP-Gal, 50 mM GlcNAc in reaction buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂). After incubation at 37 °C for 15 min, the “stop and detect” buffer was added, containing 8 nM tracer and 1.2 mg mL⁻¹ fluorescent antibody. This was incubated for 1 hr at 37 °C and read with optimised gain, excitation 580 nm and emission 620 nm.

Controls were UDP-Gal; UDP standard curve (e.g. 0 % UDP-Gal, 100 % UDP to 100 % UDP-Gal, 0 % UDP); and enzyme titration without UDP-Gal.

4.2.7 Apyrase purification

The gene for apyrase from *C. lectularius* was synthesised (Appendix 2) with overhangs for LIC into pNIC28-Bsa4 plasmid (Stavitsky *et al.*, 2010). For purification, the protocol from Valenzuela *et al.*, 1998 was followed.

4.2.7.1 Extraction from *E. coli*

The cloned gene was transformed into *E. coli* Rosetta (DE3) cells and purified from 500 mL ZYM-5052 broth cultures as described in section 3.2.1.2. with the following amendments; after harvesting the cells, the pellet was resuspended in 20 mM Tris-HCl, 500 mM NaCl, 1 mM EDTA, 1 % (v/v) Triton X-100, 10 mM DTT, pH 8.0 before sonication. After sonication, the lysate was centrifuged as normal and the pellet air dried at room temperature.

4.2.7.2 Denaturation

The pellet was resuspended in 100 mL denaturation buffer (100 mL Tris-HCl, 6 M Gu-HCl, 2 mM EDTA, 10 mM DTT, pH 8.2) and incubated at 60 °C for 10 min. After cooling to room temperature, the denatured extract was again centrifuged at 24, 000 xg for 30 min at 4 °C. The protein concentration of the supernatant was determined by Nanodrop (Thermo Fisher) and kept for dialysis.

4.2.7.3 Dialysis

For dialysis, the denatured proteins were diluted to a final concentration of 0.6 mg mL⁻¹ in 200 mL internal dialysis buffer (6 M Gu-HCl, 100 mM Tris-HCl,

pH 8.2). This was placed inside a dialysis membrane which was securely sealed. The resultant dialysis bag was placed inside a beaker containing 4 L external dialysis buffer (50 mM Tris-HCl, 250 mM NaCl, 2 mM CaCl₂, 5 % (v/v) glycerol, pH 8.0). The dialysis was run for 48 h at 4 °C with slow stirring. The external dialysis buffer was refreshed twice during the 48 h period.

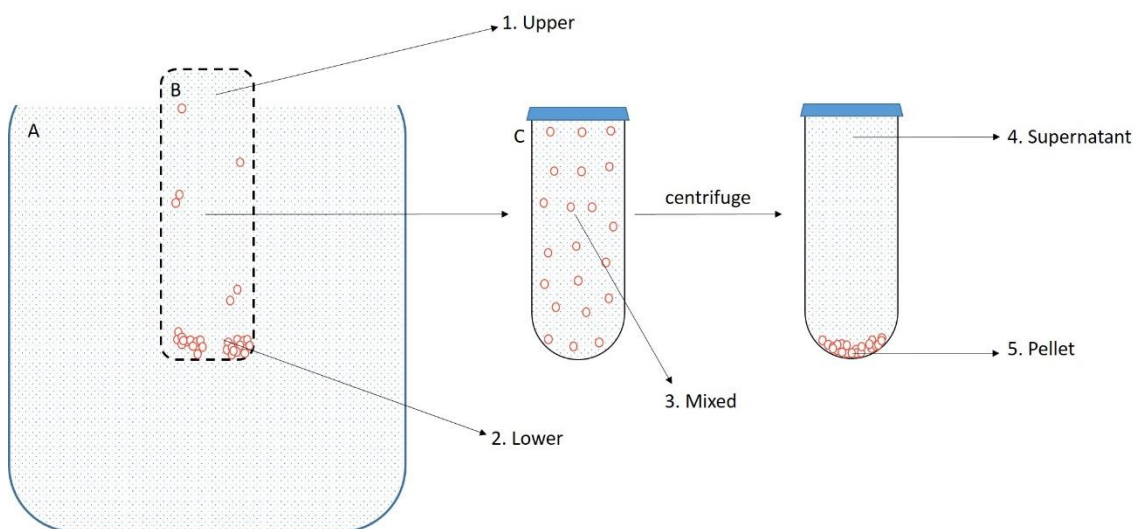


Figure 64: Dialysis of apyrase from *Cimex lectularius*. After 48 h at 4 °C, the solutes balance between the buffer (A) and sample in the dialysis bag (B) by osmosis across the membrane (dotted black line). Some visible precipitate formed (red circles). Aliquots were taken from the upper (1) and lower (2) portions of the dialysis bag before the mixture was decanted into a centrifuge tube (C) and mixed. An aliquot of mixed sample was taken before centrifugation. After centrifugation aliquots of the resultant supernatant (4) and pellet (5) were taken.

To analyse the dialysis, 20 µL aliquots were taken from the top and bottom of the bag. The dialysed sample was then decanted and centrifuged at 5,000 xg for 15 min at 4 °C. A 20 µL aliquot of the supernatant was taken. A small amount of pellet was resuspended in 1 mL 20 mM Tris-HCl, 100 mM NaCl, pH 8.0, from which a 20 µL aliquot was removed (Figure 64). These aliquots were analysed by SDS-PAGE as described (section 2.2.4.7). To each lane, 20 µL prepared sample was loaded.

4.2.7.4 TCA precipitation

As each of the aliquots in section 4.2.7.3 were taken for SDS-PAGE (Figure 64), a second 80 μL aliquot was also taken for TCA precipitation. To each aliquot, 20 μL TCA was added. The samples were incubated at 4 $^{\circ}\text{C}$ for 10 min. These were then centrifuged at 14,000 $\times g$ for 5 min at 4 $^{\circ}\text{C}$. The supernatant was removed, and the pellet washed with 200 μL cold acetone. The samples were centrifuged again under the same conditions. The acetone wash and centrifugation was repeated two more times. After the final centrifugation, the pellet was dried at 95 $^{\circ}\text{C}$ for 10 min, ensuring all the acetone had evaporated. To the remaining pellet, 20 μL SDS-PAGE loading buffer was added and the gel run as previously described (section 2.2.4.7). To each lane, 20 μL prepared sample was loaded.

4.2.7.5 Nickel affinity purification

The dialysis of extracted apyrase was repeated as previously described. After 48 h the dialysed protein was centrifuged at 24,000 $\times g$ for 30 min at 4 $^{\circ}\text{C}$. The supernatant was removed and the pellet discarded. The supernatant was incubated overnight at 4 $^{\circ}\text{C}$ with 2 mL Ni Sepharose High Performance bead slurry (GE Healthcare) and gentle rocking. After incubation, the beads were drained and 30 mL wash buffer (20 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole, pH 8.0) was added. The beads were drained and 12 mL elution buffer (20 mM Tris-HCl, 500 mM NaCl, 250 mM imidazole, pH 8.0) was added. The beads were drained and the eluate immediately diluted to 60 mL with wash buffer. The diluted protein was concentrated and stored as previously described (section 3.2.2.3), with the addition of 20 mM CaCl_2 to the final stock. Aliquots taken from the flow through, wash and elution of the chromatography, plus the concentrated protein, were analysed by SDS-PAGE (section 2.2.4.7). To each lane, 20 μL

sample treated with loading buffer was added (5 μL and 40 μL for the concentrated protein).

4.3 Results

4.3.1 Kinase purification

4.3.1.1 Mesophilic kinases

Purification of the mesophilic kinases from *E. coli* was achieved by extracting the proteins from large scale culture and performing nickel affinity exchange (Figure 65; Figure 66) followed by size exclusion chromatography (Figure 67) on an ÄKTA pure system.

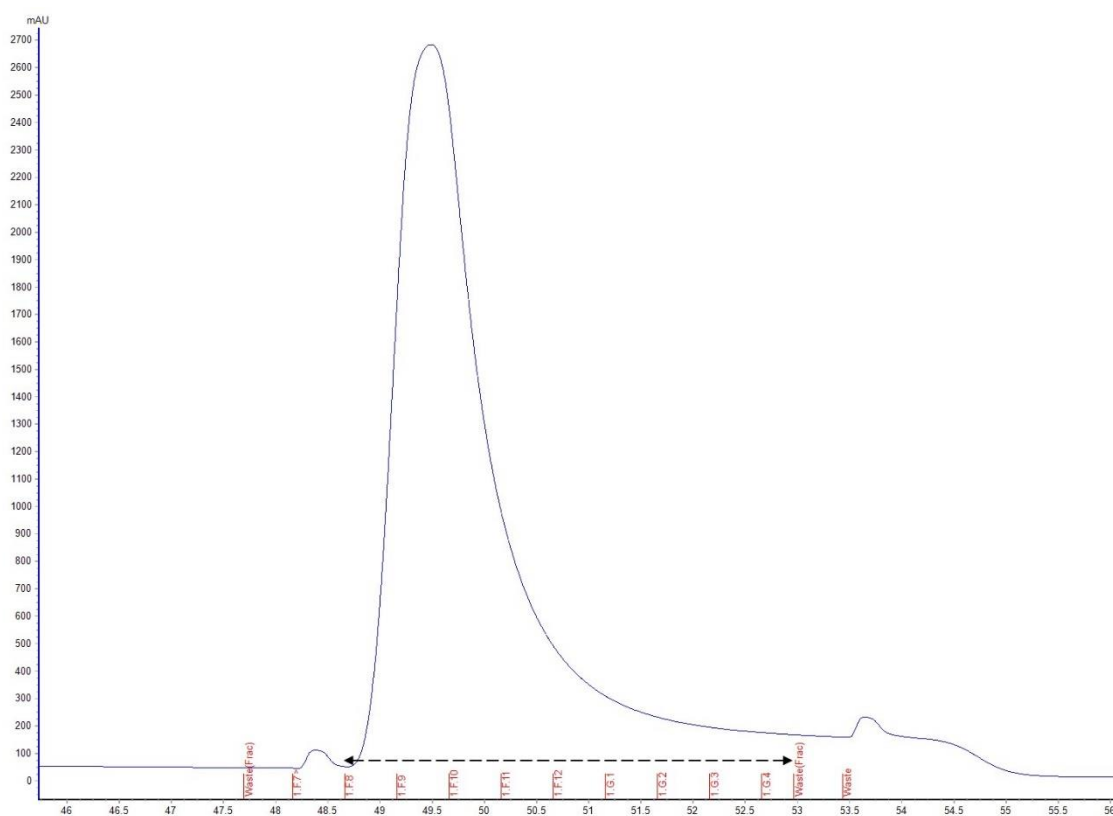


Figure 65: Elution from nickel affinity purification of mesophilic kinase, Cmk expressed in *E. coli* Rosetta (DE3) cells. Protein extract was loaded onto a 1 mL His-Trap FF column using an ÄKTA pure system. Bound protein was washed with 20 mM Tris-HCl, 100 mM NaCl, 10 mM imidazole pH 8.0 and eluted in 20 mM Tris-HCl, 100 mM NaCl, 250 mM imidazole pH 8.0. Elution begins at 1.F.7. Chromatogram shows absorbance 280 nm (mAu) (blue), fractions collected (red) and fractions analysed by SDS-PAGE (Figure 666) (black dashed arrow).

Fractions were collected after nickel affinity purification and analysed by SDS-PAGE for purity and concentration of the protein. Aliquots of the extracted protein,

flow through and wash were also run on the gel for comparison. Figure 66 shows that mesophilic Cmk is greatly over-expressed in *E. coli* Rosetta cells and nickel affinity purification removes many of the contaminating proteins. The relative weakness of the bands in the 'neat' fraction is due to the small amount loaded onto the gel (2 μ L) compared with 5 μ L for the flow through and wash fractions and 10 μ L for the eluted fractions. There is a strong band present in the 'wash' fractions which could indicate over-loading of the column but is more likely to be due to displacement of some Cmk by the 10 mM imidazole present in the buffer.

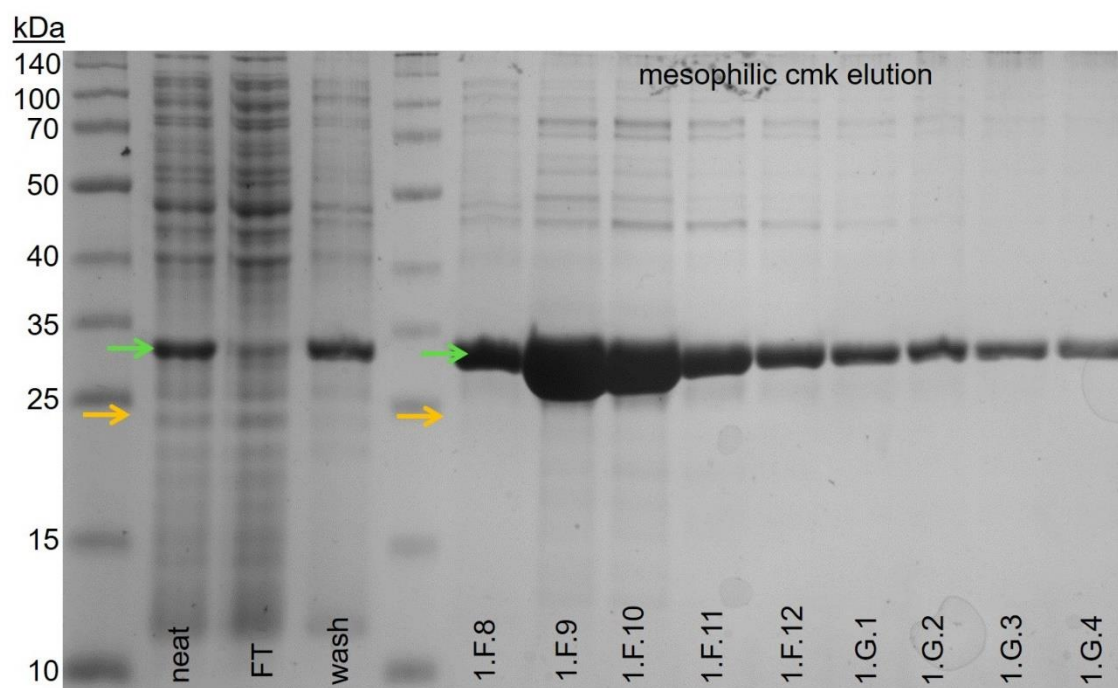


Figure 66: SDS-PAGE of nickel affinity purification of mesophilic Cmk. Neat = protein extract; FT = flow through from the nickel affinity column; wash = buffer after washing unbound protein from the column; elution = eluted fractions containing mesophilic Cmk. Orange arrows indicate expected size (24.7 kDa); green arrows indicate actual size. Each lane was loaded with 10 μ L of sample (except 2 μ L of neat and 5 μ L of FT) and the gel has 5 μ L protein ladder of defined size (kDa).

Fractions containing expressed protein (1.F.8 to 1.G.4) were pooled and applied to a size exclusion column for further purification (Figure 67).

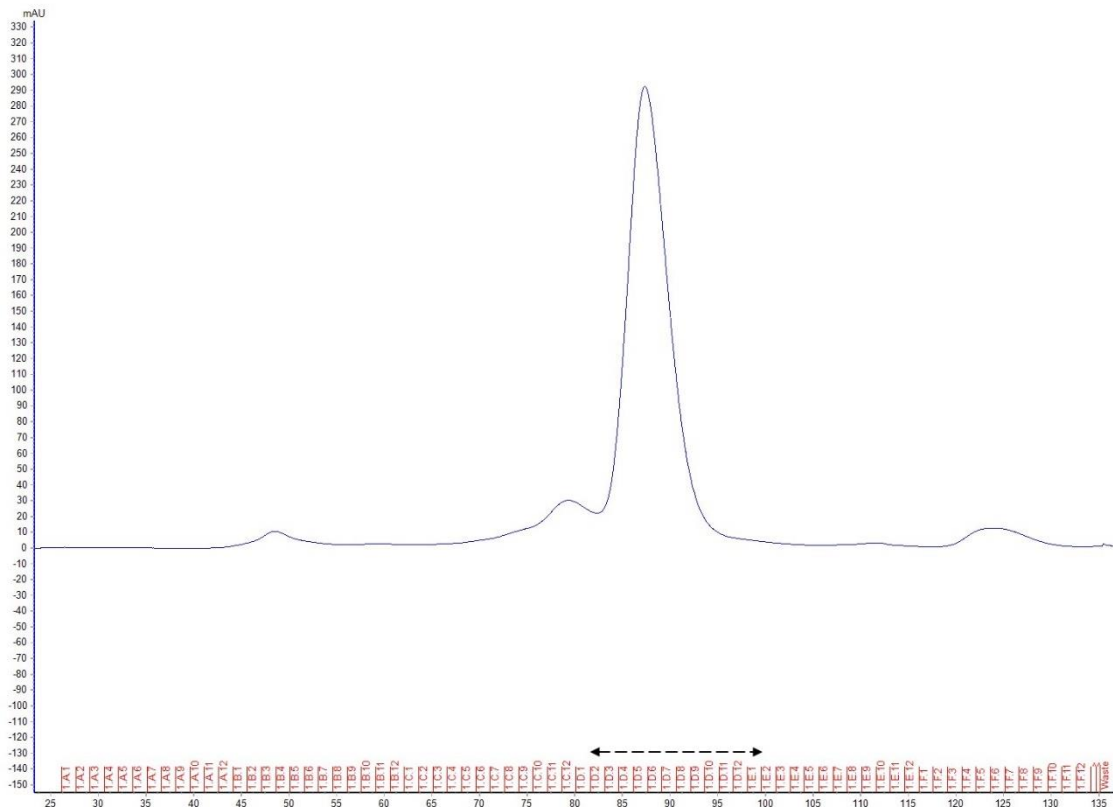


Figure 67: Elution from size exclusion purification of mesophilic kinase, Cmk, expressed in *E. coli* Rosetta (DE3) cells. Fractions of protein after nickel affinity purification were pooled and loaded onto a Superdex 200 16/600 column using an ÄKTA pure system. Isocratic elution in 10 mM HEPES pH 7.0, 100 mM NaCl. Chromatogram shows absorbance 280 nm (mAu) (blue), fractions collected (red) and fractions analysed by SDS-PAGE (Figure 68) (black dashed arrow).

Fractions collected from this purification step were also analysed by SDS-PAGE. The SDS-PAGE in Figure 68 demonstrates that Cmk (and Gmk; not shown) were prone to cleavage by proteases (revealed as double bands in the gel). It was therefore essential to include an incubation step with protease inhibitors during extraction from the cells.

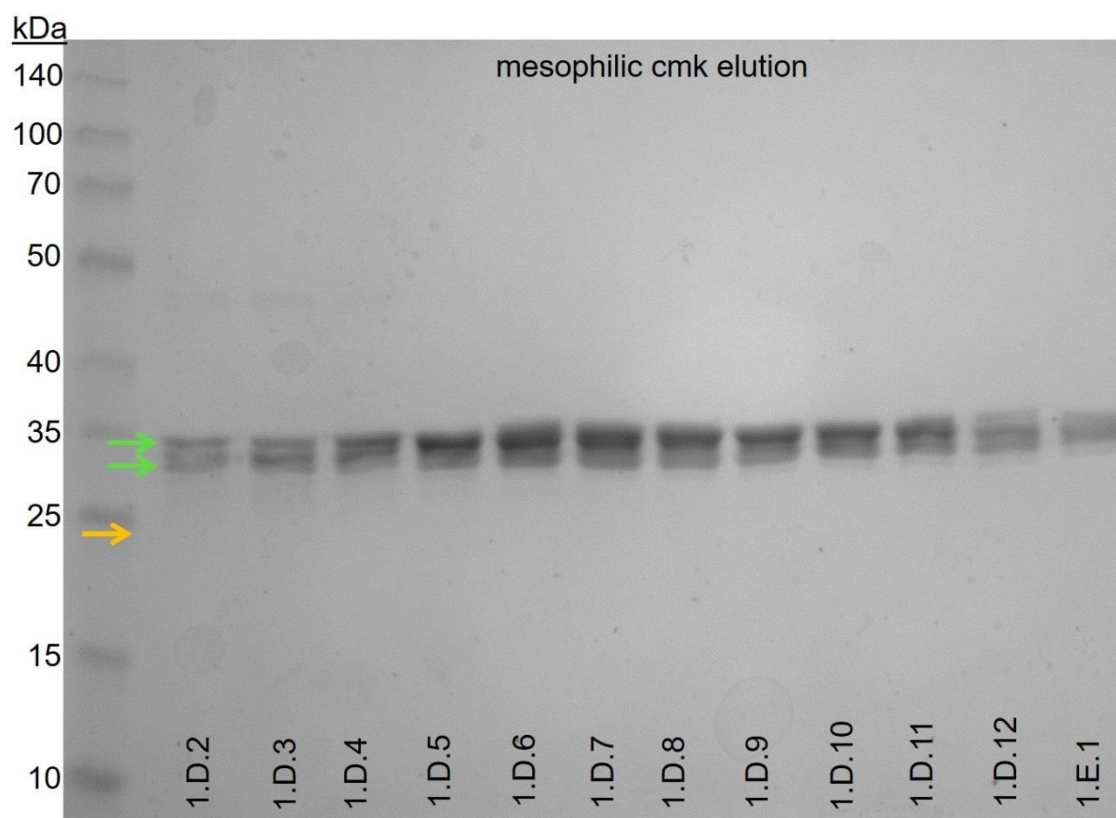


Figure 68: SDS-PAGE showing denaturation of mesophilic Cmk in fractions eluted from size exclusion chromatography without protease inhibitors added during extraction from *E. coli* Rosetta (DE3) cells. Orange arrow indicates expected size (24.7 kDa); green arrows indicate actual size. Each lane was loaded with 10 μ L of sample and the gel has 5 μ L protein ladder of defined size (kDa).

Figure 69 shows that inclusion of protease inhibitors solved the problem of denaturation. It also demonstrates that size exclusion successfully removes the majority of the contaminating proteins and the mesophilic Cmk remains at a high concentration. The bands ran a little high (~30 kDa) compared to the expected size of 24.7 kDa (plus 1 kDa His-tag), though this is consistent between gels from nickel affinity and repeats of size exclusion chromatography (Figure 66, Figure 68, Figure 69). Activity assays were the best method to test whether the purified protein was correct despite the discrepancy in the band size (sections 4.2.2 and 4.3.4). Once concentrated with 20 % glycerol into a 1.5 mL final volume, typical stock concentrations of both mesophilic kinases were approximately 1 mg mL⁻¹.

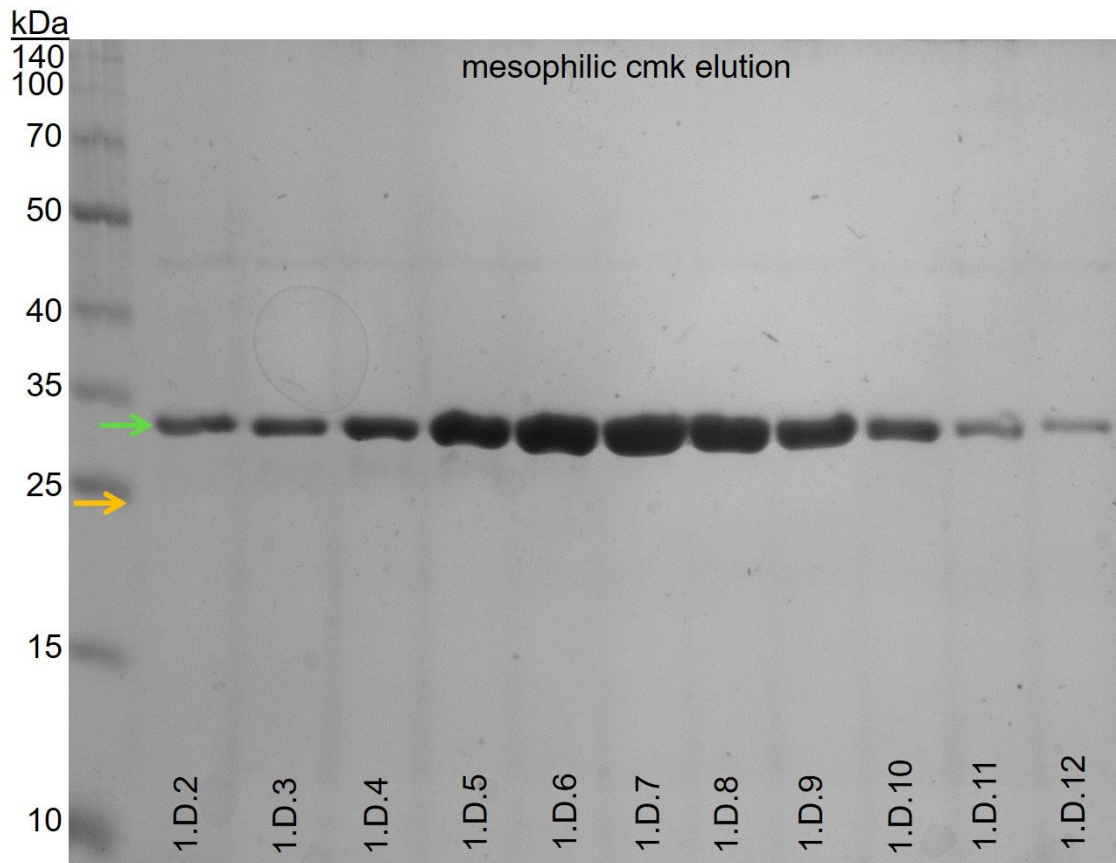


Figure 69: SDS-PAGE showing fractions eluted from size exclusion chromatography containing purified mesophilic Cmk. Extraction of the protein from *E. coli* Rosetta (DE3) cells included incubation with protease inhibitors to prevent degradation. Orange arrow indicates expected size (24.7 kDa); green arrow indicates actual size. Each lane was loaded with 10 μ L of sample and the gel has 5 μ L protein ladder of defined size (kDa).

4.3.1.2 Thermophilic kinases

Purification of thermophilic kinases from *E. coli* was more complex than the mesophilic kinases. After extraction from large scale culture and nickel affinity exchange on the ÄKTA pure system, the chromatogram for size exclusion displayed several poorly defined peaks and low absorbance at 280 nm (Figure 70).

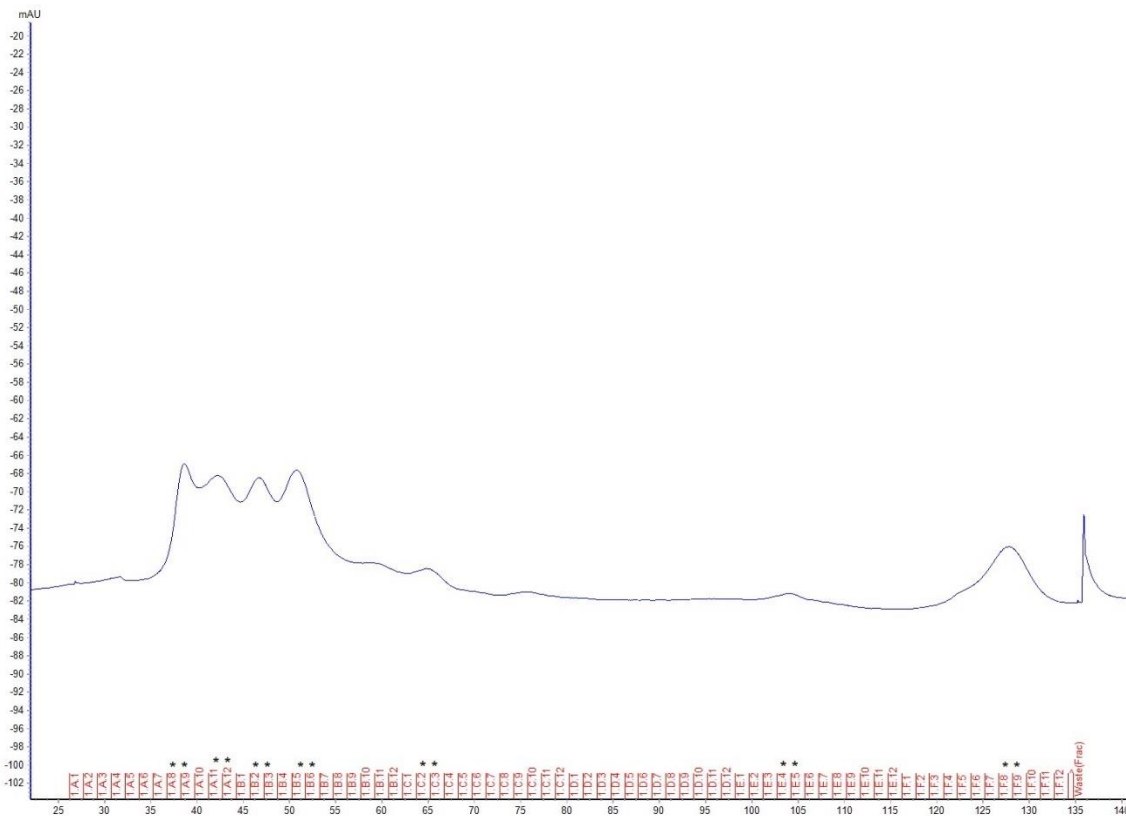


Figure 70: Elution from size exclusion purification of thermophilic kinase, Umpk, expressed in *E. coli* BL21 (DE3) cells. Fractions of protein after nickel affinity purification were pooled and loaded onto a Superdex 200 16/600 column using an ÄKTA pure system. Isocratic elution in 10 mM HEPES pH 7.0, 100 mM NaCl. Chromatogram shows absorbance 280 nm (mAu) (blue), fractions collected (red), fractions analysed by SDS-PAGE (Figure 71) (black asterisk).

Despite the negative absorbance units seen in the size exclusion chromatogram (Figure 70), SDS-PAGE analysis confirms expression of Umpk (Figure 71). The low UV absorbance may have been the result of interference by a contaminant. The gel in Figure 71 also shows several contaminating bands, and multiple bands of the protein of interest, suggesting contamination of the sample and degradation of Umpk. These results were consistent with all the thermostable kinases, though demonstrably worse for Gmk and Umpk than for Cmk and Tmk.

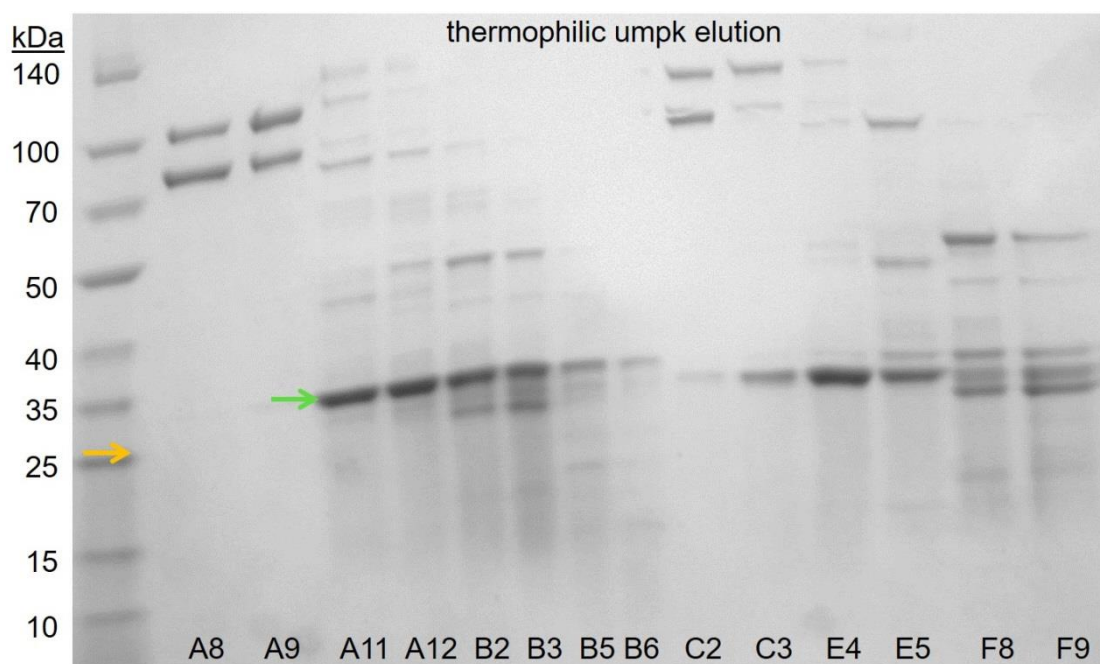


Figure 71: SDS-PAGE of thermophilic kinase Umpk extracted from *E. coli* BL21 (DE3) after nickel affinity and size exclusion chromatography. Fractions collected from size exclusion (Figure 70). Orange arrow indicates expected size (25.3 kDa); green arrow indicates actual size. Each lane was loaded with 10 μ L of sample and the gel has 5 μ L protein ladder of defined size (kDa).

To remedy this, an additional purification step, ion exchange, was included. The thermostable kinases were extracted from *E. coli* cells into buffer with no salt at pH 8.0. Anion exchange (Q column) was used for Cmk (pI 6.86), Gmk (pI 6.88) and Umpk (pI 6.76) and cation exchange (SP column) for Tmk (pI 9.67). Figure 72 shows the chromatogram from anion exchange for Umpk. The flow through was also collected.

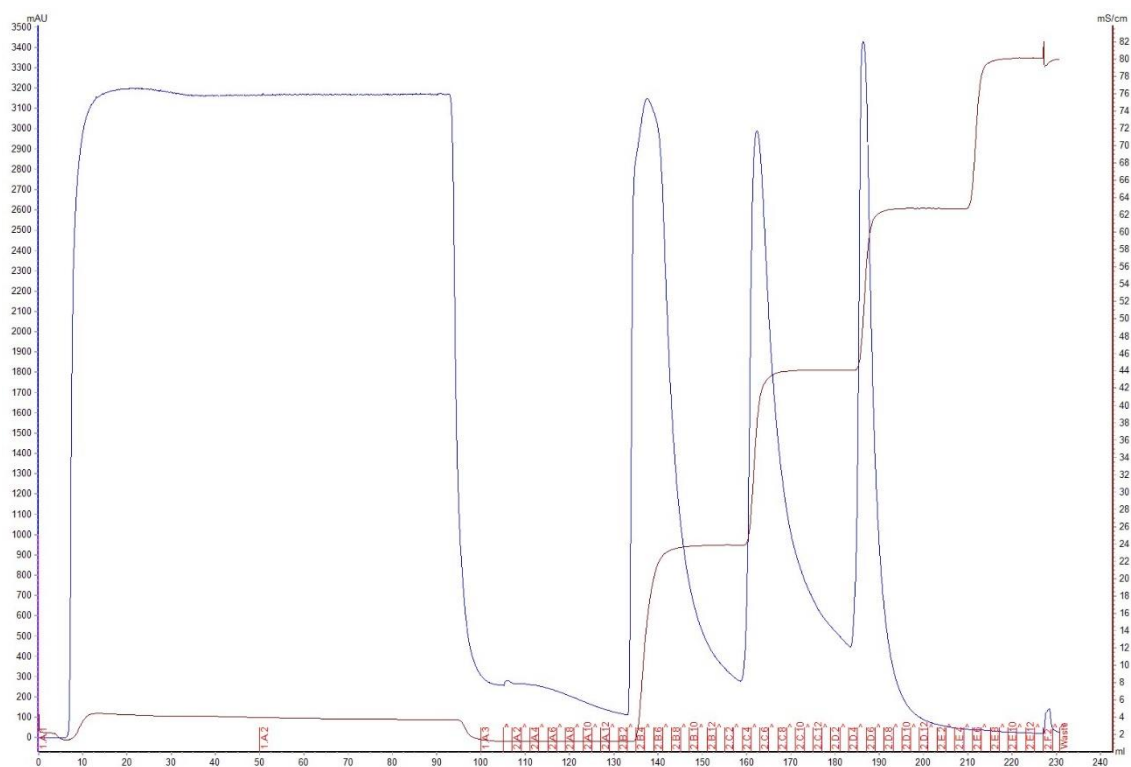


Figure 72: Elution from anion exchange of thermophilic kinase, Umpk expressed in *E. coli* BL21 (DE3) cells. Protein extract was loaded onto a 5 mL Q FF column using an ÄKTA pure system in 20 mM Tris-HCl pH 8.0. Flow through was collected in fractions 1.A.1 and 1.A.2. A gradient of NaCl from 0 to 100 mM was applied in four equal increments and applied to the column for 25 mL, beginning at 130 mL. Chromatogram shows absorbance 280 nm (mAu) (blue), fractions collected (red) and conductivity (brown).

Fractions from the flow through and each peak in Figure 72 were analysed by SDS-PAGE [not shown]. The purest fractions were pooled and run on a nickel affinity column with a stepwise gradient of imidazole (0, 25 and 250 mM) (Figure 73), collecting 4 mL for each step.

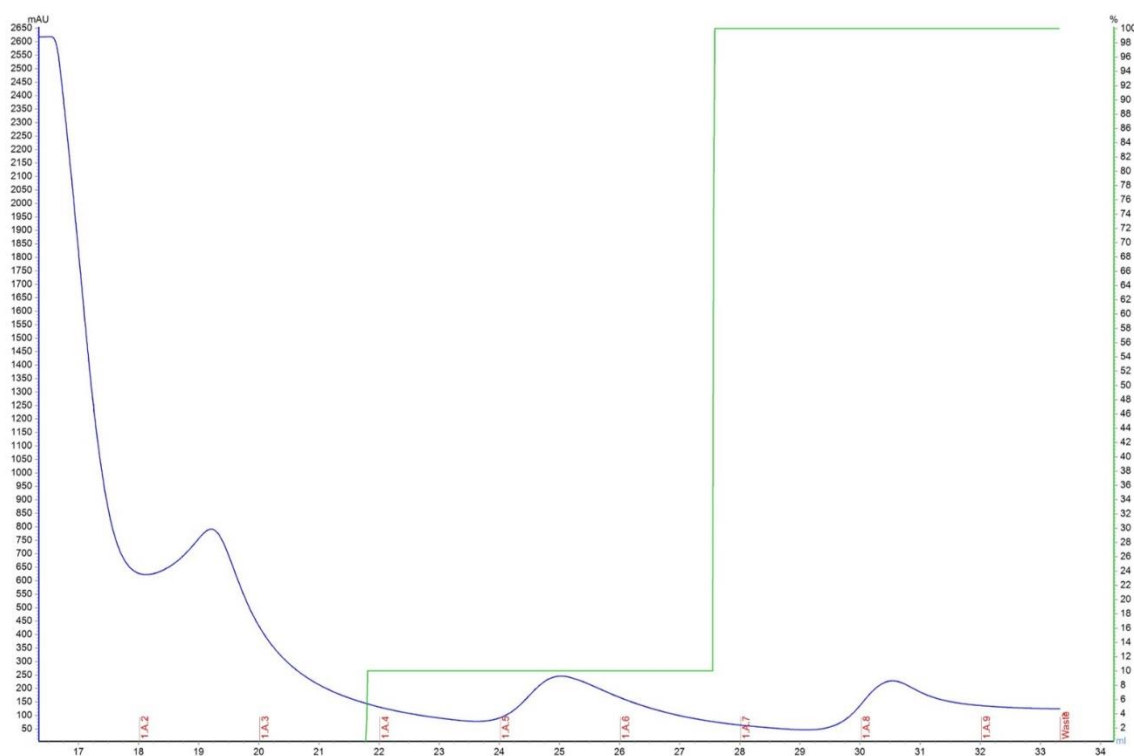


Figure 73: Elution from nickel affinity purification of thermophilic kinase, Umpk expressed in *E. coli* BL21 (DE3) cells. After anion exchange, pooled fractions were loaded onto a 1 mL His-Trap FF column using an ÄKTA pure system. Bound protein was eluted in 20 mM Tris-HCl, 100 mM NaCl with 0 mM, 25 mM and 250 mM imidazole pH 8.0. Chromatogram shows absorbance 280 nm (mAu) (blue), fractions collected (red) and buffer B concentration (green). Buffer B = 20 mM Tris-HCl, 100 mM NaCl, 250 mM imidazole, pH 8.0.

All fractions, 1.A.1 to 1.A.9, collected in the step-gradient nickel affinity purification (Figure 73) were analysed by SDS-PAGE for purity [not shown]. For all four thermostable kinases, the protein eluted in the peak with 250 mM imidazole (1.A.9 in Figure 73). This fraction was loaded onto a size exclusion column for the final step in the purification process (Figure 74).

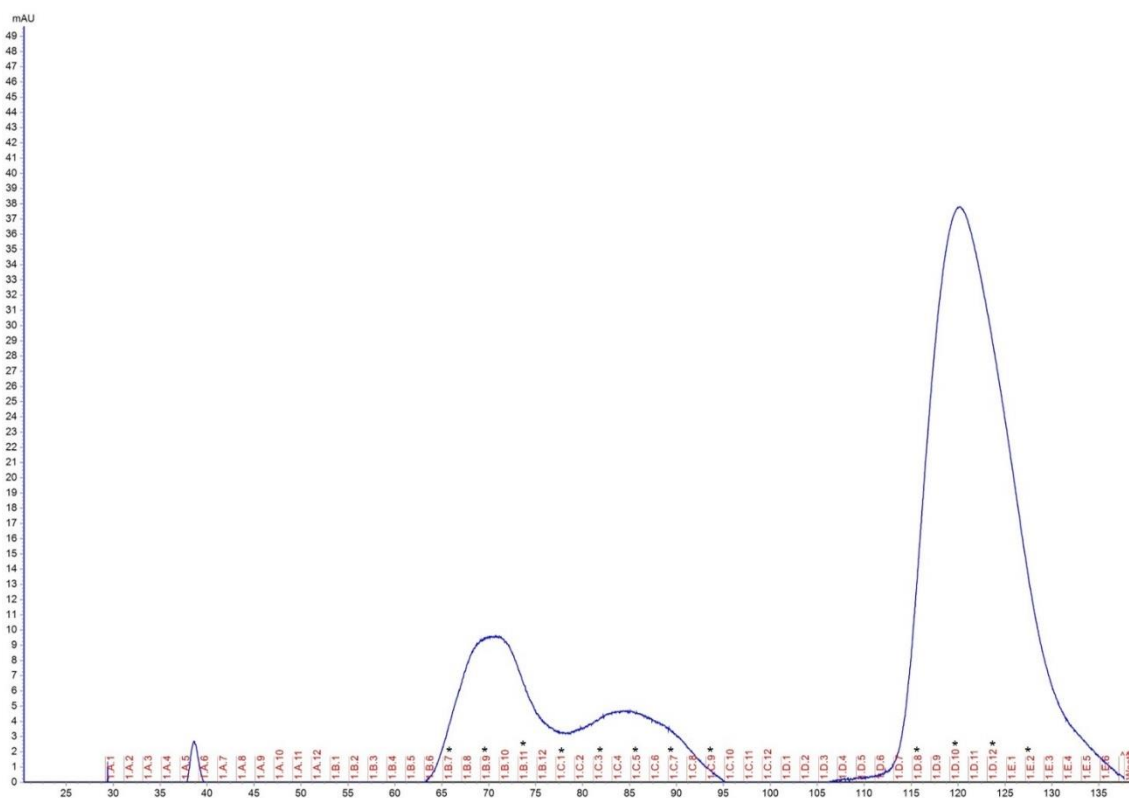


Figure 74: Elution from size exclusion purification of thermophilic kinase, Umpk, expressed in *E. coli* BL21 (DE3) cells after anion exchange and step gradient nickel affinity chromatography. Fractions of protein after nickel affinity purification were pooled and loaded onto a Superdex 200 16/600 column using an ÄKTA pure system. Isocratic elution in 10 mM HEPES pH 7.0, 100 mM NaCl. Chromatogram shows absorbance 280 nm (mAu) (blue), fractions collected (red), fractions analysed by SDS-PAGE (Figure 75) (black asterisk).

Finally, fractions from the size exclusion column (Figure 74) were analysed by SDS-PAGE (Figure 75). Fractions C3, C4, C5 and C6 were pooled and concentrated before adding glycerol, measuring protein concentration and storing at $-80\text{ }^{\circ}\text{C}$. The thermostable kinases did not purify express as much or purify as well as the mesophilic kinases and there was typically four $250\text{ }\mu\text{L}$ aliquots of approximately 0.75 mg mL^{-1} protein after each round of purification from a 500 mL culture of *E. coli*.

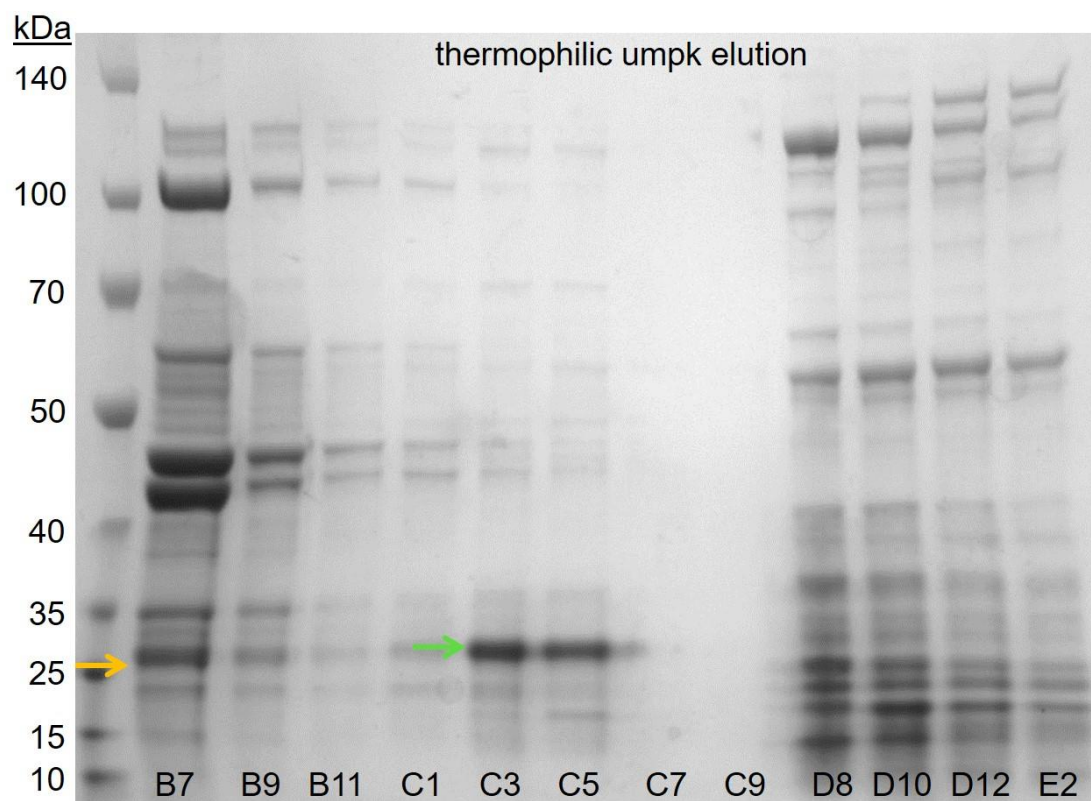


Figure 75: SDS-PAGE of thermophilic kinase Umpk extracted from *E. coli* BL21 (DE3) after ion exchange, step-gradient nickel affinity and size exclusion chromatography. Fractions collected from size exclusion (Figure 74). Orange arrow indicates expected size (25.3 kDa); green arrow indicates actual size. Each lane was loaded with 10 μ L of sample and the gel has 5 μ L protein ladder of defined size (kDa).

Despite the addition of the ion exchange step, there was still contamination of thermostable Umpk evident in the SDS-PAGE gel after size exclusion chromatography (Figure 75). However, in comparison to the gel after size exclusion chromatography without the additional ion exchange (Figure 71), there are fewer contaminants, at lower concentration and with no multiple banding of the protein of interest.

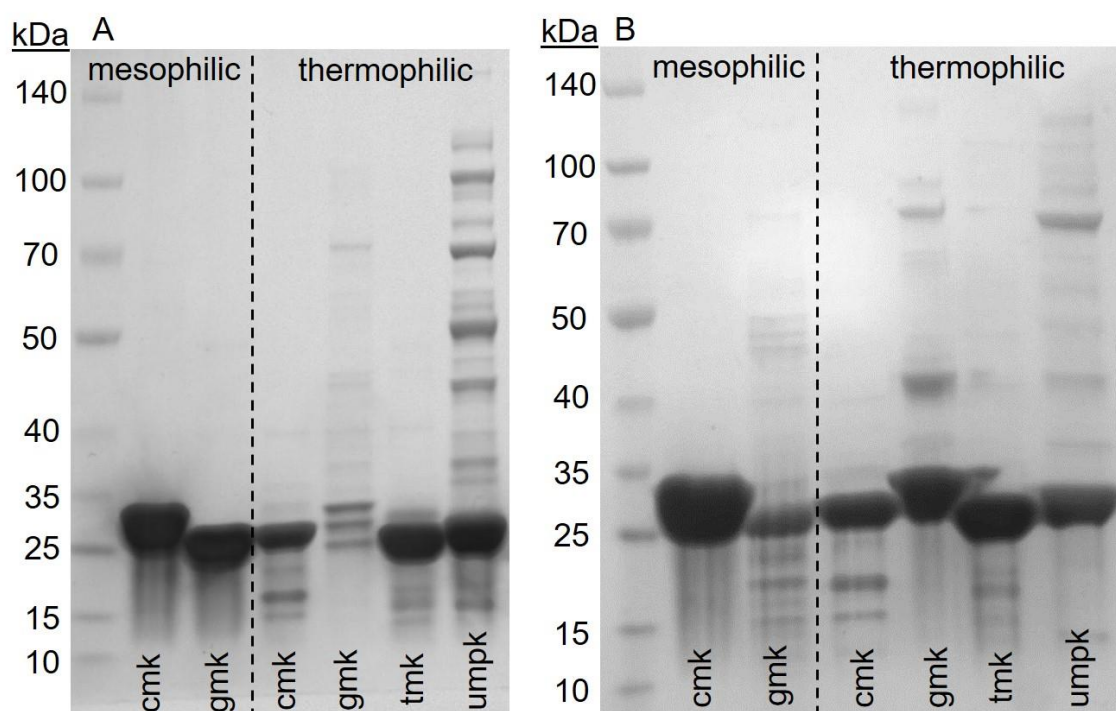


Figure 76: SDS-PAGE of purified mesophilic and thermostable kinases from *E. coli* Rosetta (DE3) and BL21 (DE3) cells. The same samples of mesophilic Cmk and Gmk are loaded on gel A then gel B after a freeze-thaw cycle at $-80\text{ }^{\circ}\text{C}$. The samples of thermophilic kinases were obtained without (gel A) and with (gel B) an ion exchange step. Each lane was loaded with $10\text{ }\mu\text{L}$ of sample and each gel has $5\text{ }\mu\text{L}$ protein ladder of defined size (kDa).

Figure 76 shows SDS-PAGE gels of the mesophilic and thermophilic kinases after purification, concentration and addition of glycerol. For the mesophilic kinases, the same sample was loaded onto each gel, with a freeze-thaw step at $-80\text{ }^{\circ}\text{C}$ in between. Whilst the mesophilic proteins purify very well, and Cmk remains stable, there is clearly degradation and aggregation of Gmk due to the freeze-thaw cycle as there are additional bands apparent in gel B that were absent in gel A. This suggested the kinases are susceptible to degradation and it was necessary to only thaw them once, keep them on ice and discard any unused sample. For the thermophilic kinases in Figure 76, the samples in gel A were purified without the ion exchange step and the samples in gel B were purified with the ion exchange. The ion exchange made no discernible difference to the purification of thermostable Cmk and Tmk. For thermostable Gmk, the ion

exchange prevented degradation of the protein as there are several bands visible in gel A but only one clean band in gel B. The Umpk enzyme had less contamination present when the ion exchange was included.

4.3.2 Kinase assay optimisation

4.3.2.1 ADP concentration

As mesophilic Cmk purified well and at high concentration it was used for development of the kinase assay. Firstly, the effect of ADP on an ATP standard curve detected with the BacTiter-Glo™ reagent was determined (Figure 77).

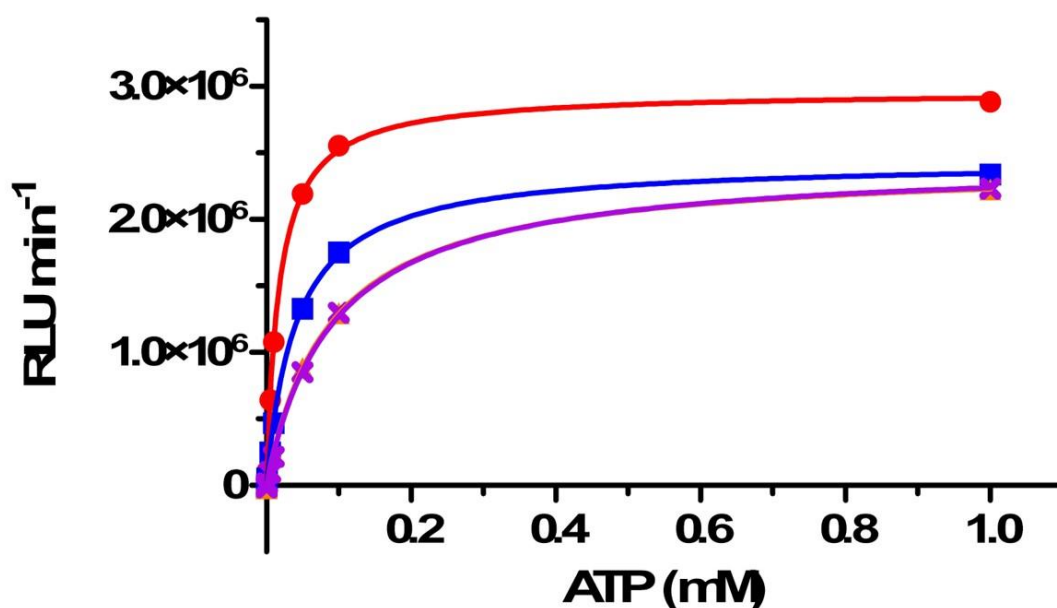


Figure 77: Standard curve of 0 - 1 mM ATP in reaction buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂) containing 0mM ADP (red circle), 0.1 mM ADP (blue square), 0.2mM ADP (orange triangle) and 0.5 mM ADP (purple cross). Tested with BacTiter-Glo™ reagent. Output in relative units of luminescence produced per minute. Michaelis-Menten analysed using GraphPad Prism® v. 5.0. Error bars indicate standard deviation from the mean; n = 2 biological replicates for 0 mM and 0.1 mM ADP and n = 3 biological replicates for 0.2 mM and 0.3 mM ADP.

Most enzymes display product inhibition as a general feature, and so a reduction in the efficiency of the luciferase enzyme in Bac-Titre Glo™ was expected.

Results in Table 14 show a reduction in reaction rate (V_{max}) and a reduction in substrate affinity (increase in K_M) in the presence of ADP. The reduction in V_{max} and increase was the same irrespective of the ADP concentration used, from 3.0×10^6 relative luminescence units (RLU) min^{-1} to 2.4×10^6 RLU min^{-1} . However, increasing the ADP concentration caused an increase in K_M . The K_M doubled in 0.1 mM ADP compared to no ADP and doubled again at 0.2 mM. For this reason, future assays contained 0.1 mM ADP for the reaction. The fit of the analyses were good ($R^2 > 0.99$ for all curves), with a standard error (SE) below 1 % for all V_{max} values and below 3 % for all K_M values.

Table 14: V_{max} and K_M for ATP standard curve with BacTiter-Glo™ reagent luciferase reaction in the presence of 0 mM, 0.1 mM, 0.2 mM and 0.5 mM ADP. Determined by Michaelis-Menten equation in GraphPad Prism® v.5.0. SE = standard error.

	0 mM ADP	0.1 mM ADP	0.2 mM ADP	0.5 mM ADP
V_{max} (RLU min^{-1})	3.0×10^6	2.4×10^6	2.4×10^6	2.4×10^6
SE: V_{max} (RLU min^{-1})	1.8×10^4	1.6×10^4	1.4×10^4	2.0×10^4
K_M (mM ATP)	0.0176	0.0414	0.0885	0.0920
SE: K_M (mM ATP)	0.0005	0.0009	0.0016	0.0024
R^2	0.9994	0.9995	0.9994	0.9988

4.3.2.2 Kinase concentration

Next, an optimal kinase concentration for turnover of UDP and ADP to UMP and ATP was determined using 0.1 mM ADP and 0 – 0.6 mM UDP, testing 0.05, 0.1, 0.2 and 0.3 mg mL^{-1} Cmk (Figure 78).

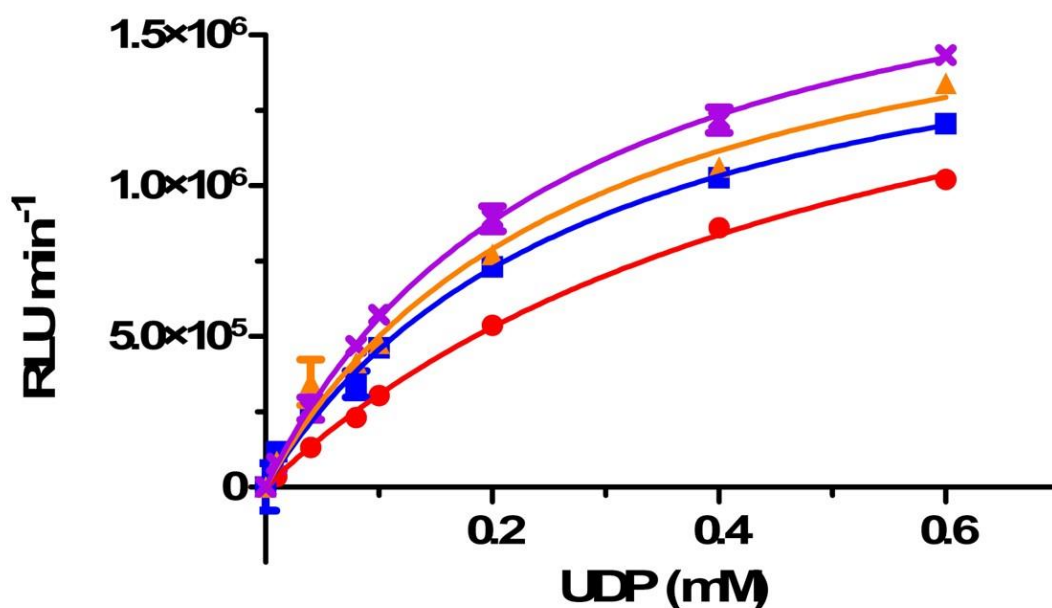


Figure 78: optimising Cmk concentration, with 0.1 mM ADP and 0-0.6 mM UDP in reaction buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂). The Cmk concentrations were 0.05 mg mL⁻¹ (red circle), 0.1 mg mL⁻¹ (blue square), 0.2 mg mL⁻¹ (orange triangle) and 0.3 mg mL⁻¹ (purple cross). Tested with BacTiter-Glo™ reagent. Output in relative units of luminescence produced per minute. Michaelis-Menten analysed using GraphPad Prism® v. 5.0. Error bars indicate standard deviation from the mean; n = 2 biological replicates.

Table 15 shows the concentration of Cmk had little impact on the apparent V_{max} for the coupled reaction (ranging from 1.8×10^6 to 2.1×10^6 RLU min⁻¹). However, the K_M reduced from 0.5486 mM UDP at 0.05 mg mL⁻¹ Cmk to 0.2913 mM UDP at 0.1 mg mL⁻¹ Cmk. A lower K_M is desirable, as this will increase the sensitivity of the reaction. Low concentrations of UDP are expected in glycosyltransferase reactions, so reducing K_M of the coupling reaction will increase the signal observed at these lower concentrations. This improvement continued, though to a lesser extent, for 0.2 mg mL⁻¹ and 0.3 mg mL⁻¹ Cmk (0.2808 mM and 0.2670 mM UDP respectively). To compromise between assay activity and enzyme expenditure, the concentration of Cmk (and other kinases) used in the glycosyltransferase characterisation assay was 0.2 mg mL⁻¹. The fit of the analyses were good ($R^2 > 0.98$ for all curves), with a SE below 7 % for all V_{max}

values. For K_M , the highest SE occurred for 0.2 mg mL⁻¹ Cmk, at 13.2 % ($K_M = 0.2808$ mM ATP; SE = 0.0370 mM ATP).

Table 15: V_{max} and K_M for various concentrations of Cmk with 0.1 mM ADP and 0 – 0.6 mM UDP. Measuring luciferase production with BacTiter-Glo™ reagent. Determined by Michaelis-Menten equation in GraphPad Prism® v.5.0. SE = standard error.

	0.05 mg mL⁻¹ Cmk	0.1 mg mL⁻¹ Cmk	0.2 mg mL⁻¹ Cmk	0.3 mg mL⁻¹ Cmk
V_{max} (RLU min⁻¹)	2.0 x 10 ⁶	1.8 x 10 ⁶	1.9 x 10 ⁶	2.1 x 10 ⁶
SE: V_{max} (RLU min⁻¹)	9.5 x 10 ⁴	8.5 x 10 ⁴	1.2 10 ⁵	4.5 x 10 ⁴
K_M (mM UDP)	0.5468	0.2913	0.2808	0.2670
SE: K_M (mM UDP)	0.0448	0.0293	0.0370	0.0128
R²	0.9970	0.9916	0.9853	0.9981

4.3.2.3 Assay range (NDP concentration)

Finally, the limits of detection of the assay were determined using 0.2 mg mL⁻¹ Cmk, 0.1 mM ADP and 0 to 5 mM UDP. As there was a clear reduction in the reaction rate, the assay was analysed with both the Michaelis-Menten and substrate inhibition equations (Figure 79).

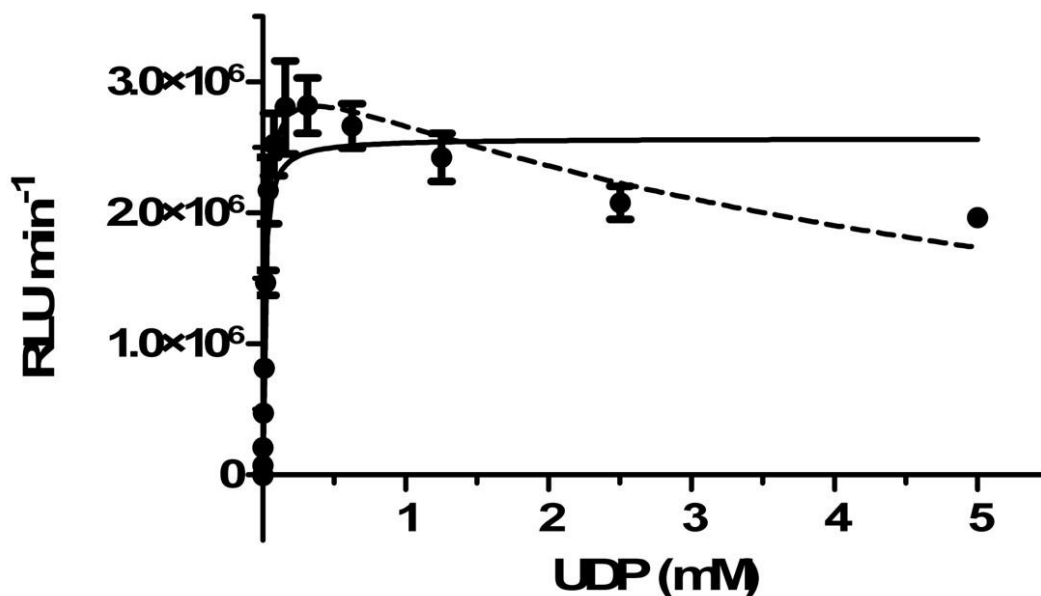


Figure 79: Michaelis-Menten (continuous line) and substrate inhibition (dashed line) of 0.2 mg mL⁻¹ Cmk, 0.1 mM ADP and 0 – 5 mM UDP in reaction buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂). Tested with BacTiter-Glo™ reagent. Output in relative units of luminescence produced per minute. Michaelis-Menten analysed using GraphPad Prism® v. 5.0. Error bars indicate standard deviation from the mean; n = 3 biological replicates.

Table 16 shows that the substrate inhibition analysis has a better fit ($R^2 = 0.9759$) than the Michaelis-Menten equation ($R^2 = 0.9210$). The K_i for UDP as an inhibitor of Cmk is 6.1510 mM UDP (SE = 0.8656; 14 %). For V_{max} , the SE was approximately 3 % for both calculations. However, for K_M the SE was 17 % for the Michaelis-Menten equation ($K_M = 0.0139$ mM UDP, SE = 0.0024 mM UDP) and 10 % for substrate inhibition ($K_M = 0.0223$ mM UDP, SE = 0.0023 mM UDP).

Table 16: Michaelis-Menten and substrate inhibition analysis of 0 – 5 mM UDP with 0.2 mg mL⁻¹ mesophilic Cmk and 0.1 mM ADP. Assay run in reaction buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂) for 10 min at 37 °C. Analysed with GraphPad Prism® v. 5.0. SE = standard error.

	Michaelis-Menten	Substrate inhibition
V_{max} (RLU min ⁻¹)	2.6 x 10 ⁶	3.2 x 10 ⁶
SE: V_{max} (RLU min ⁻¹)	7.9 x 10 ⁴	8.9 x 10 ⁴
K_M (mM UDP)	0.0139	0.0223
SE: K_M (mM UDP)	0.0024	0.0023
K_i (UDP mM)	-	6.1510
SE: K_i (UDP mM)	-	0.8656
R ²	0.9210	0.9759

4.3.3 Kinase activity by ATP production

To test luminescence output with the BacTitre Glo™ kit an ATP standard curve of 0 to 0.6 mM was analysed with the Michaelis-Menten model (Figure 80). Michaelis-Menten analysis of the ATP standard curve gave a V_{max} of 8.3 x 10⁶ RLU min⁻¹ and K_M of 0.0202 mM ATP using the BacTitre Glo™ kit. This also provided a standard curve for determining the kinetic parameters of the kinases.

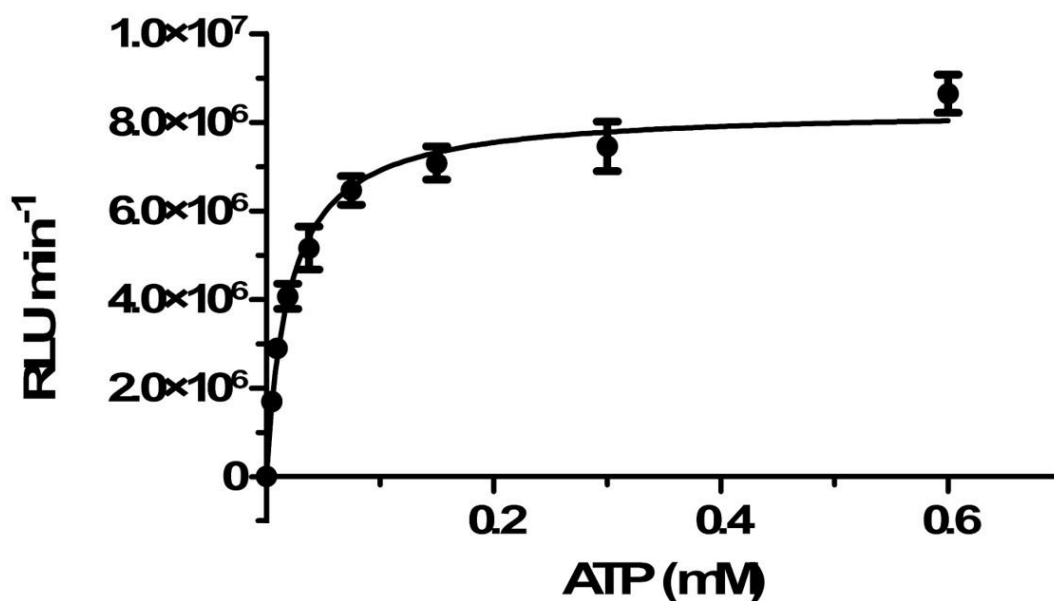


Figure 80: Standard curve of 0 - 0.6 mM ATP in reaction buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂), tested with BacTiter-Glo™ reagent. Output in relative units of luminescence produced per minute. Michaelis-Menten analysed using GraphPad Prism® v. 5.0. Error bars indicate standard deviation from the mean; n = 6 biological replicates.

This curve was used as the standard to analyse ATP production by the kinases. Using 0.2 mg mL⁻¹ of kinase, 0.1 mM ADP and 0 to 0.6 mM NDP (CDP, GDP, UDP or TDP depending on the kinase), ATP output was interpolated from the standard curve using GraphPad Prism® v.5.0 (Table 17).

Table 17: V_{max} and K_M parameters for mesophilic and thermostable kinase with 0 – 0.6 mM NDP, 0.1 mM ADP and 0.2 mg mL⁻¹ kinase. ATP produced per minute by kinases determined by Michaelis-Menten analysis of results interpolated from an ATP standard curve. Reaction run for 10 min incubation at 37 °C and 65 °C for mesophilic and thermophilic kinases respectively. SE = standard error.

Kinase	Mesophilic		Thermophilic				
	Cmk	Gmk	Cmk		Gmk	Tmk	Umpk
NDP	UDP	GDP	UDP	CDP	GDP	TDP	UDP
V_{max} (ATP μ M min ⁻¹)	2.1390	0.4170	3.1250	1.4760	3.3880	0.3977	6.4750
SE: V_{max} (ATP μ M min ⁻¹)	0.1570	0.0094	0.2138	0.0116	0.1849	0.0136	0.5893
K_M (NDP mM)	0.0361	0.0895	0.1120	0.0189	0.0946	0.0231	0.1493
SE: K_M (NDP mM)	0.0100	0.0060	0.0215	0.0007	0.0152	0.0033	0.0351
R ²	0.8502	0.994	0.9287	0.9967	0.9477	0.9496	0.9082

The highest reaction rate was demonstrated by thermophilic Umpk plus UDP with a V_{max} of $6.4750 \mu\text{M ATP min}^{-1}$ and the lowest was $0.3977 \mu\text{M ATP min}^{-1}$ for thermophilic Tmk plus TDP. However, thermophilic Umpk with UDP had the highest K_M (0.1493 mM) whereas thermophilic Cmk had the lowest K_M with CDP (0.0189 mM).

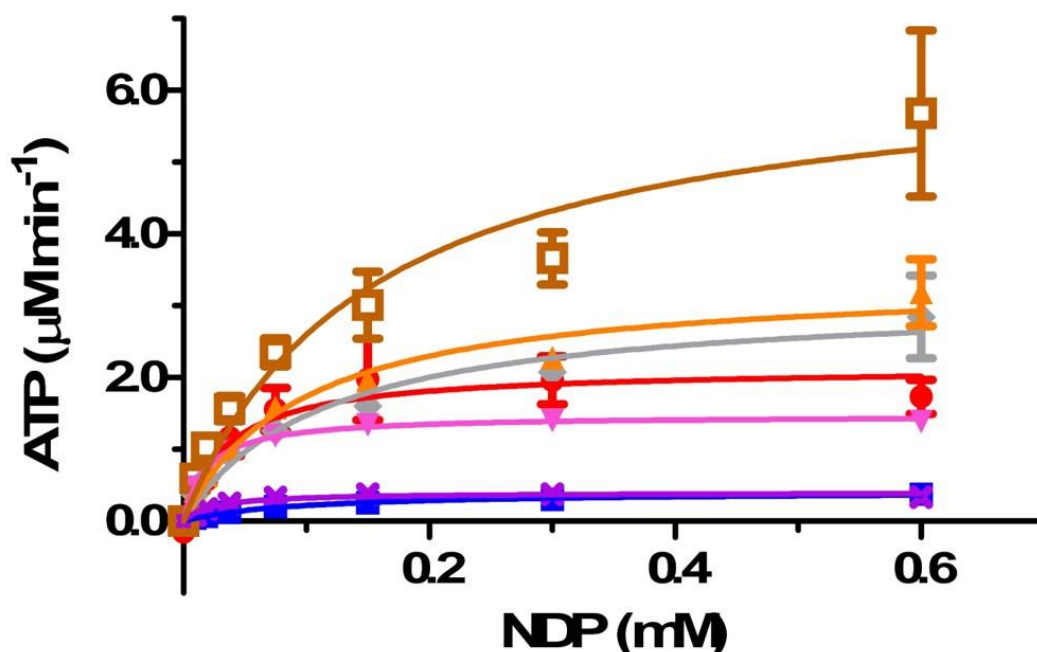


Figure 81: Michaelis-Menten analyses of ATP (μM) produced by kinases per minute, interpolated from an ATP standard curve. Mesophylic Cmk (red circle), thermophilic Cmk (grey diamond) and thermophilic Umpk (brown open square) with UDP, thermophilic Cmk with CDP (pink inverted triangle), mesophylic Gmk (blue square) and thermophilic Gmk (orange triangle) with GDP and thermophilic Tmk (purple cross) with TDP. Reactions run in buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl_2 , 10 mM MnCl_2), at 37 °C or 65 °C with 0.2 mg mL^{-1} kinase and 0.1 mM ADP for 10 min. ATP produced detected using BacTiter-Glo™ reagent. Error bars indicate standard deviation from the mean; $n = 3$ biological replicates ($n = 2$ for mesophylic Gmk).

The fit of the analyses (Figure 81) were good for most of the curves (>0.94), however it was lower for mesophylic Cmk (0.8502), thermophilic Cmk with UDP (0.9287) and Umpk (0.9082). The SE values for V_{max} were all below 10 %. However, for K_M , the best SE was 4 % ($K_M = 0.0189$, SE = 0.0007) for thermophilic

Cmk with CDP and the worst 28 % ($K_M = 0.0361$, $SE = 0.0100$) for mesophilic Cmk.

4.3.4 Kinase standard curves for coupled assay development

4.3.4.1 Mesophilic kinases

Standard curves were developed for the coupled kinase assay for glycosyltransferase characterisation. The standard curves used 0.2 mg mL⁻¹ kinase, 0.1 mM ADP and 0 – 0.6 mM NDP (UDP, CDP, TDP or GDP) in assay buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂). V_{max} and K_M were determined by Michaelis-Menten analysis (Table 18).

Table 18: V_{max} and K_M parameters for mesophilic and thermostable kinase with 0 – 0.6 mM NDP, 0.1 mM ADP and 0.2 mg mL⁻¹ kinase. Measuring relative luminescence produced per minute with BacTiter Glo™ reagent. Reaction run for 10 min incubation at 37 °C and 65 °C for mesophilic and thermophilic kinases respectively. SE = standard error.

Kinase	Mesophilic		Thermophilic				
	Cmk	Gmk	Cmk		Gmk	Tmk	Umpk
NDP	UDP	GDP	UDP	CDP	GDP	TDP	UDP
V_{max} (RLU min ⁻¹)	4.5 x 10 ⁶	1.4 x 10 ⁶	4.7 x 10 ⁶	3.5 x 10 ⁶	5.0 x 10 ⁶	1.2 x 10 ⁶	5.9 x 10 ⁶
SE: V_{max} (RLU min ⁻¹)	1.4 x 10 ⁵	2.6 x 10 ⁴	1.1 x 10 ⁵	1.8 x 10 ⁴	9.6 x 10 ⁴	3.7 x 10 ⁴	1.2 x 10 ⁵
K_M (mM NDP)	0.0379	0.0765	0.0315	0.0116	0.0291	0.0212	0.0234
SE: K_M (mM NDP)	0.0032	0.0044	0.0030	0.0003	0.0022	0.0028	0.0020
R ²	0.9816	0.9957	0.9764	0.9981	0.9841	0.9573	0.9813

Mesophilic Cmk and Gmk were compared (Figure 82). Michaelis-Menten analysis determined Cmk had a higher V_{max} than Gmk (4.5 x 10⁶ and 1.4 x 10⁶ RLU min⁻¹ respectively) and lower K_M (0.0379 mM UDP compared to 0.0765 mM GDP) (Table 18).

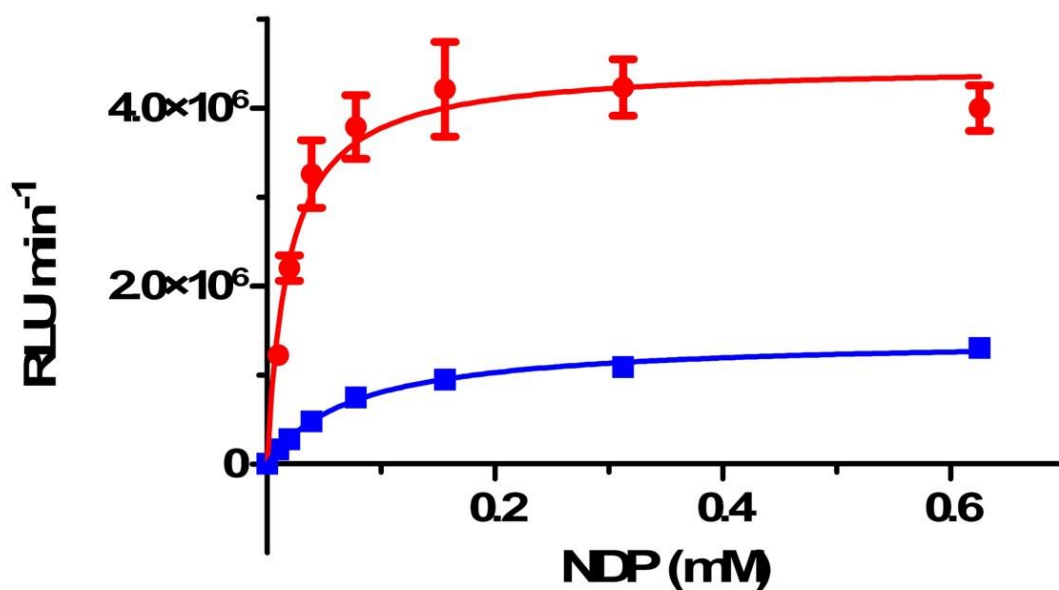


Figure 82: Michaelis-Menten curves comparing mesophilic Cmk (red circle) with UDP and Gmk (blue square) with GDP. Reactions run in buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂), at 37 °C with 0.2 mg mL⁻¹ kinase and 0.1 mM ADP for 10 min. ATP produced detected using BacTiter-Glo™ reagent giving an output of luminescence produced per minute. Error bars indicate standard deviation from the mean; n = 3 biological replicates (n = 2 for mesophilic Gmk).

4.3.4.2 Thermophilic kinases

Thermophilic Cmk (with UDP), Gmk, Tmk and Umpk were compared (Figure 83).

The V_{max} ranged from 5.9 x10⁶ RLU min⁻¹ for Umpk to 1.2 x10⁶ RLU min⁻¹ for Tmk.

The K_M ranged from 0.0315 mM UDP for Cmk to 0.0116 mM CDP for Cmk (Table 18).

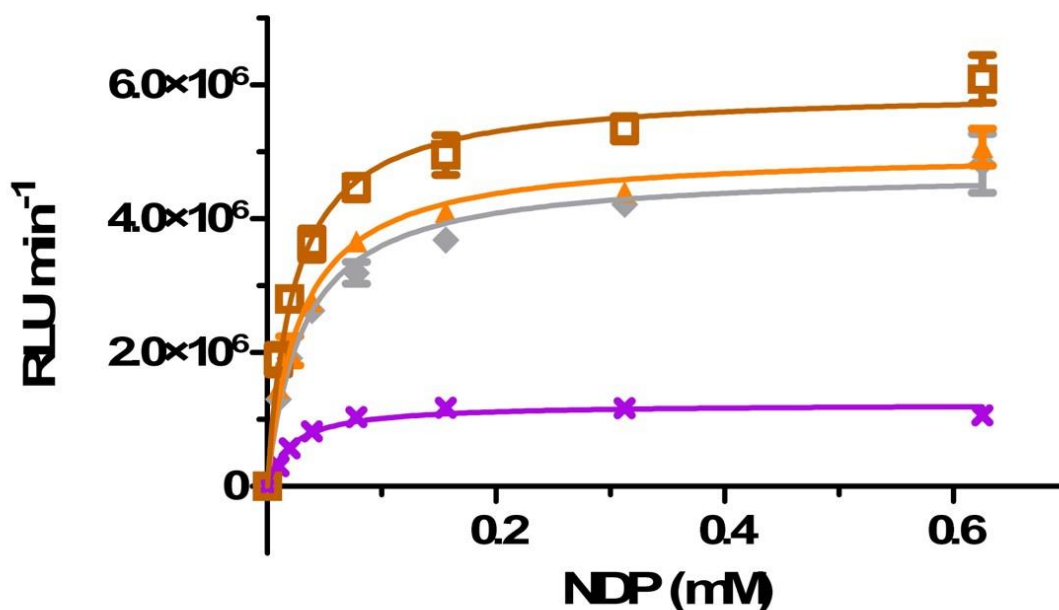


Figure 83: Michaelis-Menten curves comparing thermophilic Cmk (grey diamond) and Umpk (brown open square) with UDP, Gmk (orange triangle) with GDP and Tmk (purple cross) with TDP. Reactions run in buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂), at 65 °C with 0.2 mg mL⁻¹ kinase and 0.1 mM ADP for 10 min. ATP produced detected using BacTiter-Glo™ reagent giving an output of luminescence produced per minute. Error bars indicate standard deviation from the mean; n = 3 biological replicates.

The kinase standard curves had good accuracy, with the SE for all V_{max} and K_M values below 10 % (except for the Tmk K_M which was 13 %; $K_M = 0.0212$ mM TDP; SE = 0.0028 mM TDP) (Table 18). The R^2 values were all above 0.95.

4.3.4.3 Comparison of Cmk

The results for Cmk (thermophilic with UDP and CDP and mesophilic with UDP) can be compared (Figure 84). All the Cmk reactions had similar rates, ranging from $V_{max} = 3.5 \times 10^6$ to 4.7×10^6 RLU min⁻¹. Both mesophilic and thermophilic Cmk had similar K_M values for UDP (0.0379 mM and 0.0315 mM respectively). Thermophilic Cmk showed a higher affinity for CDP with a $K_M = 0.0116$ mM.

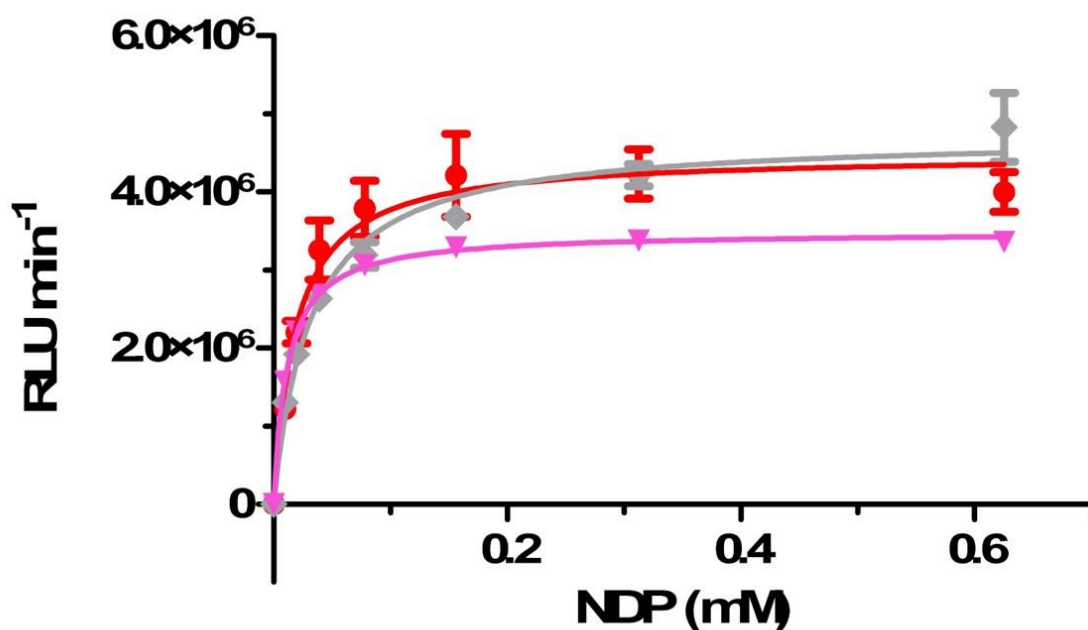


Figure 84: Michaelis-Menten curves comparing mesophilic Cmk (red circle) and thermophilic Cmk (grey diamond) with UDP and thermophilic Cmk with CDP (pink inverted triangle). Reactions run in buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂), at 37 °C or 65 °C with 0.2 mg mL⁻¹ kinase and 0.1 mM ADP for 10 min. ATP produced detected using BacTiter-Glo™ reagent giving an output of luminescence produced per minute. Error bars indicate standard deviation from the mean; n = 3 biological replicates.

4.3.4.4 Comparing mesophilic and thermophilic kinases at 37 °C

Mesophilic Cmk was compared with thermophilic Cmk and Gmk activity at 37 °C to determine whether the thermostable reactions would run at lower temperatures (usually 65 °C) (Figure 85).

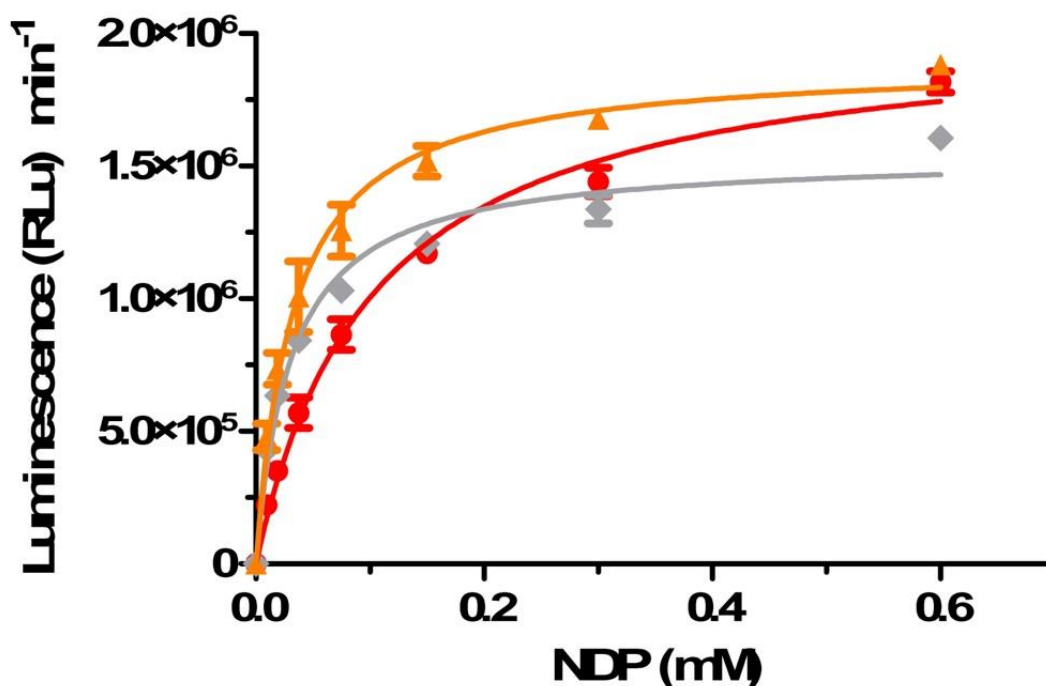


Figure 85: Michaelis-Menten curves comparing mesophilic Cmk (red circle) and thermophilic Cmk (grey diamond) with UDP and thermophilic Gmk (orange triangle) with GDP. Reactions run in buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂), at 37 °C with 0.2 mg mL⁻¹ kinase and 0.1 mM ADP for 10 min. ATP produced detected using BacTiter-Glo™ reagent giving an output of luminescence produced per minute. Error bars indicate standard deviation from the mean; n = 3 biological replicates.

The mesophilic Cmk and thermostable Gmk had similar V_{max} values (2.0×10^6 and 1.9×10^6 RLU min⁻¹ respectively) with the thermostable Cmk reaction running slightly slower (1.5×10^6 RLU min⁻¹). Accuracy for V_{max} was good (SE was 3 % or lower). For K_M the value was similar for thermostable Cmk and Gmk (0.0306 mM UDP and 0.0326 mM GDP) but higher for the mesophilic Cmk (0.1036 mM UDP). The accuracy for K_M was good (SE less than 12 % for each) and so was the curve fit (R^2 greater than 0.97 for all curves) (Table 19).

Table 19: V_{max} and K_M parameters for mesophilic Cmk and thermostable Cmk and Gmk with 0 – 0.6 mM NDP, 0.1 mM ADP and 0.2 mg mL⁻¹ kinase. Measuring relative luminescence produced per minute with BacTiter Glo™ reagent. Reaction run for 10 min incubation at 37 °C. SE = standard error.

Kinase	Mesophilic	Thermophilic	
	Cmk	Cmk	Gmk
NDP	UDP	UDP	GDP
V_{max} (RLU min ⁻¹)	2.0 x 10 ⁶	1.5 x 10 ⁶	1.9 x 10 ⁶
SE: V_{max} (RLU min ⁻¹)	6.0 x 10 ⁴	4.5 x 10 ⁴	4.4 x 10 ⁴
K_M (mM NDP)	0.1036	0.0306	0.0326
SE: K_M (mM NDP)	0.0087	0.0035	0.0029
R ²	0.9913	0.9774	0.9866

4.3.5 Characterisation of β 14GalT

For characterisation of the commercially-available β 14GalT enzyme, firstly the enzyme and substrate concentrations had to be optimised. This was done using the kinase assay with the previously defined optimum conditions with mesophilic Cmk.

4.3.5.1 β 14GalT concentration

To determine the concentration of β 14GalT, the enzyme was titrated from 0 to 2 mU mL⁻¹ and the assay run with 5 mM UDP-Gal and 5 mM GlcNAc (Figure 86).

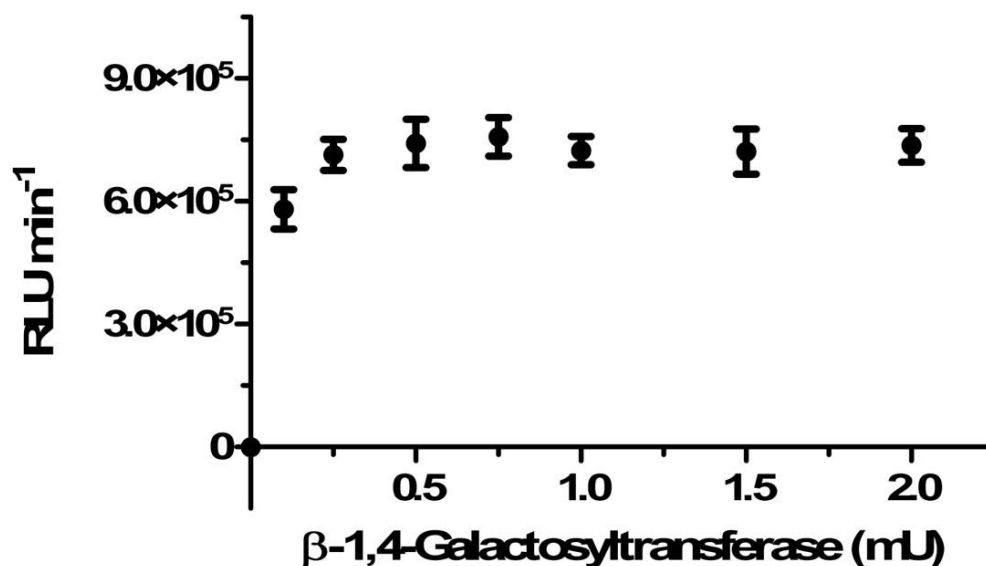


Figure 86: Enzyme titration of β 14GalT with 5 mM UDP-Gal and 5 mM GlcNAc. Measured with the kinase assay using 0.2 mg mL⁻¹ mesophilic Cmk and 0.1 mM ADP. Detected with the BacTiter-Glo™ kit, output given as relative luminescence produced per minute. Error bars indicate standard deviation from the mean; n = 3 biological replicates.

Figure 86 shows the reaction was quickly saturated with enzyme and the linear titration would occur between 0 – 0.25 mU mL⁻¹ β 14GalT. As the curve plateaued at 0.25 mU mL⁻¹ β 14GalT, this was the concentration of enzyme used for all further experiments.

4.3.5.2 Donor substrate (UDP-Gal) concentration

The range of UDP-Gal (donor substrate) to use in the assay for β 14GalT characterisation was determined. Reactions with 0.25 mU mL⁻¹ β 14GalT and 5 mM GlcNAc were tested with 0 – 2 mM UDP-Gal and measured using the kinase assay (Figure 88).

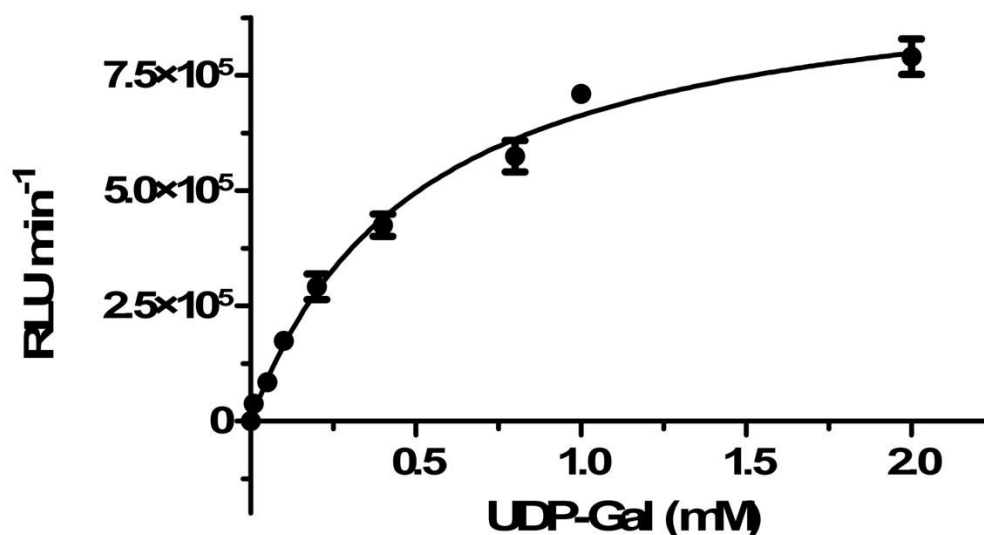


Figure 87: Michaelis-Menten analysis to optimise UDP-Gal concentration. Using 0.25 mU mL⁻¹ β14GalT and 5 mM GlcNAc with 0 – 2 mM UDP-Gal in reaction buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂). Run for 15 min at 37 °C. Measured using the coupled kinase assay and BacTiter-Glo™ reagent; output given as relative luminescence produced per minute. Error bars indicate standard deviation from the mean; n = 3 biological replicates.

Michaelis-Menten analysis determined $V_{max} = 1.5 \times 10^7$ RLU and $K_M = 0.5113$ mM UDP-Gal (Table 20). The curve was a good fit, with SE for both V_{max} and K_M below 9 % and an $R^2 = 0.9892$ (Table 20). For further assays the range of UDP-Gal used was 0 - 6 mM, as the maximum of the range was greater than 10 x the K_M .

4.3.5.3 Acceptor substrate (GlcNAc) concentration

The concentration of GlcNAc (acceptor substrate) to use in the assay was determined. Reactions with 0.25 mU mL⁻¹ β14GalT and 6 mM UDP-Gal were tested with 0 – 30 mM GlcNAc and measured using the kinase assay (Figure 88).

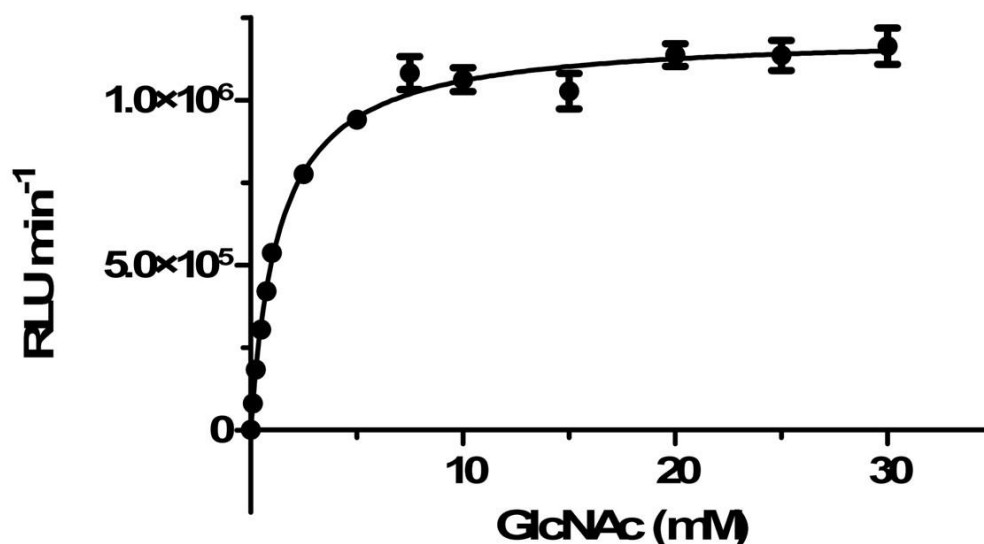


Figure 88: Michaelis-Menten analysis to optimise GlcNAc concentration. Using 0.25 mU mL⁻¹ β 14GalT and 6 mM UDP-Gal with 0 – 30 mM GlcNAc in reaction buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂). Run for 15 min at 37 °C. Measured using the coupled kinase assay and BacTiter-Glo™ reagent; output given as relative luminescence produced per minute. Error bars indicate standard deviation from the mean; n = 3 biological replicates.

The Michaelis-Menten analysis gave $V_{max} = 1.8 \times 10^7$ RLU and $K_M = 1.3420$ mM GlcNAc (Table 20). The curve was a good fit, with SE for both V_{max} and K_M below 6 % and an $R^2 = 0.9908$ (Table 20). For further assays the concentration of GlcNAc used was 15 mM, as this was greater than ten times the K_M , ensuring plenty of substrate availability for the reaction to proceed.

Table 20: V_{max} and K_M parameters of 0.25 mU mL⁻¹ β 14GalT with either 0 – 2 mM UDP-Gal (plus 5 mM GlcNAc) or 0 – 30 mM GlcNAc (plus 6 mM UDP-Gal). SE = standard error.

Substrate	0 – 2 mM UDP-Gal	0 – 30 mM GlcNAc
V_{max} (RLU min ⁻¹)	1.0×10^6	1.2×10^6
SE: V_{max} (RLU min ⁻¹)	3.3×10^4	1.3×10^4
K_M (mM)	0.5113	1.3420
SE: K_M (mM)	0.0432	0.0678
R^2	0.9892	0.9908

4.3.6 Determining kinetic parameters for β 14GalT

To characterise β 14GalT, both the novel coupled-kinase and commercial coupled-phosphatase assays were used for comparison. The reaction conditions were: 0.25 mU mL⁻¹ β 14GalT, 15 mM GlcNAc, 0 – 6 mM UDP-Gal for 10 min at 37 °C in buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂).

4.3.6.1 Novel coupled kinase assay

For the kinase assay, 0.2 mg mL⁻¹ mesophilic Cmk and 0.1 mM ADP were added. A 0 – 0.6 mM UDP standard curve was run with 0.1 mM ADP and 0.2 mg mL⁻¹ mesophilic Cmk for interpolating the reaction output into UDP produced per minute (Figure 89).

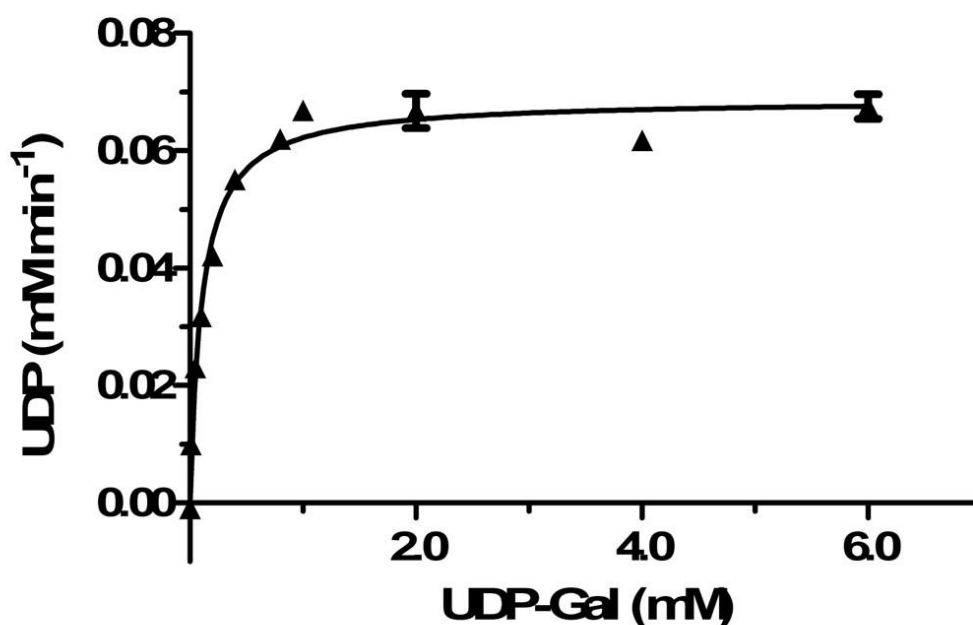


Figure 89: Michaelis-Menten curve of 0.25 mU mL⁻¹ β 14GalT with 15 mM GlcNAc and 0 – 6 mM UDP-Gal in buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂). Incubated at 37 °C for 10 min with 0.2 mg mL⁻¹ coupled mesophilic Cmk and 0.1 mM ADP. Detected using BacTiter-Glo reagent. Output given as UDP produced per minute, interpolated from UDP Michaelis-Menten standard curve. Error bars indicate standard deviation from the mean; n = 3 biological replicates.

Using the coupled kinase assay, β 14GalT measured a V_{max} of 68.7500 μ M UDP produced per minute and K_M of 0.1055 mM UDP-Gal. The accuracy and curve fit were good. The SE was 1.4 % and 7.5 % of the V_{max} and K_M respectively and the R^2 value of the curve was 0.9797 (Table 21).

Table 21: V_{max} and K_M for β 14GalT with UDP-Gal determined using a novel kinase assay and commercial phosphatase assay. Run with 0.25 mU mL⁻¹ β 14GalT and 15 mM GlcNAc at 37 °C for 10 min. * V_{max} given in UDP and phosphate produced per minute for the kinase and phosphatase assays respectively. SE = standard error. ¹ Glycosyltransferase Activity Kit, R&D Systems.

	Coupled-kinase assay	Coupled-phosphatase assay ¹
V_{max} (μ M min ⁻¹)*	68.7500	1.2070
SE: V_{max} (μ M min ⁻¹)*	0.9927	0.0239
K_M (mM)	0.1055	0.1631
SE: K_M (mM)	0.0079	0.0151
R^2	0.9797	0.9775

4.3.6.2 Commercial phosphatase assay

The same conditions were used for the coupled-phosphatase assay (R&D systems; Wu *et al.*, 2011). The results were interpolated from a linear phosphate standard titration and the given as phosphate produced per minute (Figure 90).

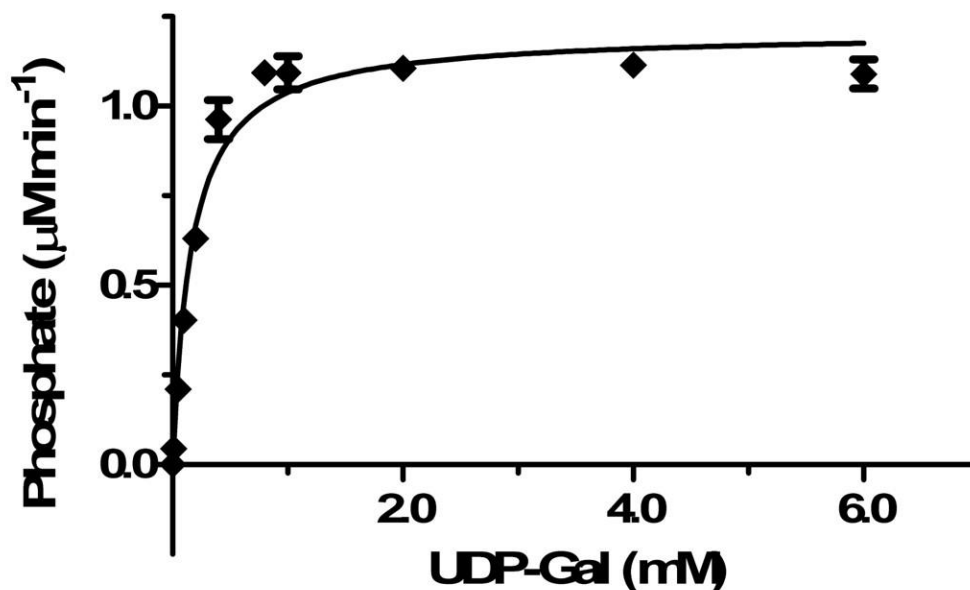


Figure 90: Michaelis-Menten curve of 0.25 mU mL⁻¹ β 14GalT with 15 mM GlcNAc and 0 – 6 mM UDP-Gal in buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂). Incubated at 37 °C for 10 min. Detected using coupled phosphatase kit (R&D systems). Output given as phosphate produced per minute, interpolated from phosphate linear standard curve. Error bars indicate standard deviation from the mean; n = 3 biological replicates.

With the coupled-phosphate assay (R&D systems; Wu *et al.*, 2011), β 14GalT demonstrated a V_{max} of 1.2070 μ M phosphate produced per minute and K_M of 0.1631 mM UDP-Gal. The accuracy and curve fit were good. The SE was 2.0 % and 9.3 % of the V_{max} and K_M respectively and the R^2 value of the curve was 0.9775 (Table 21).

4.3.7 Determining kinetic parameters for LgtC

4.3.7.1 Novel coupled kinase assay

LgtC was characterised using the same coupled-kinase protocol as β 14GalT. The reaction contained 7 g mL⁻¹ LgtC, 2 mM lactose, 0 – 0.5 mM UDP-Gal and 2 mM DTT. The concentration values for a Michaelis-Menten curve was interpolated from a UDP standard curve (Figure 91).

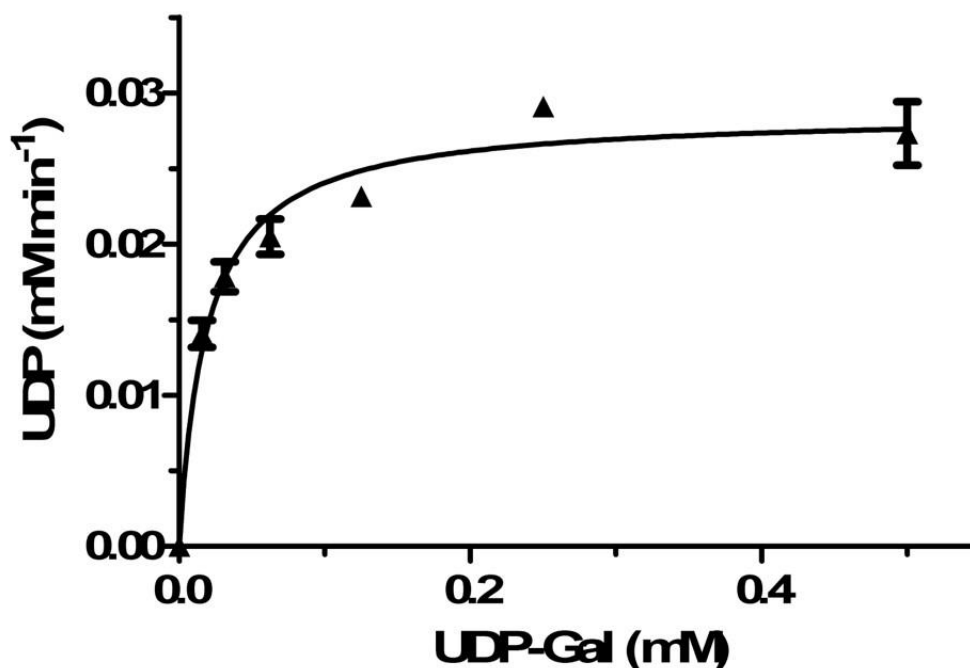


Figure 91: Michaelis-Menten curve of $7 \mu\text{g mL}^{-1}$ LgtC with 2 mM lactose and 0 – 0.5 mM UDP-Gal in buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl_2 , 10 mM MnCl_2). Incubated at 37 °C for 10 min with 0.2 mg mL^{-1} coupled mesophilic Cmk and 0.1 mM ADP. Detected using BacTiter-Glo reagent. Output given as UDP produced per minute, interpolated from UDP Michaelis-Menten standard curve. Error bars indicate standard deviation from the mean; $n = 3$ biological replicates.

LgtC, measured using the coupled kinase assay, gave a V_{max} of $0.0287 \mu\text{M UDP}$ produced per minute and K_M of $0.0191 \text{ mM UDP-Gal}$. The accuracy was variable, with a good SE of 2.8 % for V_{max} , but a higher SE of 13.1 % for K_M . The curve fit value R^2 was 0.9433 (Table22).

Table 22: V_{max} and K_M for LgtC with UDP-Gal determined using a novel kinase assay and commercial phosphatase assay. Run with $7 \mu\text{g mL}^{-1}$ LgtC and 2 mM lactose at 37 °C for 10 min. * V_{max} given in UDP and phosphate produced per minute for the kinase and phosphatase assays respectively. SE = standard error. ¹ Glycosyltransferase Activity Kit, R&D Systems.

	Coupled-kinase assay	Coupled-phosphatase assay ¹
V_{max} ($\mu\text{M min}^{-1}$)*	28.6600	0.7046
SE: V_{max} ($\mu\text{M min}^{-1}$)*	0.8078	0.0263
K_M (mM)	0.0191	0.0114
SE: K_M (mM)	0.0025	0.0021
R^2	0.9433	0.9223

4.3.7.2 Commercial phosphatase assay

The same enzyme and substrate concentrations were used for the coupled-phosphatase assay (R&D systems; Wu *et al.*, 2011). The results were interpolated from a linear phosphate standard titration and given as phosphate produced per minute (Figure 92).

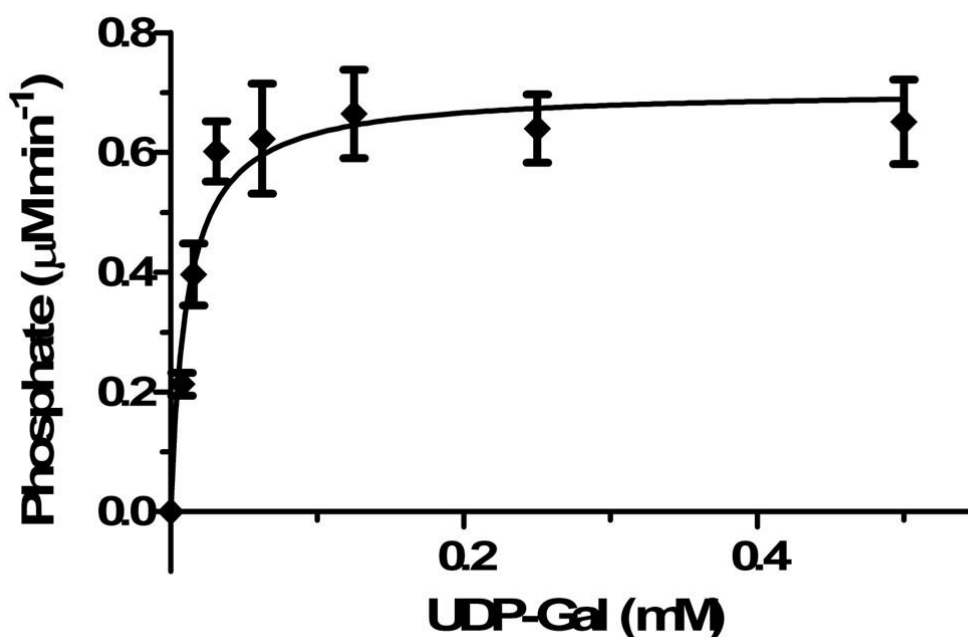


Figure 92: Michaelis-Menten curve of 7 µg mL⁻¹ LgtC with 2 mM lactose and 0 – 0.5 mM UDP-Gal in buffer (50 mM HEPES pH 8.0, 50 mM KCl, 10 mM MgCl₂, 10 mM MnCl₂). Incubated at 37 °C for 10 min. Detected using a coupled phosphatase kit (R&D systems). Output given as phosphate produced per minute, interpolated from phosphate linear standard curve. Error bars indicate standard deviation from the mean; n = 3 biological replicates.

With the coupled-phosphate assay (R&D systems; Wu *et al.*, 2011), LgtC measured a V_{\max} of 0.7046 µM phosphate produced per minute and K_M of 0.0114 mM UDP-Gal. The accuracy for V_{\max} was good (SE = 3.7 %), however it was poor for K_M (SE = 18.4 %). The R^2 value of the curve was 0.9223 (Table 22).

4.3.8 Commercial fluorescence assay

No results were obtained using the Transcreener® ADP Fluorescence Polarization Assay (BellBrook Labs). This assay required a plate reader with fluorescence polarisation: such an instrument was not available at the University of Exeter during the period of the relevant research. The same manufacturers suggested their Transcreener® ADP Fluorescence Intensity Assay (BellBrook Labs), which they market for use with UDP in fluorescence intensity mode. However, on testing, no reliable results were acquired.

4.3.9 Apyrase purification

The apyrase from *C. lectularius* was expressed in *E. coli* and purified using dialysis which was previously performed by Valenzuela *et al.* (1998). The gel in Figure 93 shows that a protein of the expected size was successfully expressed. However, there is only a weak band of protein (suggesting low concentration) from the upper portion of the dialysis bag and it is absent from the supernatant after centrifugation (Figure 64). The majority of the protein is found in the lower part of the bag, where precipitate had settled, and entirely in the pellet after centrifugation. This indicates the apyrase was insoluble upon refolding.

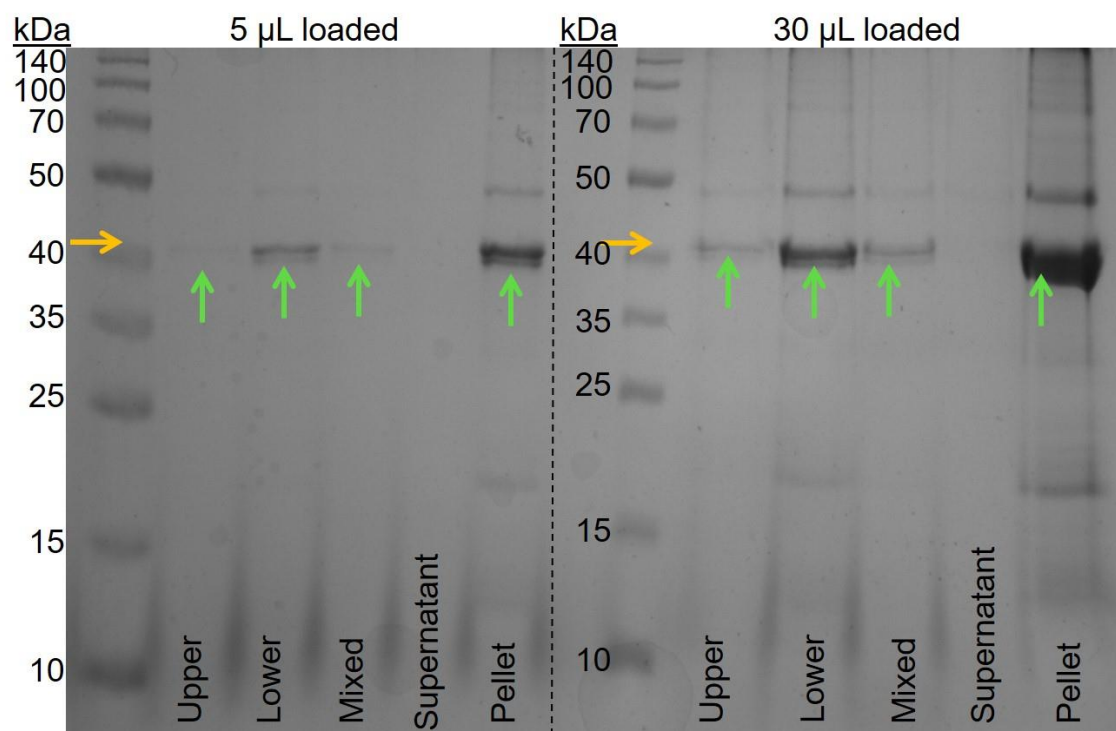


Figure 93: SDS-PAGE of apyrase after dialysis for 48 h. Expected size = 41.5 kDa (orange arrow). Observed size (green arrow). Samples taken from the upper and lower part of the dialysis bag, after mixing the bag and the supernatant and pellet after centrifuging the contents. Samples loaded in 5 μ L and 30 μ L aliquots following treatment with loading buffer.

The expression and purification was repeated, with the addition of a TCA precipitation step to concentrate the protein in the aliquots. The gel in Figure 94 demonstrates a much more concentrated protein is obtained by this method. There is also a much stronger band in the upper portion of the dialysis bag and a weak band in the supernatant after centrifugation, suggesting there may be some soluble apyrase present.

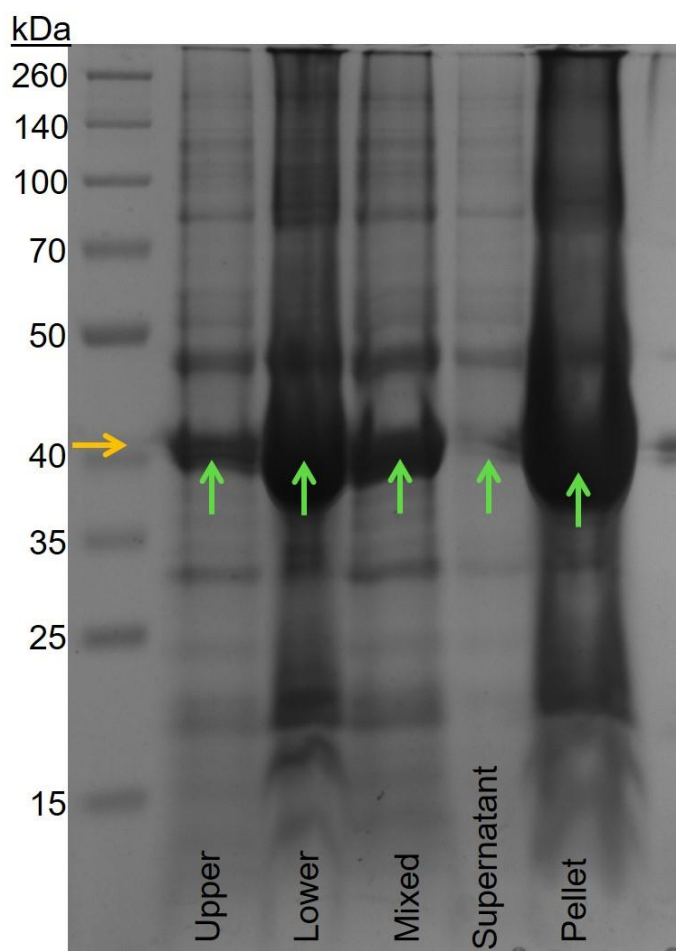


Figure 94: SDS-PAGE of apyrase after dialysis for 48 h and TCA precipitation of the samples. Expected size = 41.5 kDa (orange arrow). Observed size (green arrow). Samples taken from the upper and lower part of the dialysis bag, after mixing the bag and the supernatant and pellet after centrifuging the contents. 20 μ L sample plus loading buffer in each lane.

Due to the bands apparent in the upper portion of the dialysis bag and the supernatant, a final attempt to extract soluble apyrase from *E. coli* was made. Following dialysis, the dialysed protein was purified with nickel affinity chromatography. The gel in Figure 95 had very weak bands in the elution and concentrated protein samples. However, the bands are not of the expected size or of the size observed in Figure 93 and Figure 94.

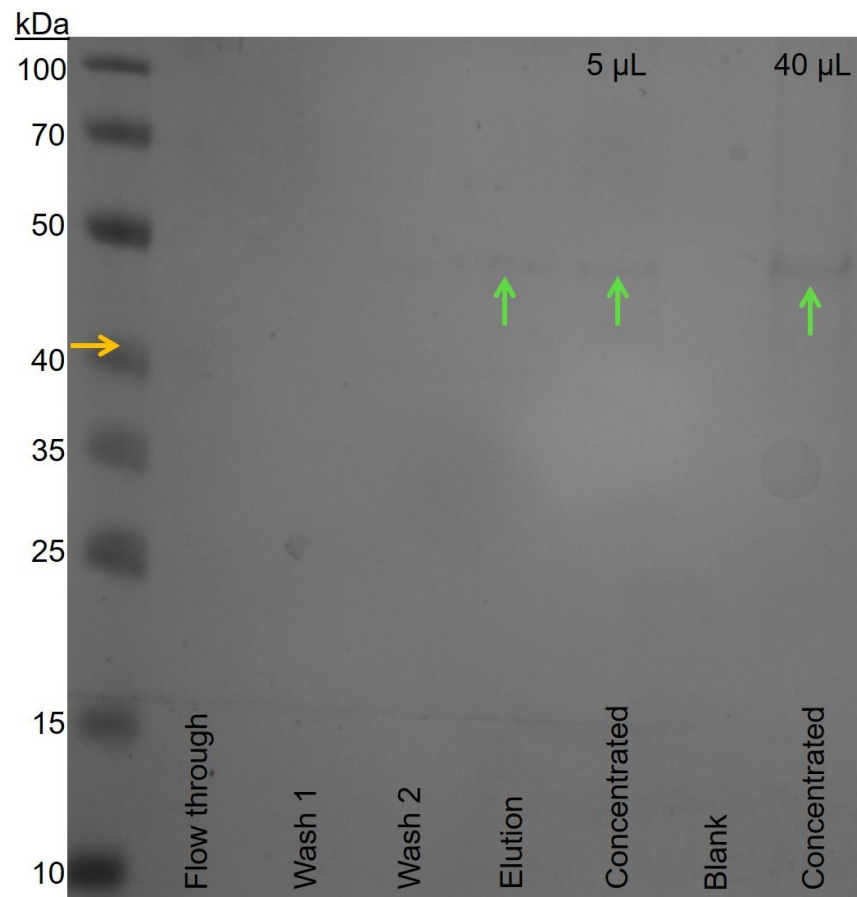


Figure 95: SDS-PAGE of apyrase after dialysis for 48 h and nickel affinity chromatography. Expected size = 41.5 kDa (orange arrow). Observed size (green arrow). Samples taken from the flow through, wash steps 1 and 2, elution and concentrated protein (5 μ L and 40 μ L aliquots of concentrated protein plus loading buffer). 20 μ L sample plus loading buffer in all other lanes.

4.4 Discussion

4.4.1 Methods for glycosyltransferase characterisation

Characterisation of glycosyltransferases has remained challenging despite the prevalence and importance of these enzymes to cellular functions. Commercial kits available are either expensive, unreliable or require specialist equipment. For example, the Glycosyltransferase Activity Kit (R&D Systems; Wu *et al.*, 2011) costs £1.65 per reaction and rapidly exhausts stocks of precious purified enzyme. The Transcreener® ADP Fluorescence Polarization Assay (BellBrook Labs) requires a polarising plate reader that is not common laboratory equipment and an effort to adapt the Transcreener® ADP Fluorescence Intensity Assay (BellBrook Labs) yielded no usable results. Finally, during the course of this project a third assay became available; the UDP-Glo™ Glycosyltransferase Assay (Promega). This assay is similar to the one developed here, as it generates ATP from the release of UDP and is measured by production of luminescence. However, this assay is only suitable for use with UDP-sugar glycosyltransferase reactions, limiting the range of glycosyltransferases that can be characterised. (This kit would not be suitable for the glycosyltransferases involved in *B. pseudomallei* CPS synthesis, for example.) The cost of the UDP-Glo™ assay is £1.18 per reaction for a 200-reaction kit to £0.39 per reaction for a 4,000-reaction kit. However, with the method developed here an inexpensive supply of kinases could be generated in-house.

Other approaches to investigating glycosyltransferase activities include radiolabelled substrates (Hartman *et al.*, 2007; Ardiccioni *et al.*, 2016) which are hazardous to handle and dispose of, fluorescent chemosensors (Wongkongkatep *et al.*, 2006; Sakamoto, Ojida & Hamachi *et al.*, 2009) which are complex to

synthesise and mass spectrometry (Zhou *et al.*,2016) which does not allow for determination of kinetic parameters.

4.4.2 Novel coupled-kinase assay

The aim of this novel assay was to provide a cost-effective, robust protocol that would not exhaust precious stocks of purified enzyme and could be adapted to a range of functions. The commercial kit used as a control in this chapter (R&D systems; Wu *et al.*, 2011) couples a phosphatase to the reaction that can remove a phosphate group from any free nucleoside diphosphate. The free phosphate is then detected using a colorimetric test. Comparatively, the concept of using a kinase to transfer a phosphate group to generate ATP is complicated by the substrate-specificity of kinases. (This problem occurs with the UDP-Glo™(Promega) kit.) However, a set of kinases were identified that enables all CDP-, GDP-, dTDP- and UDP-sugar glycosyltransferase activities to be analysed. (A robust and cost effective assay for ADP-sugar glycosyltransferase activity already exists, coupling ADP production with NADH oxidation using pyruvate kinase and lactate dehydrogenase, Wayllace *et al.*, (2012)). As these kinases are found in wild type *E. coli* K12, the genes can easily be obtained with specific primers and a destination vector. Here, LIC cloning into the pNIC28-Bsa vector was used. Alternatively, as DNA synthesis becomes increasingly affordable, the genes could be purchased in a vector with a promoter for a less laborious approach. In this study, the thermostable kinases were synthesised in a vector with a T7 promoter to eliminate the cloning steps.

To detect the generation of ATP by the coupled reactions, this study opted to use a commercially available BacTiter Glo™ kit. The kit was chosen as it is highly sensitive, enabling low quantities of glycosyltransferase and its substrates to be

used in the assay. The reaction volume for the novel assay was less than a third of the volume for the control assay (15 μL requiring 3.75 μU β14GalT , compared to 50 μL requiring 12.5 μU β14GalT). The cost of BacTiter Glo™ per reaction was £0.08. The BacTiter-Glo™ reagent is normally used for viability testing in disinfectant or antimicrobial studies (Stiefel *et al.*, 2016; Yoona *et al.*, 2015) but has also been adapted for use in an immunoassay detecting specific bacterial pathogens (Hunter & Lim, 2010) and in testing chemoluminescence in a microfluidics system (Terry *et al.*, 2008). No specialist equipment is required to measure the assay, as the BacTiter Glo™ reagent produces luminescence that can be read using most multi-mode plate readers. Five such plate readers are currently available within the Department of Biosciences at the University of Exeter, for example.

The buffer for the assay was HEPES with added magnesium (a requirement for the catalytic activity, Bertrand *et al.*, 2002) and manganese due to evidence it enhances glycosyltransferase activity (Hurwitz, 1959). Potassium was also included as some glycosyltransferases require this ion for activity. Two limitations of the assay are the ADP and NDP concentrations. For ADP, the concentration had to be a balance between activity, background noise, and product inhibition of the coupling luciferase (Figure 77). For the NDP, the concentration had to be restricted to a maximum 0.6 mM due to substrate inhibition (Figure 79), a problem noted previously (Bucurenci *et al.*, 1996). In most cases, the enzymes were not generating this quantity of NDP, so this was not an issue.

4.4.3 Purification of kinases

After synthesising or cloning the kinase genes, the next stage is to over-express and purify the enzymes. As the mesophilic kinases were from *E. coli* and genes

for the *T. thermophilus* thermostable kinases were codon-optimised for expression in *E. coli*, the enzymes all over-expressed with no issues.

For purification, it was discovered to be important to include a rigorous protease-inhibition step. Without this, the kinases underwent some proteolysis as evidenced by SDS-PAGE of mesophilic Cmk (Figure 68). Addition of either a protease inhibitor cocktail or the combination of PMSF and benzamidine was sufficient to prevent this proteolysis (Figure 69).

The mesophilic kinases purified with less contamination and denaturation than the thermostable kinases (Figure 69 (mesophilic Cmk) compared to Figure 75 (thermophilic Umpk)). There are several possible explanations for the difference in protein purity and stability, including incompatibility of the *T. thermophilus* proteins with *E. coli*, the temperature of the culture during expression, the strain of *E. coli* used (Rosetta for mesophilic, BL21 for thermophilic kinases) or the promoter used. The kinases are predicted to be unstable (according to the instability index) by ExPASy ProtParam, suggesting there may be inherent differences in the stability of the proteins. However, all the proteins were purified adequately for use in the glycosyltransferase assay development. It is worth noting that the concentration of the thermostable kinases will be inaccurate due to these purification imperfections.

Once purified, the enzymes were stored for several months. A previous study noted deterioration of mesophilic Cmk at colder temperatures (30 % activity loss after two months at -10 °C) (Hurwitz, 1959). This loss in activity may explain why the kinetic parameters were lower than expected in some assays (Figure 78; Table 15). It was also found to be important not to expose the enzymes to freeze-thaw cycles (note denaturation of one Gmk stock between the two gels in Figure 76).

Overall, purifying the mesophilic kinases was simple, requiring only nickel affinity and gel filtration chromatography. The nickel affinity chromatography provided sufficiently pure Cmk for use in the assay if gel filtration is not available (Figure 66). The thermostable kinases required an additional ion exchange step.

4.4.4 Kinase activity

Nucleotide monophosphate kinases catalyse the transfer of a phosphate group from a nucleoside triphosphate (normally ATP) to a nucleoside monophosphate (such as CMP) to produce a nucleoside diphosphate (such as CDP). Here, however, the reverse reaction is required to generate ATP. This is a reaction that has a small ΔG° ; in the cell, the equilibrium is driven by high cellular concentrations of ATP (and low NMP concentrations) to the NDP products. The literature focuses on the kinetic parameters of the kinases in the forward direction, so all K_M values available are for the nucleoside monophosphate and ATP (Bucurenci *et al.*, 1996; Bertrand *et al.*, 2002). The data obtained here provides these parameters for the nucleoside diphosphate in the reverse reaction (Table 17). It is interesting to note that although the human Cmk enzyme has demonstrated similar affinities for CMP and UMP (Van Rompay, Johansson & Karlsson, 1999), the *E. coli* homologue has displayed selectivity for CMP. Previous studies have demonstrated a 99 % (Briozzo *et al.*, 1998) and 96 % (Bucurenci *et al.*, 1996; Ofiteru *et al.*, 2007) loss in affinity for UMP compared to CMP. The data obtained in this study suggest the same is not true for the reverse reaction. Mesophilic Cmk had a K_M of 0.0379 mM UDP (Table 18) compared with 0.035 mM UMP (Bucurenci *et al.*, 1996; Ofiteru *et al.*, 2007). The thermophilic Cmk from *T. thermophilus* had a K_M of 0.0315 mM for UDP compared with 0.0116 mM for CDP at 65 °C (Table 18). This is a 68 % loss of affinity for UDP

compared with CDP, though due to the lower purity of the thermostable kinase preparations it would be pertinent to investigate this further. Nevertheless, in view of these data it would be advisable to use prokaryotic Umpk instead of Cmk in future assays for glycosyltransferase characterisation.

When comparing the mesophilic kinases (at 37 °C) from *E. coli* to the thermophilic kinases (at 65 °C) from *T. thermophilus*, there is no difference in the V_{max} or K_M values (T-test = 0.1933 and 0.3621 respectively; values below 0.05 indicate statistical significance). This indicates that all the kinases have similar rates of reaction and substrate affinity, despite deriving from different organisms. With the mesophilic kinases, the glycosyltransferase and kinase reactions run concurrently at 37 °C. The thermostable kinases were introduced to the study to enable the glycosyltransferase reaction to run at 37 °C then the kinase reaction at 65 °C while the glycosyltransferase activity is stopped by denaturation of the enzyme. However, when comparing mesophilic Cmk with thermophilic Cmk and Gmk at 37 °C it was demonstrated that the thermophilic kinases worked well at the lower temperature (Table 19). In this experiment the thermophilic kinases demonstrated higher affinity for the substrates ($K_M = 0.0306$ mM UDP for Cmk and 0.0326 mM GDP for Gmk) than the mesophilic kinase ($K_M = 0.1036$ mM UDP). This possibly suggests the thermophilic are more stable when stored and more tolerant of freeze-thaw cycles, as newly-prepared mesophilic Cmk established a K_M of 0.0379 mM UDP at 37 °C in a later assay (Table 18). The rate of reaction was highest for mesophilic Cmk (2.0×10^6 RLU min⁻¹). However, the reaction rates were the same magnitude for the thermophilic kinases (1.5×10^6 and 1.9×10^6 RLU min⁻¹ for Cmk and Gmk respectively) (Table 19). These kinases performed slightly better at 65 °C ($K_M = 0.0315$ mM UDP, $V_{max} = 4.7 \times 10^6$ RLU min⁻¹ for Cmk, $K_M = 0.0291$ mM GDP, $V_{max} = 5.0 \times 10^6$ RLU min⁻¹ for Gmk).

These data indicate that the activity of the thermophilic kinases is too great at 37 °C with too little improvement at 65 °C to use in place of the mesophilic kinases. The thermophilic kinases may have improved tolerance to storage and freeze-thaw cycles, however this was not tested in this study.

4.4.5 Characterisation of glycosyltransferases

The novel assay was compared with an existing assay and data to determine its suitability for characterising glycosyltransferases. A commercially-available glycosyltransferase, β 14GalT from *B. taurus* milk, was chosen for developing and testing the assay. Firstly, the enzyme concentration was optimised by titrating the glycosyltransferase (Figure 86). The concentration 0.25 mU was chosen as it is the point where the graph plateaued, indicating saturation of the reaction. Next, the K_M values for the donor (Figure 87) and acceptor substrates (Figure 88) were determined. The optimal concentrations for the final assay were chosen at greater than ten times the determined K_M (15 mM GlcNAc, $K_M = 1.3420$ mM; 0 - 6 mM UDP-Gal, $K_M = 0.5113$ mM). This ensured there was sufficient substrate for the reaction to proceed.

Previously-determined K_M values for β 14GalT for UDP-Gal were 0.043 mM (Pâquet & Moscarello, 1984) and 0.025 mM (Nakazawa *et al.*, 1993). A more recent study (Tedaldi *et al.*, 2014) calculated an average K_M of 0.0220 mM UDP-Gal for β 1,4GalT using a modified version of the coupled-phosphatase assay. However, the results are not directly comparable as a chicken egg-white lysozyme was included as a carrier protein to stabilise β 1,4GalT and enhance function. In this study, the control phosphatase assay (R&D systems; Wu *et al.*, 2011) measured a K_M of 0.1631 mM and the novel kinase assay 0.1055 mM UDP-Gal. The value determined by both assays was far greater than the previously

measured K_M . However, as the K_M determined here was inconsistent with the literature for both the novel and control assays, the error may exist in the optimisation or operation of the experiments rather than the design of the novel assay. It may be worth noting that during this study production of β 1,4GalT was discontinued and the enzyme was no longer attainable commercially. The K_M obtained using the novel assay was more favourable than the with the control assay, suggesting the coupled-kinase assay is at least as robust as commercially available kits.

Optimisation of the LgtC assay had already been conducted by Dr. Mirella Vivoli (University of Exeter). Existing kinetic data for LgtC reported a range of K_M values from 0.0044 – 0.0200 mM UDP-gal (Wakarchuk *et al.*, 1998; Presson *et al.*, 2001; Lairson *et al.*, 2004) and the recent Tedali *et al.*, (2014) paper determined an average K_M of 0.180 mM UDP-gal. The coupled-kinase assay measured a K_M of 0.0191 mM UDP-Gal (Figure 91), compared to 0.0114 mM UDP-Gal with the control phosphatase assay (Figure 92). The data collected in this study for LgtC is consistent with the existing literature and once again the coupled-kinase assay generated a more favourable K_M than the control assay. The results for LgtC are perhaps further indication of an underlying issue with the β 14GalT experiments.

4.4.6 Apyrase purification

Another approach was to find an alternative phosphatase to the one in the commercial kit (R&D systems; Wu *et al.*, 2011). This would have the advantage of one enzyme being suitable for all glycosyltransferase reactions (i.e. has the ability to remove a phosphate from any free nucleoside diphosphate). An apyrase from the common bedbug, *C. lectularius*, was identified as a potential candidate (Valenzuela *et al.*, 1996). Unlike the mesophilic kinases, the gene had to be

synthesised rather than cloned as *C. lecturlarius* gDNA was not available. It was also more complex to purify, particularly when compared with the straightforward nickel affinity and size exclusion chromatography for the mesophilic kinases. However, the apyrase required refolding. The protocol from Valenzuela *et al.* (1998) was followed. The apyrase was clearly expressed, with bands seen in the insoluble fractions in SDS-PAGE analysis (Figure 93), particularly after concentration of the protein with TCA precipitation (Figure 94). A small amount of protein was seen in the soluble fraction so an attempt at purification by nickel affinity was made. However, the only protein recovered from this was the incorrect size and very low in concentration (Figure 95). At this point, as the kinase assay was developing well, the purification of apyrase was abandoned. It is still a viable approach if the purification of the enzyme could be simplified. This study was unable to reproduce purification of apyrase (despite correspondence with the published author) and characterisation of β 14GalT kinetic parameters. The irreproducibility of published data is currently an issue of much discussion within the scientific community and was the subject of a recent Nature article (Baker, 2016).

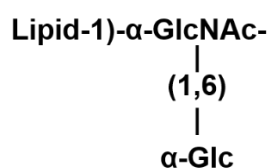
Chapter 5

Polysaccharide synthesis: glycosyltransferase purification and plasmid construction

5.1 Introduction

As previous attempts to clone and express the entire CPS coding locus from *B. pseudomallei* into *E. coli* were unsuccessful (unpublished data), this project proposed a different approach. Only the essential genes for CPS expression were selected; six enzymes for GDP-6_DHep biosynthesis (*gmhA*, *wcbL*, *wcbN*, *wcbM*, *wcbK* and *wcbJ*) and three glycosyltransferases for polymer assembly (*wcbB*, *wcbE* and *wcbH*). However, this would not be sufficient to produce an extracellular capsule. Other genes on the CPS locus encode proteins for a putative mannose priming sugar, lipid anchor biosynthesis and chaperones for transporting the polymer to the cell surface (Cuccui *et al.*, 2012). Rather than cloning all the genes from *B. pseudomallei*, this project proposes to engineer *E. coli* to attach the nascent CPS to truncated native polysaccharides. This will exploit the existing lipid anchor and transport mechanisms in *E. coli*, reducing the number of foreign genes required.

Two potential polysaccharides that could be exploited in *E. coli* are the LPS and CPS. In laboratory strain *E. coli* K-12 the LPS is truncated (Prehm *et al.*, 1976):



Lipid A anchors the polysaccharide to the cell membrane and a GlcNAc residue is available for accepting a new sugar linkage. The K12 capsule antigen comprises of an acetylated 3-Deoxy-D-manno-oct-2-ulosonic acid (KDO) moiety and two rhamnose residues (Schmidt & Jann, 1983). There are no known strains available with a truncated CPS but it may be possible to make a mutant expressing the structure:

-2)- β -KDO(7/8 Oac)-(1,5)- α -Rha-

KDO provides the link to a lipid anchor in the K12 CPS, while a rhamnose residue is available to accept a donated sugar.

In addition to the truncated polysaccharides, *E. coli* requires extra glycosyltransferases to link the nascent *B. pseudomallei* CPS to the native polysaccharide to (Figure 96).

LPS



CPS






-  additional sugar (*E. coli* or *Y. pseudotuberculosis*)
-  mannose primer (*B. pseudomallei*)
-  manno-heptose (*B. pseudomallei*)

Figure 96: Hybrid *B. pseudomallei* CPS attached to truncated *E. coli* polysaccharides. Additional sugars (grey) are added to the truncated polysaccharides in *E. coli*. The *B. pseudomallei* priming mannose (blue) is added and then the manno-heptose (orange) polymerised to form the final hybrid CPS.

Six glycosyltransferases with putative functions in O-antigen biosynthesis in different strains of *E. coli* were identified (Table 23). They are WbaD, a mannosyltransferase in O77, WbnI, an α -1,3-galactosyltransferase in O86, WbuP, a β -1,3-galactosyltransferase in O114, WffQ, a mannosyltransferase in O40, WbbK, a glucosyltransferase in O16 and WaaB, an

α -1,6-galactosyltransferase for K12 O-antigens (Yi *et al.*, 2006; Zhao *et al.*, 2007; Zhou *et al.*, 2013; Hong & Reeves, 2014; Qian *et al.*, 2014; Zhou *et al.*, 2016). A seventh glycosyltransferase, WbyB, putatively adds heptose to galactose in IA, IIA, and IVB O-antigen biosynthesis in *Y. pseudotuberculosis* (Pacinelli *et al.*, 2002).

Table 23: Additional glycosyltransferases for assembly of hybrid *B. pseudomallei* CPS in *E. coli*. Includes gene/protein name, species of origin and function (adds donor sugar to acceptor sugar).

Gene/Protein	Species of origin	Donor sugar	Acceptor sugar
WbaD	<i>E. coli</i>	β -mannose	α -GlcNAc
WbuP	<i>E. coli</i>	β -galactose	α -GlcNAc
WbnI	<i>E. coli</i>	α -galactose	β -galactose
WbbK	<i>E. coli</i>	α -glucose	α -rhamnose
WaaB	<i>E. coli</i>	α -galactose	α -glucose
WffQ	<i>E. coli</i>	β -mannose	α -galactose
WbyB	<i>Y. pseudotuberculosis</i>	heptose	α -galactose

Priming of the native truncated polysaccharides by the additional glycosyltransferases ends with either a terminal mannose (to simulate the mannose primer found in *B. pseudomallei*) or heptose, onto which the *manno*-heptose from *B. pseudomallei* can be polymerised (Figure 97).

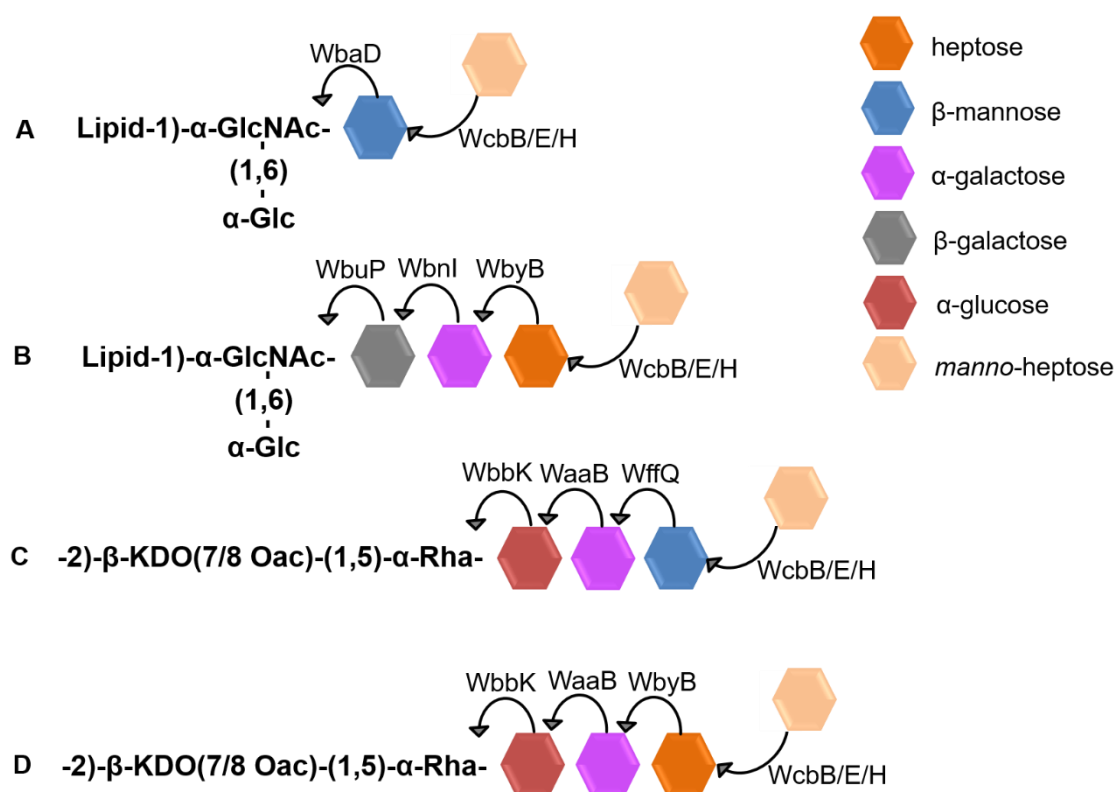


Figure 97: Biosynthesis of hybrid polysaccharides in *E. coli*. Truncated LPS (A & B) or CPS (C & D) in *E. coli* are primed with different sugars (heptose, orange; β -mannose, blue; α -galactose, purple; β -galactose, grey; α -glucose, red) by glycosyltransferases from different *E. coli* strains (WbaD, WbuP, Wbni, WbbK, WaaB, WffQ) or *Y. pseudotuberculosis* (WbyB). The manno-heptose from *B. pseudomallei* (light orange) can then be polymerised to produce the final hybrid polysaccharide.

This approach removes the need to include the *B. pseudomallei* glycosyltransferase responsible for attaching mannose to a lipid, as the native *E. coli* lipid moiety would be primed with either α -GlcNAc or KDO- α -Rhamnose (Figure 97). For hybrid polysaccharides B and D with a terminal heptose (Figure 97), only the *B. pseudomallei* glycosyltransferase that polymerises manno-heptose is required. For hybrid polysaccharides A and C with a terminal mannose (Figure 97), two *B. pseudomallei* glycosyltransferases for linking manno-heptose to the mannose primer and for manno-heptose polymerisation is required.

To achieve this, three plasmids need to be co-expressed in *E. coli* for the expression of the six GDP-6dDHep biosynthetic genes from *B. pseudomallei*

(*wcbJ*, *wcbK*, *wcbM*, *wcbL*, *gmhA* and *wcbN*), a putative acetyltransferase (*wcbI*) and the glycosyltransferases from *B. pseudomallei* (*wcbB*, *wcbE* and *wcbH*) and additional glycosyltransferases (either *wbaD* only, *wbuP/wbnI/wbyB*, *wbbK/waaB/wffQ* or *wbbK/waaB/wbyB*) (Figure 98).

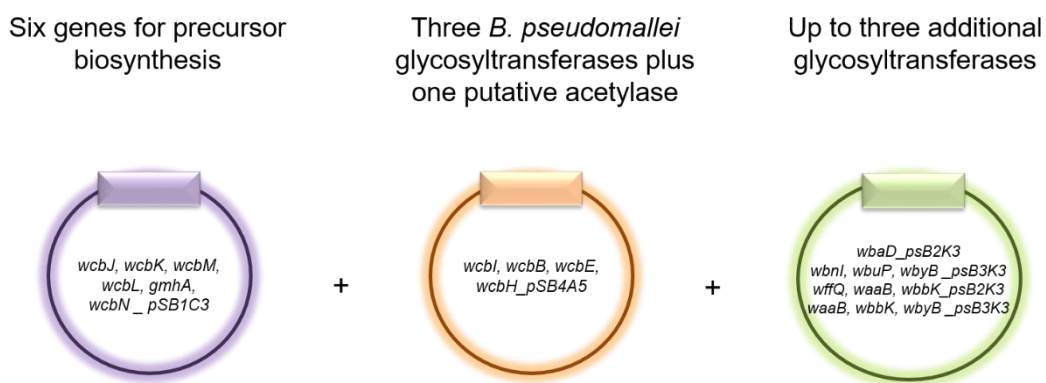


Figure 98: Co-expression of three plasmids in *E. coli* for the synthesis of hybrid CPS. The first plasmid (purple) is chloramphenicol resistant and encodes the six GDP-6dHep biosynthesis genes. The second plasmid (orange) is ampicillin resistant and encodes a putative acetyltransferase and three glycosyltransferases from *B. pseudomallei*. The third plasmid (green) encodes additional glycosyltransferases from either *Y. pseudotuberculosis* (*wbyB*) or *E. coli* and is kanamycin resistant.

Each vector requires a different antibiotic resistance to ensure all three plasmids are retained by *E. coli*. They must also have different origins of replication, firstly to ensure each plasmid is expressed and secondly because there are different copy requirements (Table 24). For the first plasmid expressing the six biosynthetic genes, a high copy number is desirable to maximise biosynthesis of the sugar nucleotide precursor. For the *B. pseudomallei* glycosyltransferases a low copy number would be beneficial to minimise cytotoxic effects. A low copy number is also desirable for the additional glycosyltransferases for the same reason and this can be achieved by two kanamycin-resistant plasmids (though one is inducible to high copy number in the presence of IPTG) (Table 24).

Table 24: Vector properties for co-expression of three plasmids in *E. coli*. All vectors deposited in the BioBricks™ Registry of Standard Biological Parts.

Vector	Antibiotic resistance	Origin of replication	Copy number	
pSB4A5	ampicillin	pSC101	low	~5
pSB1C3	chloramphenicol	pUC19-derived pMB1	high	100-300
pSB2K3	kanamycin	oriS	low (inducible high)	<10 (>100 with IPTG)
pSB3K3	kanamycin	p15A pMR101- derived	low / medium	20-30

There is evidence to suggest that 2-O-acetylation of the *manno*-heptose moiety of *B. pseudomallei* CPS is essential for eliciting protective antibody response (Marchetti *et al.*, 2015). Therefore, it is desirable to acetylate the hybrid CPS for use in *B. pseudomallei* vaccine development. *WcbI*, encoded on the CPS locus, is a putative acetyltransferase (Vivoli *et al.*, 2014). The *wcbI* gene can be included on the plasmid encoding *B. pseudomallei* glycosyltransferases to search for evidence of acetylation.

5.1.1 Aims

There were two aims for this chapter. The first was to purify *B. pseudomallei* glycosyltransferases *WcbB*, *WcbE* and *WcbH* with the intention of determining their kinetic parameters using the assay developed in Chapter 4. The second aim was to construct two plasmids (one encoding *B. pseudomallei* glycosyltransferases and the other encoding additional glycosyltransferases from *E. coli* and *Y. pseudotuberculosis*) for co-expression in *E. coli* with the GDP-6dDHep biosynthetic plasmid assembled in Chapter 2.

The purpose of these aims is to contribute to the development of a novel platform for the biosynthesis of bespoke polysaccharides. Glycosyltransferase characterisation is necessary for the design of novel systems for polysaccharide

biosynthesis and information is relatively scarce in current literature. It is also a demonstration of the ability to engineer *E. coli* for the expression of non-native polysaccharides.

5.2 Materials and Methods

All reagents were purchased from Sigma-Aldrich unless otherwise stated.

5.2.1 Purification of glycosyltransferases

5.2.1.1 Gene synthesis for protein purification

Genes for purification of *B. pseudomallei* glycosyltransferases *wcbB*, *wcbE* and *wcbH* were amplified by PCR (section 2.2.3.4) from plasmids kindly provided by Dr. Harmer, University of Exeter. The primers (Appendix 1) contained flanking sequences for LIC (Savitsky *et al.*, 2010) and genes were sub-cloned into the pNIC28-Bsa4 vector.

5.2.1.2 Bacterial cultures and detergents

The glycosyltransferases were prepared and stored as described in section 2.2.2.1. *E. coli* cultures expressing glycosyltransferases were grown in LB plus IPTG-induction and ZYM-5052 auto-induction media with the appropriate antibiotics. Small scale cultures (section 2.2.4.5) were prepared for mini nickel column purification (section 5.2.1.3). Large scale cultures (section 3.2.1) were grown for gravity flow and ÄKTA nickel affinity purification (section 5.2.1.4).

After sonication, detergents were added to lysed cultures for extracting the glycosyltransferases. The lysate was incubated with the detergent at room temperature for 20 min with mild agitation. The detergents tested were sodium lauroyl sarcosinate (Sarkosyl), 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), Triton X-100 and octyl β -D-glucopyranoside (octyl glucoside). All detergents were purchased from Sigma-Aldrich.

5.2.1.3 Mini nickel column purification

His-tagged proteins extracted from small cultures were purified using mini-nickel columns and microcentrifuge. Firstly, the columns were prepared by washing with 650 μ L 1M NaOH (1,800 xg, 2 min). Next, the columns were washed twice with 650 μ L buffer (20 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole pH 8.0) (1,800 xg, 2 min). Finally, if testing detergents, the columns were washed twice with 650 μ L buffer (20 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole pH 8.0) plus detergent at desired concentration (1,800 xg, 2 min). Once the detergent wash was complete, 650 μ L samples were added to the columns by gentle pipetting. The samples were centrifuged at 640 xg for a minimum 10 min. If the flow rate was slow, the samples were spun for a longer duration at the same speed. Bound protein was washed three times with buffer (20 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole pH 8.0), plus detergent if required, for 5 min at 1,800 xg. After washing, a final dry spin at 1,800 xg for 2 min removed any excess buffer and detergent. The columns were then placed into fresh 1.5 mL microcentrifuge tubes. Proteins were eluted into 650 μ L aliquots elution buffer (20 mM Tris-HCl, 500 mM NaCl, 250 mM imidazole pH 8.0) plus detergent if desired. The columns were spun at 1,800 xg for 3 min for the first aliquot and 5 min for the second. Columns were cleaned twice with 650 μ L elution buffer without detergent at 1,800 xg for 3 min, then 0.05 % sodium azide for 2 min. The columns were stored in 0.05 % sodium azide at 4 °C to prevent microbial growth. Chromatography with the mini nickel columns could be performed at either room temperature or 4 °C. An aliquot was taken from each wash, flow through and elution for analysis by SDS-PAGE.

5.2.1.4 ÄKTA and gravity flow nickel affinity chromatography

Extractions from larger cultures were purified using the Äkta and gravity flow nickel columns as details in sections 3.2.2.1 and 3.2.2.2. Detergents were added to the wash and elution steps (with gravity flow columns only) where indicated.

5.2.1.5 Differential scanning fluorimetry

WcbB was analysed with differential scanning fluorimetry (DSF) using a pre-prepared set of 24 substrate and buffer conditions. WcbB (purified without glycerol) was diluted to 0.133 mg mL⁻¹, with 5,000x SYPRO® Orange in dH₂O. A 15 µL aliquot of the protein and dye mix was added to each well of a 48-well U-bottomed PCR plate (Applied Biosystems) on ice. A 96 well-plate containing the pre-prepared substrates and buffers was centrifuged at 1,500 xg for 5 min. A 5 µL aliquot of each mix was taken from the plate and added in duplicate to the protein samples. The plate was sealed with optically transparent sealing tape (Applied Biosystems) and centrifuged for 2 min at 300 xg. The samples were run on a StepOne Real-Time PCR machine using the StepOne v2.2.2 software (Applied Biosystems). The temperature of the samples was raised from 25 °C to 99 °C in one 50 min cycle.

5.2.2 Additional biosynthetic constructs

5.2.2.1 Gene synthesis for co-expression studies

For construction of additional plasmids for co-expression in *E. coli*, genes were synthesised as previously described (section 2.2.3.1) with flanking restriction sites for BioBricks™ 3A assembly. The glycosyltransferase genes synthesised were: *wcbB*, *wcbE* and *wcbH* from *B. pseudomallei*, *waaB*, *wbaD*, *wbbK*, *wbnl*,

wbuP and *wffQ* from *E. coli* and *wbyB* from *Y. pseudotuberculosis*. The putative acetyltransferase gene, *wcbI*, was also synthesised (Appendices 2-5).

5.2.2.2 BioBricks™ 3A assembly of additional plasmids

Additional plasmids for co-expression in *E. coli* were created using BioBricks™ 3A assembly as previously described in section 2.2.3. Vector backbones used in the assembly were from the BioBricks™ 2012 kit; pSB4A5, pSB2K3 and pSB3K3. The constitutive promoter (BBa_J23118) and double terminator (BBa_B0015) were also used. Final constructs were: p-*wbaD*-tt_2K3, p-*wbnI-wbuP-wbyB*-tt_2K3; p-*wffQ-waaB-wbbK*-tt_2K3, p-*waaB-wbbK-wbyB*-tt_3K3, and p-*wcbI-wcbB-wcbE-wcbH*-tt_4A5 (Appendices 2-5).

5.2.2.3 Co-expression in *E. coli*

E. coli DH5α cells were transformed with plasmids as described in section 2.2.3.2. For transformations with two or three plasmids, the same protocol was followed using 2 L of each plasmid. After heat shock, the *E. coli* cells were incubated without antibiotics in LB for at least 1 hr at 37 °C then plated onto agar containing the appropriate combination of antibiotics. The concentration of amp, chl and kan used varied as indicated. Plates were incubated for at least 18 h at 37 °C. Single colonies were collected and grown in LB with appropriate antibiotics overnight at 37 °C. Plasmids were extracted using the kit and PCR and RD analysis were performed as previously described in sections 2.2.3.4 and section 2.2.3.5. The DNA was visualised according to section 2.2.3.7.

5.2.2.4 Enzyme-linked immunosorbent assay

An enzyme-linked immunosorbent assay (ELISA) was used to screen for synthesis of the *B. pseudomallei* CPS in *E. coli* transformed with three plasmids. To prepare extracts for the assay, the transformed *E. coli* were grown overnight in 5 mL LB with antibiotics ($30 \mu\text{g mL}^{-1}$ Amp, $30 \mu\text{g mL}^{-1}$ Chl and $30 \mu\text{g mL}^{-1}$ Kan unless otherwise stated). A 1 mL aliquot was taken from each culture and the cells were lysed by heating to $80 \text{ }^{\circ}\text{C}$ for 10 min.

To prepare the ELISA (Figure 99), a 96-well plate was coated with $100 \mu\text{L}$ per well $5 \mu\text{g mL}^{-1}$ 4VIH12 monoclonal antibody (mAb) and incubated over night at $4 \text{ }^{\circ}\text{C}$. The mAb was removed and the plate washed three times with phosphate buffered saline (PBS), pH 7.4 (GE Healthcare) plus 0.05 % Tween-20 and dried by patting onto absorbent paper.

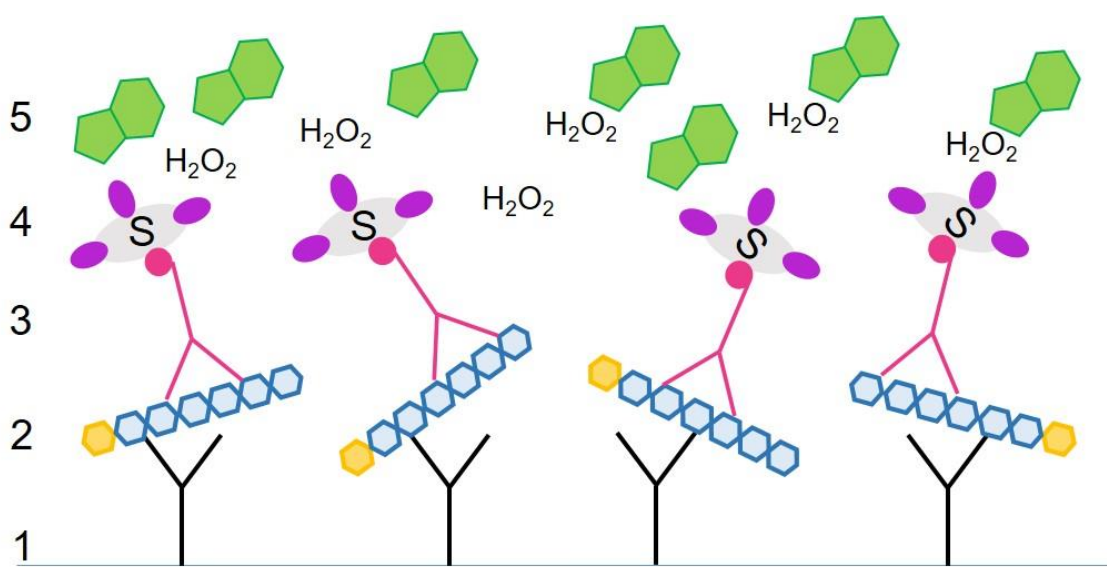


Figure 99: Enzyme-linked immunosorbent assay for detecting *B. pseudomallei* CPS. 1: 4VIH12 monoclonal antibodies (black) are adsorbed onto the surface of a 96-well plate. 2: The antibodies bind CPS (blue and orange hexagons). 3: Bound CPS is detected by biotinylated 4VIH12 monoclonal antibodies (pink). 4: streptavidin (grey oval with S) with horse radish peroxidase (HRP; purple ovals) bind to biotinylated antibodies. 5: ATBS reagent is oxidised by HRP in the presence of peroxide (H_2O_2) to produce blue/green colour which can be measured at 414 nm absorbance.

The plate was blocked with 200 μL per well PBS (pH 7.4) plus 2 % (w/v) skimmed milk powder for one hour at 37 °C. The plate was then washed and dried as previously described.

All wells in columns 2-12 were filled with 100 μL PBS (pH 7.4). A 200 μL aliquot of each standard and sample was then added to the first column. Wells A1 and B1 contained a standard of 100 ng mL^{-1} *Burkholderia thailandensis* E555 CPS (prepared at Dstl) in PBS (pH 7.4). Wells C1 – F1 contained the neat heat-lysed samples of *E. coli* co-expressing three plasmids. The standards and samples were diluted 1:1 (moving 100 μL each time) across the plate to column 11. A 100 μL aliquot was discarded from column 11 to ensure a consistent volume across the plate. The final column of wells remained blank (PBS, pH 7.4) as a negative control. The plate was incubated at 37 °C for 1 h, then washed and dried as previously described.

To each well, 100 μL 5 $\mu\text{g mL}^{-1}$ biotinylated 4VIH12 mAb was added and the plate incubated at 37 °C for 1 h. The plate was washed and dried as before. Next, 100 μL streptavidin peroxidase was added to each well and the plate was again incubated at 37 °C for 1 h. The plate was washed as previously described, with six washes instead of three and dried. During incubation, detection reagent was prepared (0.3 mg mL^{-1} 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS), 100 mM citric acid (pH 4.4). To this, 10 μL 30 % (v/v) hydrogen peroxide (H_2O_2) was added per 10 mL ABTS reagent immediately prior to use.

Finally, 100 μL ABTS reagent was added to each well and the plate incubated at room temperature for 20 min. The plate was then read at absorbance 414 nm.

5.2.2.5 Immunoblot assay

Samples of *E. coli* transformed with three plasmids were prepared as described in section 5.2.2.3. For the immunoblot assay (Figure 100), nitrocellulose membrane (GE Healthcare) was cut to fit a circular petri dish and a 3 x 3 grid was drawn in pencil. A 2 μ L aliquot of sample or standard was carefully dropped into the centre of each section of the grid (Table 25);

Table 25: Samples for immunoblot assay for *B. pseudomallei* CPS expression in *E. coli* DH5 α .

Grid	Plasmid 1	Plasmid 2	Plasmid 3
1	p-wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_1C3	p-wcbI-wcbB-wcbE-wcbH-tt_4A5	p_wbaD_tt_3K3
2	p-wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_1C3	p-wcbI-wcbB-wcbE-wcbH-tt_4A5	p_wbnI_wbuP_wbyB_tt_2K3
3	p-wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_1C3	p-wcbI-wcbB-wcbE-wcbH-tt_4A5	p_wffQ_waaB_wbbK_tt_2K4
4	p-wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_1C3	p-wcbI-wcbB-wcbE-wcbH-tt_4A5	p_waaB_wbbK_wbyB_tt_3K3
5	Wild type <i>B. thailandensis</i> E555 expressing the <i>B. pseudomallei</i> CPS		
6	LB media only		
7	Wild type <i>E. coli</i> DH5 α		
8	p-wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_1C3	-	-
9	p-wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_1C3	p-wcbI-wcbB-wcbE-wcbH-tt_4A5	-

The membrane was thoroughly dried at room temperature, then blocked with 5 mL PBS (pH 7.4) plus 5 % (w/v) skimmed milk at room temperature for 1 hr. After blocking, the membrane was washed with PBS (pH 7.4) plus 0.05 % (v/v) Tween-20. Next, 10 μ L sera obtained from mice challenged with *B. pseudomallei* CPS (with and without Crm-conjugate; kindly gifted by Dstl) was diluted in 1 mL PBS (pH 7.4). This was coated onto the membrane and incubated at room temperature for 1 h. The membrane was washed as before, then incubated with HRP-conjugated anti-mouse mAb (either anti-IgG or anti-IgM) (Promega) at room temperature for 1 h. Finally, the membrane was thoroughly washed and stained using a colorimetric Pierce DAB (3,3'-diaminobenzidine) Substrate kit (Thermo

Fisher). The DAB produces a dark brown colour when oxidised by the HRP conjugated to the secondary anti-mouse antibodies

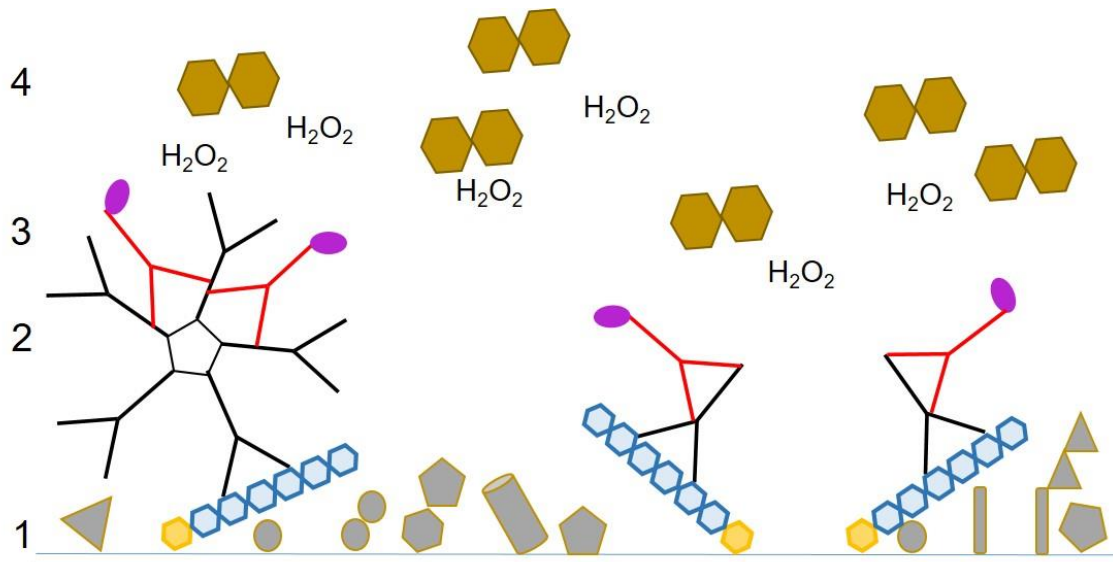


Figure 100: Immunoblot for detecting *B. pseudomallei* CPS. 1: Whole cell extract (grey shapes) potentially expressing CPS (blue and orange hexagons) from transformed *E. coli* is adsorbed onto nitrocellulose membrane. 2: Antibodies (IgM, black pentamer; IgG, black monomer) from mouse sera challenged with *B. pseudomallei* CPS bind to adsorbed CPS. 3: The mouse antibodies are detected with HRP-conjugated (purple ovals) anti-mouse antibodies (red). 4: DAB reagent is oxidised by HRP in the presence of peroxide (H_2O_2) to produce a brown colour.

5.3 Results

5.3.1 Purification of glycosyltransferases

5.3.1.1 High-throughput screening for purification conditions of WcbH

The mini-nickel column purifications were useful for high-throughput screening of different conditions. Here, data for optimising WcbH expression is presented.

The first step was to determine whether detergent could improve purification of the protein. 0.5 % (w/v) Sarkosyl, 10 % (w/v) CHAPS and 10 % (v/v) Triton X-100 were tested with WcbH (Figure 101):

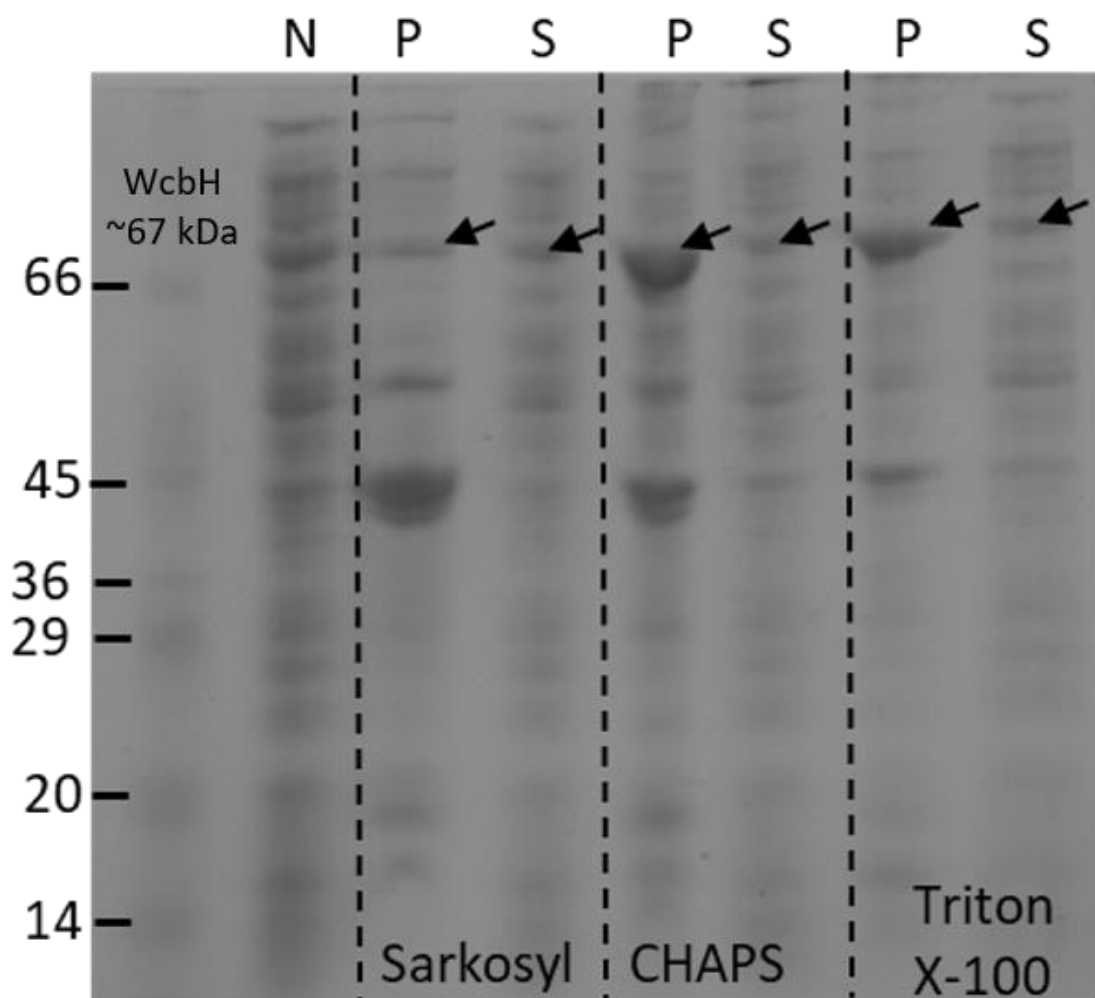


Figure 101: Testing detergents for purification of WcbH. SDS-PAGE showing (N) neat supernatant; (P) pellet and (S) supernatant of mini-nickel column purification with 0.5 % (w/v) Sarkosyl; 10 % (w/v) CHAPS; 10 % (v/v) Triton X-100. Arrows indicate bands of expected size.

Initial attempts to purify WcbH yielded very weak bands (Figure 101). There was little protein in the neat supernatant due to no induction of protein expression using ZYM-5052 auto-induction media. Most of the protein that was expressed was lost in the pellet. As WcbH is putatively membrane-associated it is likely to require presence of a membrane to maintain structure and stability (Breton et al., 2006; Albesa-Jové et al., 2014). Furthermore, over-expression of proteins in *E. coli* can result in their precipitating out in inclusion bodies, requiring resolubilisation and refolding (Singh & Panda, 2005).

To combat these challenges, three surfactants (Sarkosyl, CHAPS and Triton X-100) were tested to determine whether they improved WcbH purification. Although the bands were still faint, and most of the protein was lost using CHAPS and Triton X-100, the bands seen in the Sarkosyl fractions were evenly distributed between the pellet and supernatant (Figure 101).

Next, IPTG-induction of WcbH expression was tested to determine whether the quantity of protein produced could be improved (Figure 102).

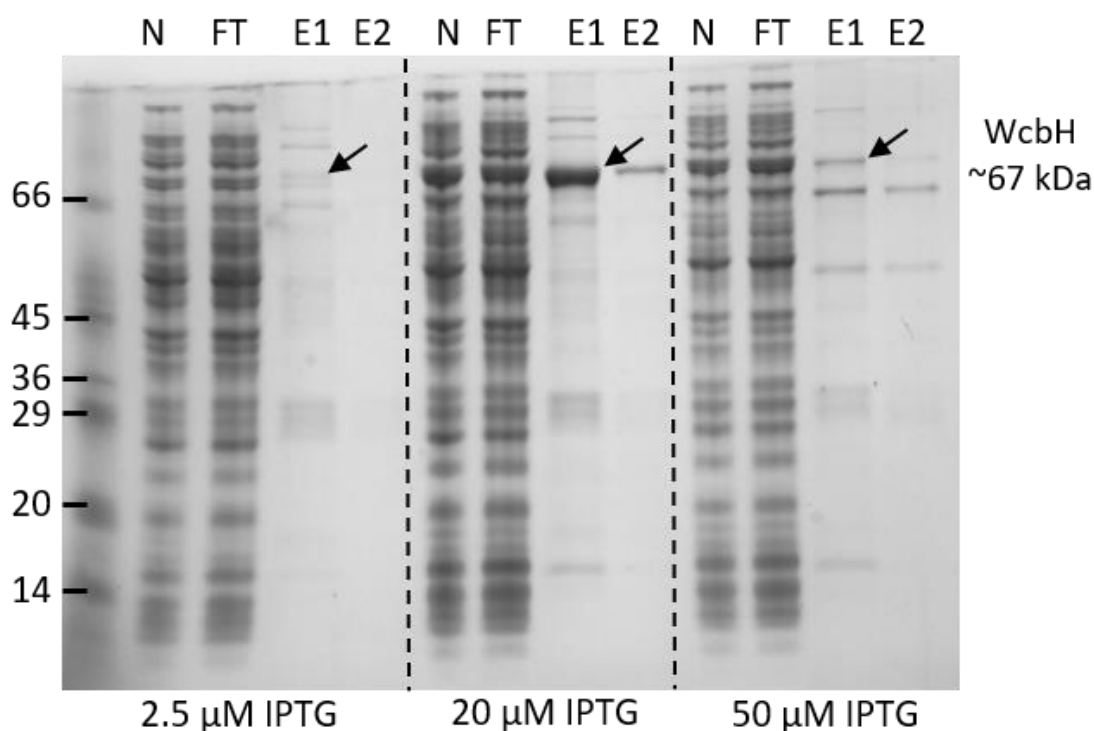


Figure 102 Mini nickel column high-throughput screening for IPTG-induction of WcbH expression. Purification included 0.25 % Sarkosyl. SDS-PAGE showing (A) Screening concentration of IPTG for induction of protein expression. Concentrations tested were 2.5 μM, 20 μM and 50 μM. Arrows indicated band of expected size.

For WcbH, the optimal concentration of IPTG to induce protein expression was 20 μM as this produced the most protein (Figure 102). Lower concentrations were not sufficient to increase protein expression whereas it is likely higher concentrations caused over-expression of the protein leading to cytotoxicity.

5.3.1.2 Gravity flow column purification

The glycosyltransferases were expressed with 20 μM IPTG induction and 0.25 % Sarkosyl for purification on a larger-scale using gravity flow nickel affinity chromatography (Figure 103).

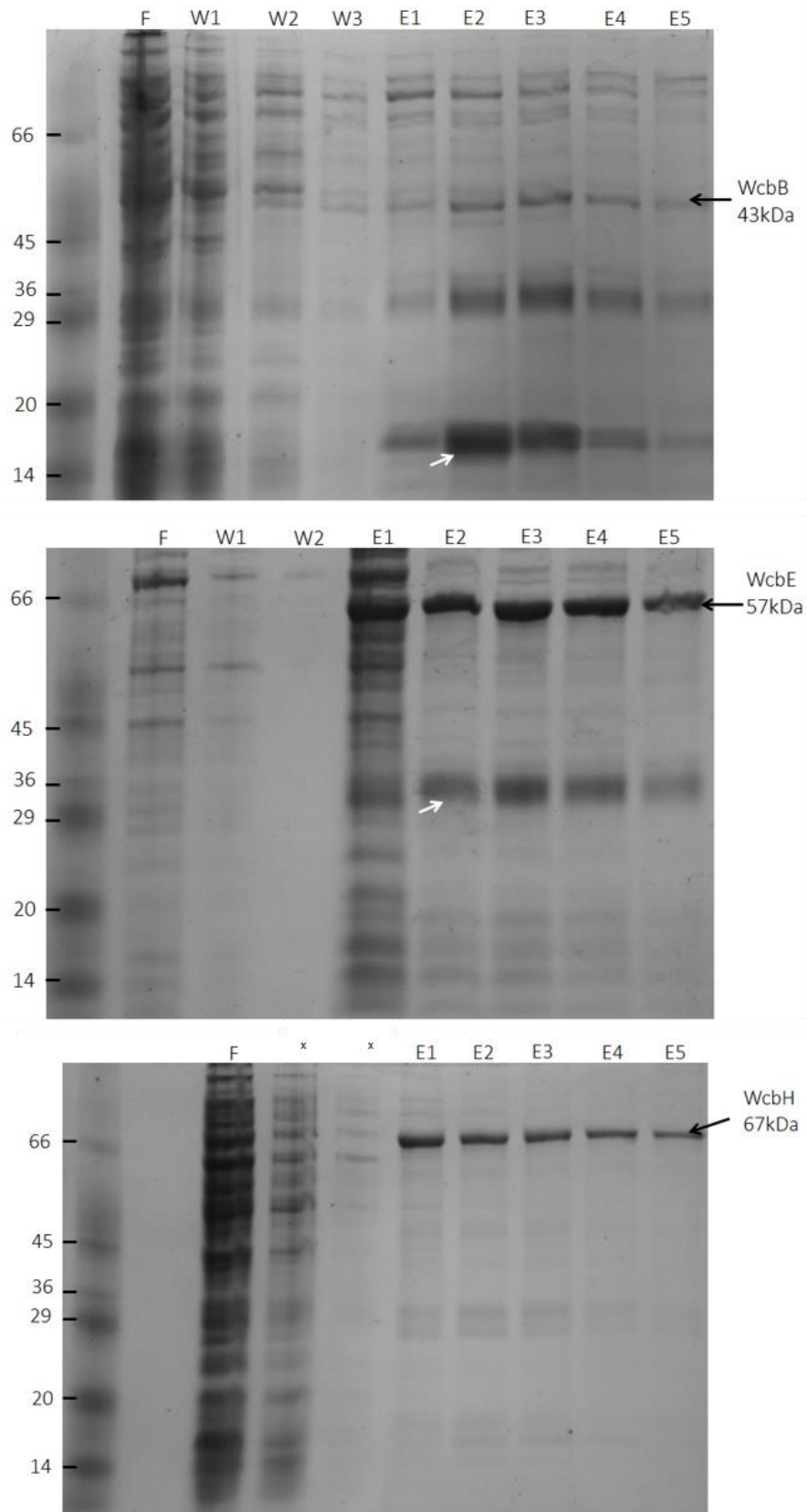


Figure 103: Gravity-flow nickel column purification of WcbB, WcbE and WcbH. Using 20 μ M IPTG-induction and 0.25 % Sarkosyl for purification. SDS-PAGE showing (F) = flow through; (W1-3) = wash 1-3; (E1-5) elutions 1-5. Black arrow indicates band of correct size. White arrow indicates large band at incorrect size. x indicates wrong protein loaded.

The optimised conditions of IPTG and Sarkosyl enabled purification of WcbH by nickel affinity chromatography. As detergent was used this was performed by gravity flow to prevent damage to the Äkta. The elutions in show very clean samples after nickel affinity purification for WcbH. However, for WcbB and WcbE the conditions were sub-optimal. Strong bands much lower than the expected size indicate the proteins were denaturing during the purification process. Addition of protease inhibitors to the purification of WcbB and WcbE prevented the denaturation and all three proteins could be further purified by size exclusion chromatography.

5.3.1.3 Purification by size exclusion chromatography

Elutions for each protein from nickel affinity columns were pooled and further purified by size exclusion chromatography. No detergent was used in the elution buffer for size exclusion as it may interfere with downstream applications (Figure 104).

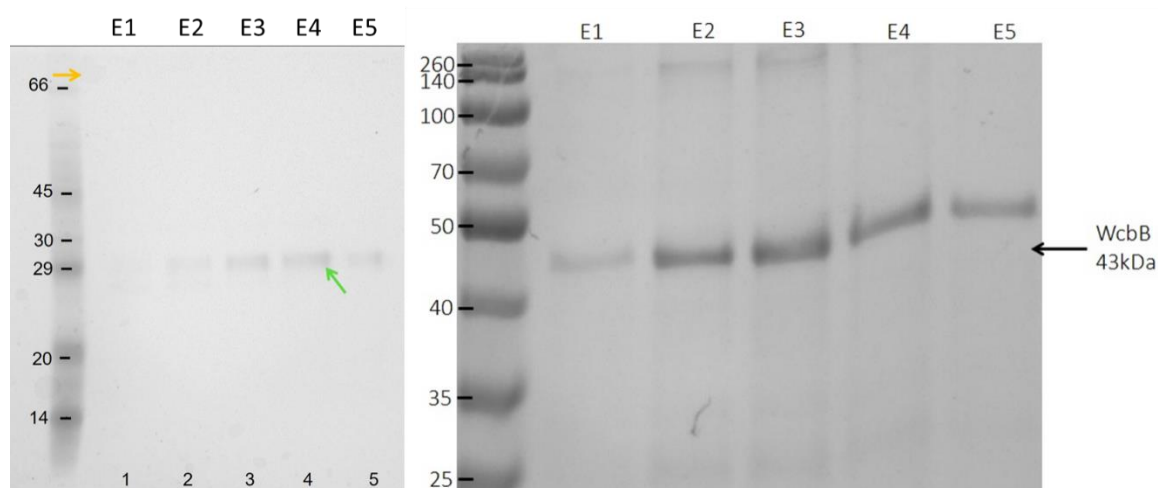


Figure 104: SDS-PAGE of size exclusion chromatography of WcbH and WcbB in 10 mM HEPES, pH 8.0, 500 mM NaCl with no detergent. WcbH and WcbB extracted from *E. coli* DH5 α in LB, induced with 2.5 μ M IPTG and purified by gravity nickel column chromatography with 0.25 % (w/v) Sarkosyl. (E1-5) = elutions 1-5.

The SDS-PAGE for WcbH purified by size exclusion chromatography shows that WcbH was lost. There is a contaminating protein, which can be seen in SDS-PAGE for both nickel affinity and size exclusion (Figure 104). However, WcbB purified at the correct size following size exclusion chromatography and with no contaminants (Figure 104). It was therefore analysed by differential scanning fluorimetry (DSF).

5.3.1.4 Protein stability analysed by differential scanning fluorimetry

WcbB purified by nickel affinity and size exclusion chromatography was mixed with a set of 24 substrates and buffer conditions for DSF analysis (Figure 105).

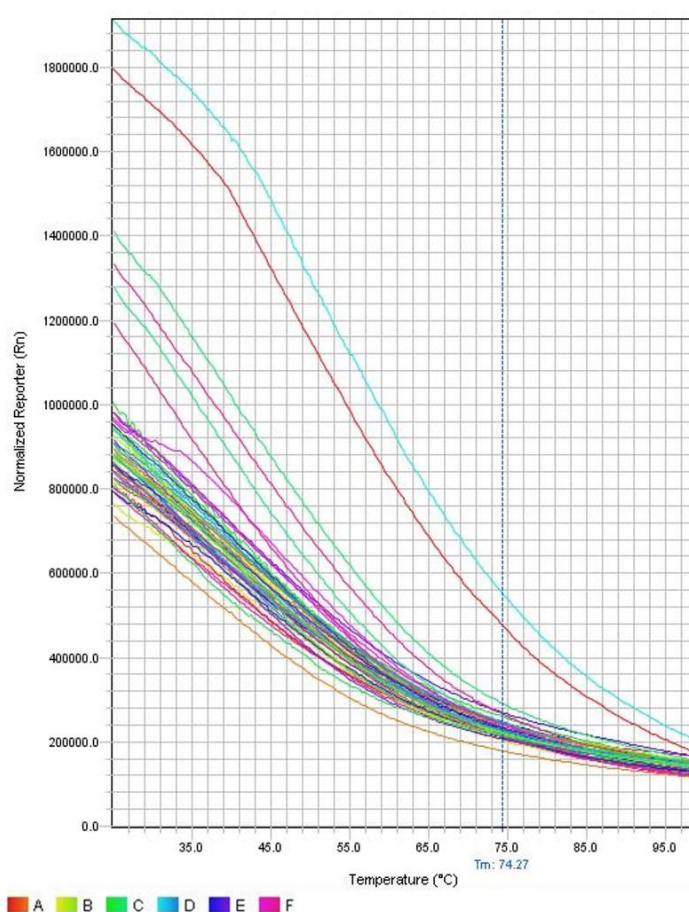


Figure 105: Differential scanning fluorimetry curve for WcbB. 0.133 mg mL⁻¹ WcbB was mixed with 5000x SYPRO® Orange and added to 24 solutions containing different substrates and buffer conditions. Stability of the protein in temperatures ranging from 25 °C to 100°C was monitored.

Typically, in a DSF assay the normalised reporter signal starts low, then peaks when the protein becomes unstable and unfolds, exposing more of the bound SYPRO® Orange dye. Bound to a substrate in ideal buffer conditions proteins become more stable and the peak 'shifts' to a higher temperature. For WcbB, the signal starts high. This strongly suggests that the protein was already denatured and the purification was unsuccessful.

None of the three *B. pseudomallei* glycosyltransferases were successfully purified.

5.3.2 Additional plasmid construction

Construction of the additional plasmids for CPS assembly followed the BioBricks™ method that successfully constructed the biosynthetic plasmid for GDP-6dbHep. The pSB4A5, pSB2K3 and pSB3K3 vectors from the Registry of Standard Biological Parts were chosen and five new plasmids were constructed.

5.3.2.1 *B. pseudomallei* glycosyltransferases

The *B. pseudomallei* glycosyltransferase genes, *wcbB*, *wcbE* and *wcbH* and the putative acetyltransferase gene *wcbI* were assembled onto the low copy number, ampicillin-resistant pSB4A5 vector backbone (Figure 106). The constitutive BBa_J23118 promoter and BBa_B0015 double terminator sequences were used to complete the operon.

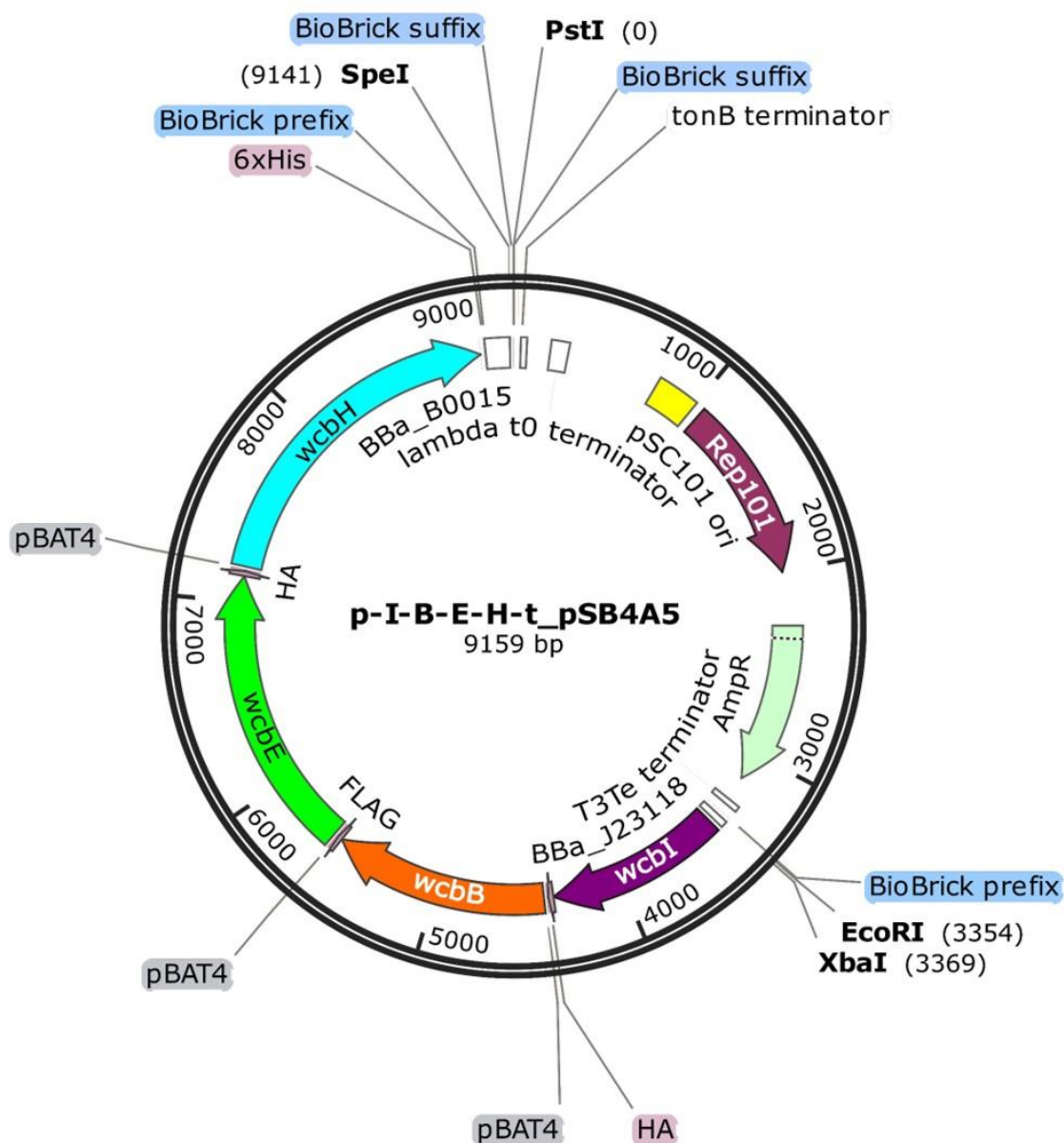


Figure 106: Plasmid map for *wcbI*, *wcbB*, *wcbE* and *wcbH* operon in BioBricks™ vector pSB4A5.

5.3.2.2 Additional glycosyltransferases

The additional glycosyltransferases from *E. coli* and *Y. pseudotuberculosis* were cloned onto either the pSB2K3 or pSB3K3 vector backbone (Figure 107). They are kanamycin resistant. Copy number for pSB2K3 is low (less than 10) unless induced by IPTG (more than 100) and is low/medium for pSB3K3 (20-30).

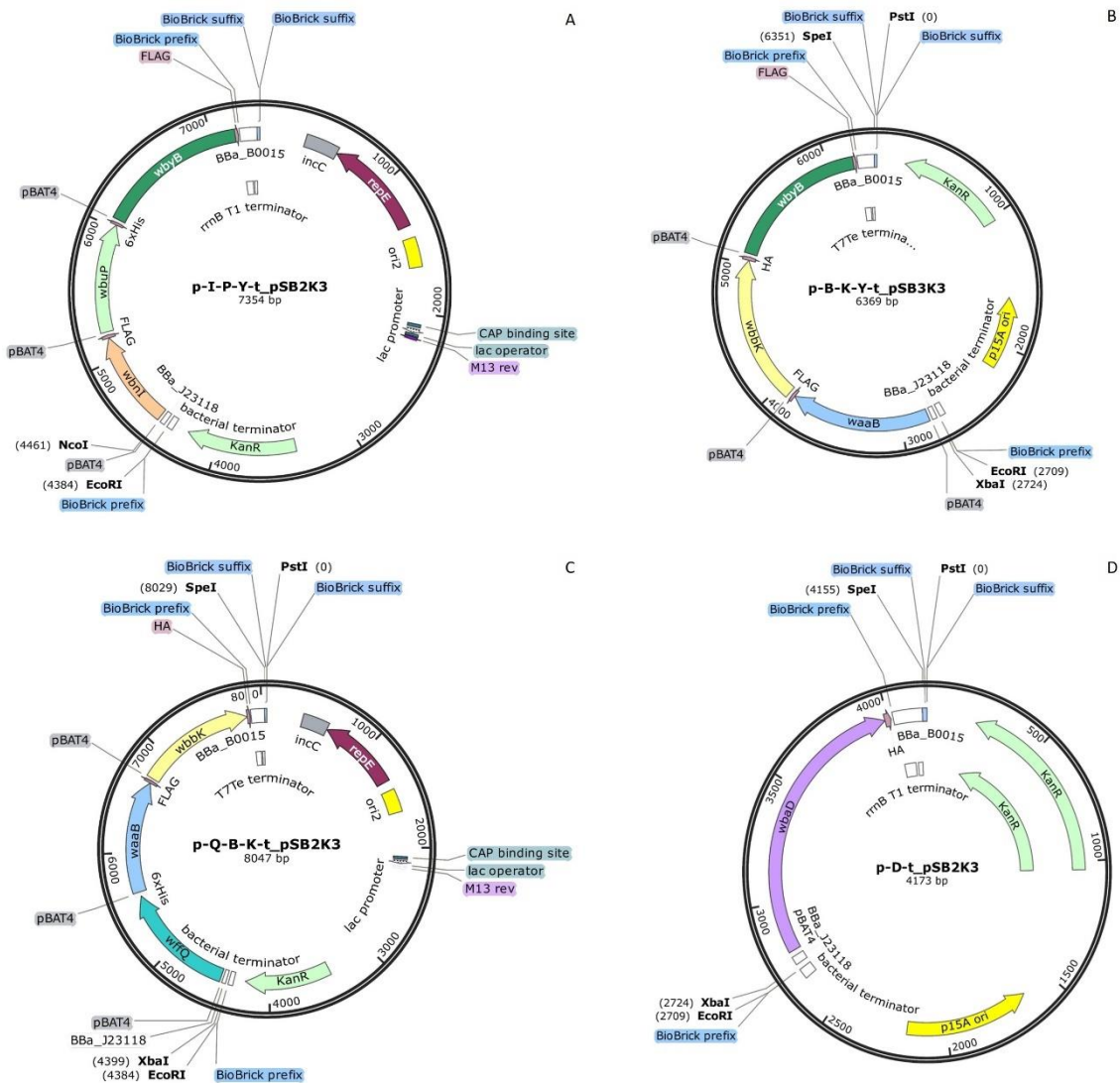


Figure 107: Plasmid constructs for expression of additional glycosyltransferases in *E. coli*. (A) *wbnI-wbuP-wbyB_pSB2K3*; (B) *waaB-wbbK-wbnI_pSB3K3*; (C) *wffQ-waaB-wbbK_pSB2K3* and (D) *wbaD_pBB2K3*.

5.3.3 Immunoblot and ELISA

The immunoblot assay of *E. coli* DH5 α cultures co-transformed with different combinations of the plasmids (Table 25) were all negative for *B. pseudomallei* CPS antigen. The positive control *B. thailandensis* E555 gave a positive indication for antigen, confirming the immunoblot assay worked (Figure 108). This suggests none of the *E. coli* cultures were expressing CPS with antigen recognisable by the sera from mice challenged with *B. pseudomallei*.

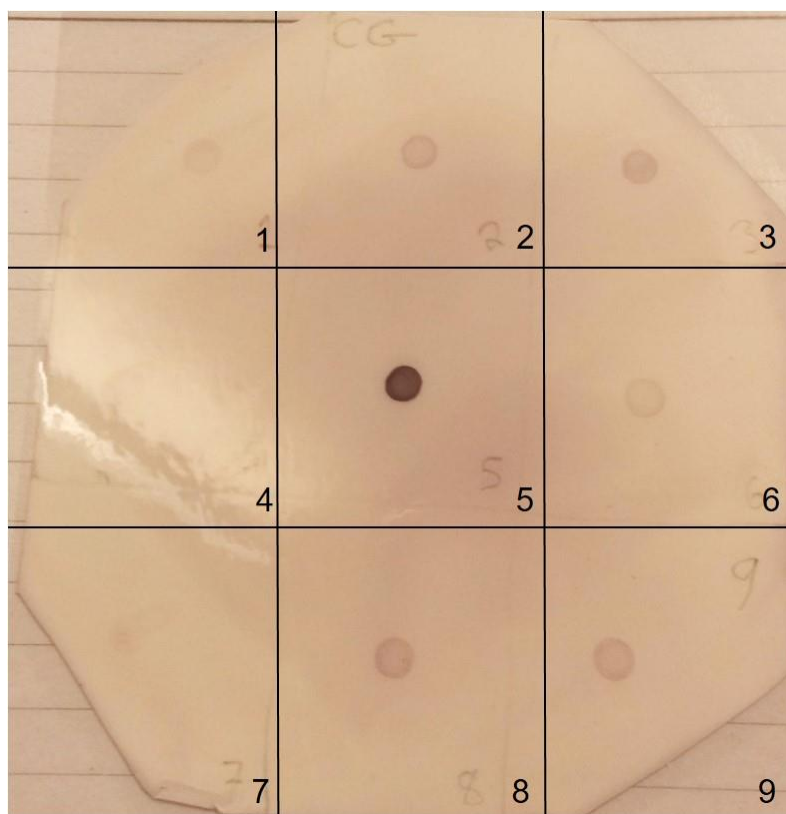


Figure 108: Immunoblot of IgG positive sera from *B. pseudomallei*-challenged mice against extracts from *E. coli* 1-4: co-expressing p-*wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_1C3*, p-*wcbI-wcbB-wcbE-wcbH-tt_4A5* and one of the four additional glycosyltransferase plasmids; 5: *B. thailandensis* positive control; 6-9: negative controls. (1) p-*waaB-wbbK-wbyB-tt_3K3*; (2) p-*wbaD-tt_2K3*; (3) p-*wbnI-wbuP-wbyB-tt_2K3*; (4) p-*wffQ-waaB-wbbK-tt_2K3*; (5) wild type *B. thailandensis* E555; (6): LB media; (7) wild type *E. coli* DH5 α , (8) *E. coli* expressing p-*wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_1C3* and (9) *E. coli* expressing p-*wcbJ-wcbK-wcbM-wcbL-gmhA-wcbN-tt_1C3* and p-*wcbI-wcbB-wcbE-wcbH-tt_4A5*.

ELISA analysis of *E. coli* DH5 α co-expressing different combinations of the plasmids was also negative, confirming there was no antigen recognisable to the anti-CPS antibody to bind to. The standard control worked as expected.

5.4 Discussion

In this chapter, the plasmid expressing six genes for GDP-6dbHep biosynthesis was co-expressed with two additional plasmids, one expressing the *B. pseudomallei* CPS glycosyltransferases (and putative acetyltransferase) and one expressing additional glycosyltransferases to facilitate assembly of the nascent CPS onto truncated native *E. coli* polysaccharides.

Leloir glycosyltransferases are essential for the assembly of polysaccharides as they facilitate the creation of glycosidic bonds using a nucleotide sugar donor and an acceptor molecule. The reaction releases a nucleotide diphosphate and the sugar is linked to the donor. Unfortunately, due to difficulties with purification of glycosyltransferases and the number of potential isomers in a reaction involving saccharides, there is a shortage of characterised glycosyltransferases in the literature. This includes the three putative glycosyltransferases involved in *B. pseudomallei* CPS biosynthesis, WcbB, WcbE and WcbH. Whilst their functions have been speculated (mannosyltransferase to prime the lipid anchor with a mannose, heptosyltransferase to add a *manno*-heptose to the primer and a polymerase to assemble the *manno*-heptose O-antigen), they have not been determined. The first requirement for characterisation of an enzyme is to obtain a pure sample.

5.4.1 Glycosyltransferase purification

Several attempts at purifying WcbB, WcbE and WcbH glycosyltransferases were made. Firstly, using small-scale nickel affinity chromatography to determine optimum conditions for expression and purification. The glycosyltransferases expressed very well under low-level IPTG-induction (Figure 102). Without IPTG induction, no protein was expressed. Whereas at higher IPTG concentrations,

over-expression of the glycosyltransferase and its activity likely became cytotoxic.

During initial purification attempts, most of the protein precipitated and was recovered in the pellet rather than solubilised in the supernatant (Figure 101). This probably occurred because the glycosyltransferases are membrane-associated and require the presence of a surface to remain stable (Breton *et al.*, 2006; Albesa-Jové *et al.*, 2014). Assembly of the CPS occurs on the inner leaflet of the inner membrane. Association with the membrane may help localise the glycosyltransferases to improve efficiency of capsule expression. In an attempt to stabilise the enzymes, different detergents were tested during the purification process. (Figure 103). Sarkosyl appeared to reduce the amount of unfolded protein and was used in subsequent purifications. However, despite clean bands obtained after both nickel affinity and size exclusion chromatography, no functional enzyme was obtained. The removal of the Sarkosyl to prevent it damaging the size exclusion column, was the likely cause of the proteins unfolding. A different approach that could have been attempted is to recover the unfolded protein and refold it via dialysis. This is a method often used for rescuing proteins from inclusion bodies (Singh & Panda, 2005).

5.4.2 Glycosyltransferase function from bioinformatics

Without pure enzyme to assay, elucidation of function can sometimes be achieved using bioinformatics. The nucleotide sequences of related glycosyltransferases have little to no homology and are not useful for classification or characterisation purposes. Amino acid sequences between glycosyltransferases which facilitate the same glycosidic linkage can also have low homology (Brockhausen, 2014). Despite this, they can be used to predict a

general structure and classify them into families. The amino acid sequences of WcbB, WcbE and WcbH indicate they are GT-B glycosyltransferases with the $\beta/\alpha/\beta$ Rossmann fold. The Uniprot and CAZy databases classify them as belonging to family 1 or family 4 respectively.

Analysis of the amino acid sequences for WcbB, WcbE and WcbH using ProtParam (www.Expasy.com) revealed that all of the glycosyltransferases were predicted to be unstable (where less than 40 = stable the results were 41.83, 45.24 and 43.02 respectively). They are also aliphatic and hydrophobic, confirming their association with the membrane. Analysis with TMHMM (prediction of transmembrane helices in proteins) showed that although closely associated with the membrane, there are no transmembrane regions.

Subfamily classification can sometimes be achieved by comparing similarities in the fold, active site and mechanism of action (Coutinho *et al.*, 2003). However, this is not possible for WcbB, WcbE and WcbH.

Although the broad action of glycosyltransferases can be elucidated from their structure the precise mechanism can be difficult to define, as the stereochemistry of sugars means substrates can have identical molecular weights and chemical formulae but very different structures. (Or, subtly different structures, such as α/β anomers).

As characterisation of glycosyltransferases using bioinformatics is not as forthcoming as with other enzyme classes, more work needs to be done in developing a method for purification. Chapters 3 and 4 explored the development of novel assay methods to provide an *in vitro* lipid membrane for stability and to minimise the quantity of pure glycosyltransferase required for characterisation. However, the new methods still require the production of a small amount of purified glycosyltransferase.

5.4.3 Assembly of CPS on truncated *E. coli* polysaccharides

Despite transforming *E. coli* with three plasmids for the biosynthesis of GDP-6dHep, glycosyltransferases for polymer assembly and for expression on truncated polysaccharides, no CPS antigen was detected by antibodies. There are several reasons why the CPS may not have been produced. Firstly, only low amounts of GDP-6dHep was being synthesised (Chapter 1) and may have been insufficient for the glycosyltransferases to utilise.

Secondly, as the glycosyltransferases were separated from the biosynthetic genes and expressed on a different plasmid, there may have been temporal or spatial issues in expression, meaning the glycosyltransferases weren't close enough to the substrate to react.

Thirdly, there may be a requirement for the native lipid anchor to assemble the CPS (though this was supposed to be circumvented by exploitation of the native *E. coli* polysaccharide mechanisms). Potential lipid anchors include undecaprenyl phosphate which is successfully utilised by the PglB assembly method (section 5.4.4), dolichol-phosphate which has been observed with a mannose priming sugar attached (as is suspected to be the case with the *B. pseudomallei* CPS lipid) (Maeda & Kinoshita, 2008) or a unique lipid, which is why identification is troublesome (Cocoran *et al.*, 2006; Liston *et al.*, 2015).

Finally, there may have been cross-talk (or lack of it) between the native and non-native pathways. As S7P is diverted from the pentose phosphate pathway into capsule production, it would be logical for mechanisms within the cell to exist that prevent cessation of the metabolic pathway. It is conceivable that this is mediated by the activities of glycosyltransferases. Perhaps the reason *E. coli* has an extended lag-time (Figure 35) and low GDP-6dHep yield is that it lacks the cross-talk and diverts too much S7P into biosynthesis of the precursor. WcbM

could be a regulatory point in the pathway. A similar mechanism has been identified in *E. coli* whereby increased levels of WaaA are prevented from increasing LPS expression despite the fact KDO sugars are over-expressed (Emiola *et al.*, 2016).

5.4.4 Alternative methods for CPS biosynthesis for vaccine development

Alternative methods are available for the creation of a glycoconjugate vaccine with *B. pseudomallei* CPS, such as the PglB system. This is a process for conjugating a polysaccharide to a carrier protein within a bacterial cell to create an entire product for contribution to vaccine development (Cuccui *et al.*, 2013; Figure 109). The carrier protein is the important co-stimulant in glycoconjugate vaccines for T-lymphocyte-mediated B-lymphocyte class-switching (Figure 15). Conjugating the polysaccharide and protein *in vivo* removes the need for a chemical conjugation step. It may also be possible to tag the carrier protein with a signal peptide for direct secretion into the extracellular matrix.

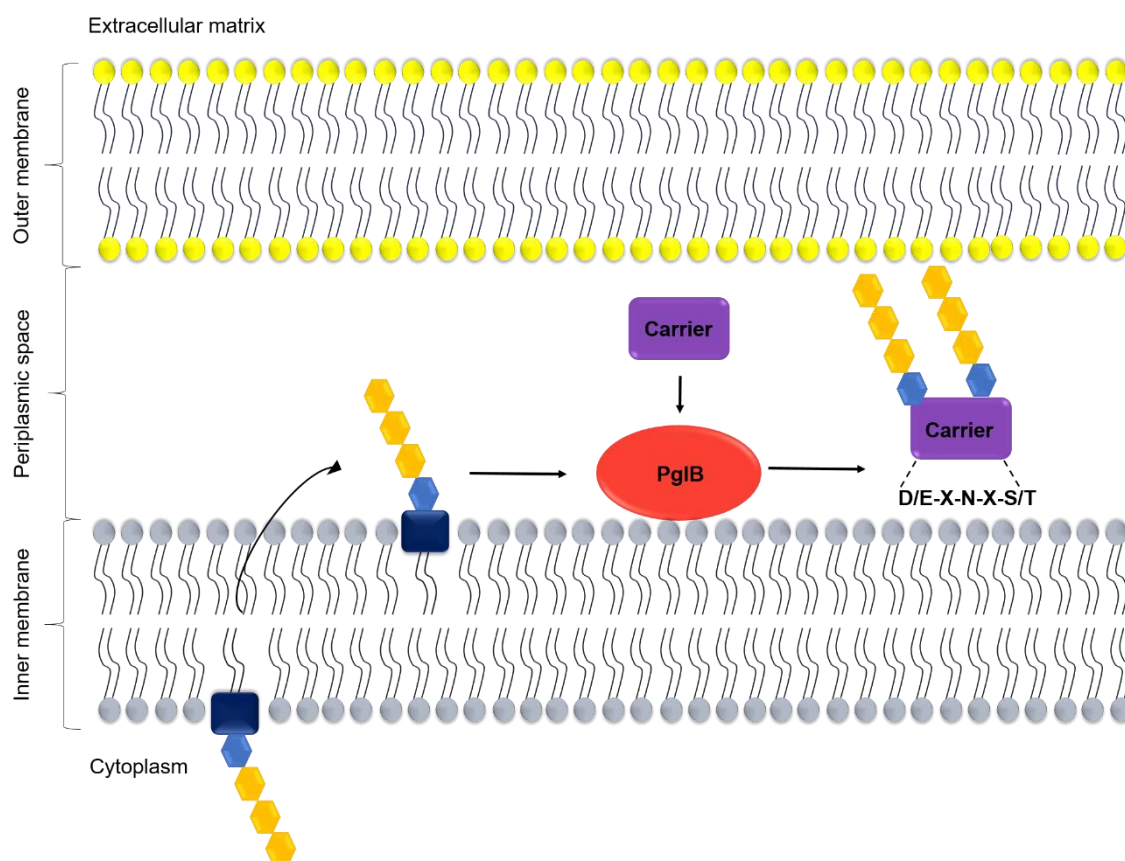


Figure 109: PglB assembly of glycoconjugates in a bacterial cell. The nascent polysaccharide (blue and orange hexagonals) is assembled onto undecaprenol pyrophosphate (dark blue rectangle) in the inner leaflet of the inner membrane. It is transported to the periplasm where PglB (red oval) recognises the undecaprenol-polysaccharide conjugate. It removes the polysaccharide and transfers it to the D/E-X-N-X-S/T recognition site on a carrier protein (purple rectangle). Adapted and reproduced from Cuccui *et al.* (2013).

The system also uses co-expression of three plasmids, one for the biosynthesis of the polysaccharide, one for the carrier protein and a third for PglB which conjugates polysaccharide to the carrier protein. In this system, the polysaccharide is synthesised directly onto undecaprenol pyrophosphate. The lack of a direct lipid anchor for the *B. pseudomallei* CPS in *E. coli* could be a reason for the lack of CPS antigen expression (Figure 108). PglB (from *C. jejuni*) is expressed in the periplasm and awaits the polysaccharide-undecaprenol pyrophosphate conjugate. Upon recognition, it transfers the entire polysaccharide from the lipid anchor to the D/E-X-N-X-S/T moiety on the carrier protein (Cuccui *et al.*, 2013; Figure 109). The system works well, providing a suitable Pgl protein

can be found, and has been used in the development of glycoconjugates for vaccines against *F. tularensis*, *S. flexneri* and *N. meningitidis* (Cuccui *et al.*, 2013; Kämpf *et al.*, 2015; Pan *et al.*, 2016).

Another potential method is to extract the CPS from *B. thailandensis* E555 and chemically conjugate to a carrier protein (Scott *et al.*, 2013). Ideal carrier proteins are LolC and Hcp from *B. pseudomallei* as these are not only immunostimulatory but provide an additional pathogen-specific antigen against which the immune system can generate an antibody response (Harland *et al.*, 2007).

This chapter lays out the foundations for a platform for bespoke polysaccharide biosynthesis in *E. coli* and identifies some of the challenges that need to be overcome.

Chapter 6

Final discussion and conclusions

6.1 Project background

B. pseudomallei is the aetiological agent of melioidosis and a potential bioterror threat (Gilligan, 2002; Butler, 2012). In addition to the severe, acute disease usually caused, *B. pseudomallei* can also establish latent, persistent and chronic infections that are difficult to treat. It is inherently resistant to a broad spectrum of antibiotics and there is no prophylactic vaccine available. The CPS, a prominent surface antigen and virulence factor, has been shown to elicit protective immunity in mammalian animal models, particularly when conjugated to a carrier protein (Burtnick *et al.*, 2012; Scott *et al.*, 2014). However, despite the capsule being strongly expressed, extraction often results in low yields (Burtnick *et al.*, 2012). Furthermore, it is highly undesirable to culture large volumes of ACDP hazard group 3 bacteria to harvest the CPS for vaccine development.

Structural analysis of the CPS revealed it is a comparatively simple homopolymer of 2-O-acetyl-deoxy-D-manno-heptopyranose (Knirel *et al.*, 1992; Perry *et al.*, 1995), linked to an undefined lipid anchor by a putative mannose priming sugar (Cuccui *et al.*, 2012). All known genes required for CPS biosynthesis are clustered on a single locus on chromosome 1 of *B. pseudomallei* and genes for the precursor sugar nucleotide and glycosyltransferases for assembly have been putatively annotated (Cuccui *et al.*, 2012).

For these reasons, the CPS of *B. pseudomallei* was chosen to contribute to the development of a synthetic biology platform for the production of bespoke polysaccharides.

The concept for this platform was to engineer an adaptable strain of *E. coli* capable of expressing any novel or bespoke polysaccharide required. Ideally, the *E. coli* would be modified to produce only the metabolites required for growth and survival and a set of truncated exopolysaccharides onto which the bespoke

polysaccharides could be assembled. (Alternatively, a system for tagging nascent polysaccharide with signal peptide for secretion into the extracellular environment could be utilised (Cuccui *et al.*, 2013)).

However, before the platform could be developed, there was a requirement to demonstrate expression of a novel precursor sugar nucleotide in *E. coli*.

6.2 Biosynthesis of GDP-6dHep

The first aim of this project was to clone six biosynthetic genes for GDP-6dHep into *E. coli*. This aim was achieved using synthetic, codon-optimised genes designed for use with the BioBricks™ 3A assembly method (Figure 18; Figure 19). Sequential assembly of the operon resulted in a construction of a high-copy number, chloramphenicol-resistant plasmid with the operon controlled by a strong constitutive promoter (Figure 24; Figure 27). GDP-6dHep was purified from *E. coli* transformed with the biosynthetic plasmid and its structure comprehensively verified by LC-MS and GC-MS (section 2.3.5). However, in the effort to simplify the process as far as possible, the six genes were all synthesised with the same c-Myc tag and RBS. This prevented translation of the proteins from being verified by SDS-PAGE and Western blotting (Figure 32). It also made the assumption that expressing all the genes with with the same strong RBS would produce the most GDP-6dHep. Furthermore, the promoter chosen was strongly and constitutively expressed. The result of this design was slowed growth with a significantly increased lag time in *E. coli* (Figure 35) and a low GDP-6dHep yield that could not be detected by NMR (section 2.3.5.4).

To improve expression of GDP-6dHep, the operon design should be optimised. The BioBricks™3A assembly method was chosen as it enables combination of components to be analysed and is easily adaptable to determine the optimal

arrangement (Shetty *et al.*, 2008). This approach could have been better exploited here, including testing different strength promoters and RBSs, different copy number plasmids and inducible promoters. Optimisation should also include analysing a more natural operon design by maintaining the gene order found on the *B. pseudomallei* CPS coding locus and without codon optimisation (as the locus is not GC-rich) (Nishizaki *et al.*, 2006; Kovács *et al.*, 2009; Cuccui *et al.*, 2012).

However, the BioBricks™ 3A assembly method was used to analyse and store each new fragment as it was assembled. For example, from a study by Butty *et al.* (2009), it was suspected that WcbM may be cytotoxic. The BioBricks™ method enabled the effect of expression to be tested before the remainder of the operon was assembled.

6.3 Characterising GDP-6d_BHep biosynthesis *in vitro*

It was accepted that an integral part of the BioBricks™ methodology is to characterise all the constituent parts to contribute to an open source database and help the synthetic biology field flourish. To characterise the GDP-6d_BHep biosynthetic pathway, *in vitro* assays were conducted. It was confirmed that although not cytotoxic, WcbM activity does have an inhibitory effect on the pathway (Table 9; Table 10; Butty *et al.*, 2009). It was further determined that WcbM inhibits WcbL activity (Table 11).

The assay was then conducted in a microfluidics system, where enzyme activity proceeds in a turbulent aqueous environment encapsulated in a lipid droplet (Leptihn *et al.*, 2013; Baxani *et al.*, 2016). Products were analysed by mass spectrometry and although most of the intermediates were absent from the sample (as expected due to being unstable or consumed by the following

reaction) there was an accumulation of the product of WcbM activity, GDP-D-glycero- α -D-manno-heptopyranose (Table 12). This could be the possible cause of the pathway inhibition.

The microfluidics system offered two potential benefits to the characterisation of glycosyltransferase activity. Firstly, the lipid layer of the droplet could offer stability to membrane-associated glycosyltransferases that quickly denature when removed from the cellular environment and are notoriously complex to purify (Breton *et al.*, 2006; Albesa-Jové *et al.*, 2014). Secondly, the droplet size enabled enzyme reactions to take place in volumes as low as 1 μ L with the GDP-6dbHep product detectable by mass spectrometry (Table 12). This could enable high-throughput kinetic analysis of glycosyltransferases, using a minimal amount of precious purified enzyme.

To improve on this area of work, adjacent droplets forming a lipid bilayer could be tested. The GmhA, WcbL and WcbN reactions could occur in one compartment and, after transfer of the products via α -haemolysin, the WcbM, WcbK and WcbJ reactions proceeding in the second compartment. The bilayer may also provide extra stability to the glycosyltransferases. Finally, issues such as glycerol from the enzyme stocks interfering with droplet formation and accurate separation of the aqueous and lipid fractions for mass spectrometry need to be resolved.

6.4 A novel assay for glycosyltransferase activity

In parallel with the microfluidics system, a novel glycosyltransferase assay was developed. The Leloir glycosyltransferase reaction releases a free nucleotide diphosphate (Leloir, 1971; Figure 60). Coupling this activity to a kinase plus ADP, produces ATP which can be measured with a luciferase reaction (BacTiter Glo™

kit) (Figure 60; Figure 61). By purifying several different kinases, which utilise different nucleotide diphosphates, a wide range of glycosyltransferase activities could be characterised by the assay, making it commercially viable. Furthermore, thermostable kinases could be used to create a one-pot assay, allowing the kinase to turnover the nucleotide diphosphate as its produced and increasing its usability (Figure 62). Development of the assay was a success and it was used to determine the kinetic parameters of commercially-available β 14GalT and a gifted sample of LgtC (section 4.3.6; section 4.3.7). However, there were no purified *B. pseudomallei* glycosyltransferases to test.

An attempt was made to purify *B. pseudomallei* glycosyltransferases WcbB, WcbE and WcbH but was not successful despite the addition of a surfactant (section 5.3.1). To improve on this, other surfactants could be tested (such as octyl β -D-glucopyranoside) or an attempt at re-folding the purified proteins could be made.

6.5 Exploiting native *E. coli* polysaccharide biosynthesis

Finally, a plan was designed to attach nascent *B. pseudomallei* CPS to truncated native polysaccharides in *E. coli* using additional glycosyltransferases from *E. coli* and *Y. pseudotuberculosis* (Figure 96; Figure 97; Table 23). Additional plasmids were constructed via the BioBricks™ method, ensuring they had different origins of replication and antibiotic resistances (Figure 98; Figure 107). Co-expression of three plasmids containing the six biosynthetic genes, three glycosyltransferases and additional glycosyltransferases produced no detectable *B. pseudomallei* CPS antigen (Figure 108). This could be due to lack of an acetylation, as this has been identified as an essential recognition site for anti-CPS antibody (Marchetti *et al.*, 2015). Despite previous in-house data and several experiments in this

study, no demonstration of acetyltransferase activity by Wcbl has been observed. This may be because Wcbl is not the correct enzyme, it may require a polymer of *deoxy-D-manno*-heptose for activity or it may require co-purification with other proteins (Cuccui *et al.*, 2012; Vivoli *et al.*, 2014). Other possibilities for lack of detectable CPS antigen are the low yield of GDP-6dDHep being insufficient to produce the CPS, cross-talk between the glycosyltransferases and native polysaccharide biosynthetic pathways or lack of the correct lipid anchor. To ensure fully-assembled CPS is achieved, future work should focus on optimising the plasmids to mimick the natural operon found on the coding locus. Not only with the aim of improving GDP-6dDHep yield, but also to increase the likelihood of harnessing the glycosyltransferase activity.

For development of a platform for bespoke polysaccharide biosynthesis, an *E. coli* strain with truncated polysaccharides and only the most essential metabolic pathways should be engineered. This would provide acceptors for nascent polysaccharides and reduce potential inhibitory cross-talk with between native and non-native glycosyltransferase pathways (Emiola *et al.*, 2016).

To manufacture a glycoconjugate vaccine for melioidosis, the *B. thailandensis* E555 strain expressing *B. pseudomallei* capsule could be used for harvesting CPS (Sim *et al.*, 2010; Scott *et al.*, 2013). Alternatively, the CPS could be engineered onto OMVs (Price *et al.*, 2016; Valentine *et al.*, 2016). However, as the genes for *B. pseudomallei* CPS expression are known and the structure is relatively simple, it would be worth pursuing the acetyltransferase to complete the antigen and continue developing the platform. The CPS could either be expressed on the cell surface and enzymatically cleaved or secreted via the PglB system to avoid LPS contamination (Cuccui *et al.*, 2013).

6.6 Final conclusion

Polysaccharides are found throughout nature and have a great number of functions in both pathogens and hosts. It is certain that manufacturing a wide range of polysaccharide structures will be an ongoing requirement in research and industry. Due to the complexity of chemical synthesis, engineering microorganisms to express bespoke polysaccharides is a viable approach. As synthetic biology tools and techniques improve, the development of a strain or collection of strains of *E. coli* that can be rapidly adapted to produce any polysaccharide desired is achievable. *E. coli* could be further adapted to secrete polysaccharide into the extracellular matrix to circumvent the issue of LPS contamination (which is a barrier to extracting the CPS from *B. thailiandensis* E555). Purified CPS from *B. pseudomallei* would be a great contribution to development of a glycoconjugate vaccine for melioidosis and, if successful, could pave the way for other biosynthetically-produced polysaccharide vaccines.

Future work should focus on establishing the correct acetyltransferase enzyme (or enzyme complex), optimising the streamlined operon and creating an *E. coli* with truncated polysaccharides and pared-down metabolism.

References

- Adler, N. R. L., Stevens, M. P., Dean, R. E., Saint, R. J., Pankhania, D., Prior, J. L., Atkins, T. P., Kessler, B., Nithichanon, A. *et al.* (2015). Systematic Mutagenesis of Genes Encoding Predicted Autotransported Proteins of *Burkholderia pseudomallei* Identifies Factors Mediating Virulence in Mice, Net Intracellular Replication and a Novel Protein Conferring Serum Resistance. *PLoS One* **10** (4), e0121271
- Ahimou, F., Semmens, M. J., Haugstad, G. & Novak, P. J. (2007). Effect of Protein, Polysaccharide, and Oxygen Concentration Profiles on Biofilm Cohesiveness. *Appl Environ Microbiol* **73** (9), 2905–2910.
- Ahmed, K., Enciso, H. D., Masaki, H., Tao, M., Omori, A., Tharavichikul, P. & Nagatake, T. (1999). Attachment of *Burkholderia pseudomallei* to pharyngeal epithelial cells: a highly pathogenic bacteria with low attachment ability. *Am J Trop Med Hyg* **60** (1), 90-93.
- Akira, S., Uematsu, S., & Takeuchi, O. (2006). Pathogen Recognition and Innate Immunity. *Cell* **124** (4), 783-801.
- Albesa-Jové, D., Giganti, D., Jackson, M., Alzari, P. M. & Guerin, M. E. (2014). Structure-function relationships of membrane-associated GT-B glycosyltransferases. *Glycobiol* **24** (2), 108-124.
- Amadasi, S., Zoppo, S. D, Bonomini, A., Bussi, A., Pedroni, P., Balestrieri, G., Signorini, L. & Castelli, F. (2015). A Case of Melioidosis Probably Acquired by Inhalation of Dusts During a Helicopter Flight in a Healthy Traveler Returning From Singapore *J Travel Med* **22** (1), 57–60.
- Amado, M., Almeida, R., Schwientek, T. & Clausen, H. (1999). Identification and characterization of large galactosyltransferase gene families: galactosyltransferases for all functions. *Biochim Biophys Acta* **1473**, 35-53.
- Anderson, J. C., Dueber, J. E., Leguia, M., Wu, G. C., Goler, J. A., Arkin, A. P. & Keasling, J. D. (2010). BglBricks: a flexible standard for biological part assembly. *J Biol Eng* **4**, 1-12.
- Angus, A. A., Agapakis, C. M, Fong, S., Yerrapragada, S., Estrada-de los Santos, P., Yang, P., Song, N., Kano, S., Caballero-Mellado, J. *et al.* (2014). Plant-Associated Symbiotic *Burkholderia* Species Lack Hallmark Strategies Required in Mammalian Pathogenesis. *PLoS One* **9** (1), e83779
- Anuntagool, N., Wuthiekanun, V., White, N. J., Currie, B. J., Sermswan, R. W., Wongratanacheewin, S., Taweekhaisupapong, S., Chaiyaroj, S. C. & Sirisinha, S (2006). Lipopolysaccharide heterogeneity among *Burkholderia pseudomallei* from different geographic and clinical origins. *Am J Trop Med Hyg* **74** (3), 348-352.
- Ardiccioni, C., Clarke, O. B., Tomasek, D., Issa, H. A., von Alpen, D. C., Pond, H. L., Banerjee, S., Rajashankar, K. R., Liu, Q. *et al.* (2016). Structure of the polyisoprenyl-

phosphate glycosyltransferase GtrB and insights into the mechanism of catalysis. *Nat Commun* **7**, e1-9.

Artenstein, M. S., Gold, R., Zimmerly, J. G., Wyle, F.A., Schneider, H. & Harkins. C. (1970). Prevention of meningococcal disease by group C polysaccharide vaccine. *N Engl J Med* **282(8)**,417-420.

Atkins, T., Prior, R. G., Mack, K., Russell, P., Nelson, M., Oyston, P. C. F., Dougan, G. & Titball, R. W. (2002a). A Mutant of *Burkholderia pseudomallei*, Auxotrophic in the Branched Chain Amino Acid Biosynthetic Pathway, Is Attenuated and Protective in a Murine Model of Melioidosis. *Infect Immun* **70 (9)**, 5290–5294.

Atkins, T., Prior, R., Mack, K., Russell, P., Nelson, M., Prior, J., Ellis, J., Oyston, P. C. F., Dougan, G. & Titball, R. W. (2002b). Characterisation of an acapsular mutant of *Burkholderia pseudomallei* identified by signature tagged mutagenesis. *J Med Microbiol* **51**, 539-547.

Attree, O. & Attree, I. (2001). A second type III secretion system in *Burkholderia pseudomallei*: who is the real culprit? *Microbiol* **147 (12)**, 3197-3199.

AuCoin, D. P., Reed, D. E., Marlenee, N. L., Bowen, R. A., Thorkildson, P., Judy, B. M., Torres, A. G. & Kozel, T. R. (2012). Polysaccharide Specific Monoclonal Antibodies Provide Passive Protection against Intranasal Challenge with *Burkholderia pseudomallei*. *PLoS ONE* **7**, e35386.

Avcı, F. Y., Li, X., Tsuji, M. & Kasper, D. L. (2011). A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. *Nat Med* **17 (12)**,1602-1609.

Baillie, L., Townend, T., Walker, N., Eriksson, U. & Williamson, D. (2004). Characterization of the human immune response to the UK anthrax vaccine. *FEMS Immunol Med Microbiol* **42**, 267–270.

Baker, A., Pearson, T., Price, E. P., Dale, J., Keim, P., Hornstra, H., Greenhill, A., Padilla, G. & Warner, J. (2011a). Molecular Phylogeny of *Burkholderia pseudomallei* from a Remote Region of Papua New Guinea. *PLoS One* **6 (3)**, e18343.

Baker, A., Tahani, D., Gardiner, C., Bristow, K. L., Greenhill, A. R. & Warner, J. (2011b). Groundwater Seeps Facilitate Exposure to *Burkholderia pseudomallei*. *Appl Environ Microbiol* **77 (20)**, 7243–7246.

Baker, M. (2016). Is there a reproducibility crisis? *Nat* **533** 452-454.

Balder, R., Lipski, S., Lazarus, J. J., Grose, W., Wooten, R. M., Hogan, R. J., Woods, D. E. & Lafontaine, E. R. (2010). Identification of *Burkholderia mallei* and *Burkholderia pseudomallei* adhesins for human respiratory epithelial cells. *BMC Microbiol* **10**, 250.

- Barnes, J. L. & Ketheesan, N. (2007).** Development of protective immunity in a murine model of melioidosis is influenced by the source of *Burkholderia pseudomallei* antigens. *Immunol Cell Biol* **85**, 551-557.
- Baxani, D., Morgan, A., Jamieson, W., Allender, C., Barrow, D. & Castell, O. (2016).** Bilayer networks within a hydrogel shell: A robust chassis for artificial cells and a platform for membrane studies. *Angew Chem Int Ed Engl* **55** (46), 14240-14245.
- Belshe, R. B., Edwards, K. M., Vesikari, T., Black, S. V., Walker, R. E., Hultquist, M., Kemble, G. & Connor, E. M. (2007).** Live Attenuated versus Inactivated Influenza Vaccine in Infants and Young Children. *N Engl J Med* **356**, 685-696.
- Berg, G., Eberl, L. & Hartmann, A. (2005).** The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. *Environ microbiol* **7** (11), 1673-1685.
- Bertrand, T., Briozzo, P., Assairi, L., Ofiteru, A., Bucurenci, N., Munier-Lehmann, H., Golinelli-Pimpaneau, B., Bârză, O. & Gilles, A. M. (2002).** Sugar specificity of bacterial CMP kinases as revealed by crystal structures and mutagenesis of *Escherichia coli* enzyme. *J Mol Biol* **315**(5), 1099-110.
- Bigger, J. (1944)** Treatment of staphylococcal infections with penicillin. *Lancet* **244**, 497-500
- Bjune, G., Høiby, E. A., Grønnesby, J. K., Arnesen, Ø, Fredriksen, J H., Halstensen, A., Holten, E., Lindbak, A. K., Nøkleby, H. et al. (1991).** Effect of outer membrane vesicle vaccine against group B meningococcal disease in Norway. *Lancet* **338**, 1093-1096.
- Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K. et al. (1997).** The complete genome sequence of *Escherichia coli* K-12. *Science* **277**(5331), 1453-1462.
- Boël, G., Letso, R., Neely, H., Price, W. N., Wong, K. H., Su, M. Luff, J. D., Valecha, M., Everett, J. K. et al. (2016).** Codon influence on protein expression in *E. coli* correlates with mRNA levels. *Nature* **529**, 358-363.
- Bondi, S. K. & Goldberg, J. B. (2008).** Strategies toward vaccines against *Burkholderia mallei* and *Burkholderia pseudomallei*. *Expert Rev Vaccines* **7** (9), 1357-1365.
- Bottero, D., Gaillard, M. E., Zurita, E., Moreno, G., Martinez, D. S., Bartel, E., Bravo, S., Carriquiriborde, F., Errea, A. & Castuma, C. (2016).** Characterization of the immune response induced by pertussis OMVs-based vaccine. *Vaccine* **34** (28), 3303-3309.
- Breitbach, K., Klocke, S., Tschernig, T., van Rooijen, N., Baumann, U & Steinmetz, I. (2006).** Role of Inducible Nitric Oxide Synthase and NADPH Oxidase in Early Control of *Burkholderia pseudomallei* Infection in Mice. *Infect Immun* **74** (11), 6300-6309.
- Breton, C., Snajdrová, L., Jeanneau, C., Koca, J. & Imberty, A. (2006).** Structures and mechanisms of glycosyltransferases. *Glycobiol* **16** (2), 29R-37R.

Brett, P. J., Mah, D. C. & Woods, D. E. (1994). Isolation and characterization of *Pseudomonas pseudomallei* flagellin proteins. *Infect immun* **62**, 1914-1919.

Brett, P. J. & Woods, D. E. (1996). Structural and immunological characterization of *Burkholderia pseudomallei* O-polysaccharide-flagellin protein conjugates. *Infect Immun* **64 (7)**, 2824-2828.

Briozzo, P., Golinelli-Pimpaneau, B., Gilles, AM., Gaucher, JF., Burlacu-Miron, S., Sakamoto, H., Janin, J. & Bârză, O. (1998). Structures of *Escherichia coli* CMP kinase alone and in complex with CDP: a new fold of the nucleoside monophosphate binding domain and insights into cytosine nucleotide specificity. *Structure* **6 (12)**, 1517-1527.

Brockhausen, I. (2014). Crossroads between bacterial and mammalian glycosyltransferases. *Front Immunol* **5**, 492

Brook, M. D., Currie, B. & Desmarchelie, P. M. (1997). Isolation and identification of *Burkholderia pseudomallei* from soil using selective culture techniques and the polymerase chain reaction. *J Appl Microbiol* **82**, 589-596.

Brudal, E., Lampe, E. O. Reubsaet, L., Roos, N., Hegna, I. K., Thrane, I. M., Koppang, E. O. & Winther-Larsen, H. C. (2015). Vaccination with outer membrane vesicles from *Francisella noatunensis* reduces development of francisellosis in a zebrafish model. *Fish Shellfish Immunol* **42 (1)**, 50-57.

Bucurenci, N., Sakamoto, H., Briozzo, P., Palibroda, N., Serina, L., Sarfati, R. S., Labesse, G., Briand, G., Danchin, A. et al. (1996). CMP kinase from *Escherichia coli* is structurally related to other nucleoside monophosphate kinases. *J Biol Chem* **271(5)**, 2856-2862.

Burntack, M. N., Brett, P. J., Harding, S. V., Ngugi, S. A., Ribot, W. J., Chantratita, N., Scorpio, A., Milne, T. S., Dean, R. E. et al. (2011). The cluster 1 Type VI secretion system is a major virulence determinant in *Burkholderia pseudomallei*. *Infect immun* **79**, 1512-1525.

Burntack, M. N., Heiss, C., Roberts, R. A., Schweizer, H. P., Azadi, P., & Brett, P. J. (2012). Development of capsular polysaccharide-based glycoconjugates for immunization against melioidosis and glanders. *Front Cell Infect Microbiol.* **2**, 108.

Burntack, M. N. & Brett, P. J. (2013). *Burkholderia mallei* and *Burkholderia pseudomallei* Cluster 1 Type VI Secretion System Gene Expression Is Negatively Regulated by Iron and Zinc. *PLoS ONE* **8 (10)**, e76767.

Butler, J. C., Breiman, R. F., Campbell, J. F., Lipman, H. B., Broome, C. V. & Facklam, R. R. (1993). Pneumococcal Polysaccharide Vaccine Efficacy An Evaluation of Current Recommendations. *JAMA.* **270 (15)**, 1826–1831.

Butler, D. (2012). Viral research faces clampdown. US biosecurity concerns could harm collaboration. *Nature* **490**, 456.

- Butt, A., Higman, V. A., Williams, C., Crump, M. P., Hemsley, C. M., Harmer, N. & Titball, R. W. (2014).** The HicA toxin from *Burkholderia pseudomallei* has a role in persister cell formation. *Biochem J* **459** (2), 333-344.
- Campbell, J. A., Davies, G. J., Bulone, V. & Henrissat, B. (1997).** A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid similarities. *Biochem J* **326**, 929-939.
- Campos, M. A., Vargas, M. A., Regueiro, V., Llompарт, C. M., Albertí, S. & Bengoechea, J. A. (2004).** Capsule Polysaccharide Mediates Bacterial Resistance to Antimicrobial Peptides. *Infect Immun*, **72** (12), 7107–7114.
- Casey, W. T., Spink, N., Cia, F., Collins, C., Romano, M., Berisio, R., Bancroft, G. J. & McClean, S. (2016).** Identification of an OmpW homologue in *Burkholderia pseudomallei*, a protective vaccine antigen against melioidosis. *Vaccine*. **34** (23), 2616-2621.
- Cescutti, P., Impallomeni, G., Garozzo, D., Sturiale, L., Herasimenka, Y., Lagatolla, C., Suwannasaen, D., Mahawantung, J., Chaowagul, W. et al. (2011).** Human Immune Responses to *Burkholderia pseudomallei* Characterized by Protein Microarray Analysis. *J Infect Dis* **203** (7), 1002–1011.
- Chan, Y. Y., Tan, T. M. C., Ong, Y. M. & Chua K. L. (2004).** BpeAB-OprB, a Multidrug Efflux Pump in *Burkholderia pseudomallei*. *Antimicrob Agents Chemother* **48** (4), 1128–1135.
- Chan, P. H., Weissbach, S., Okon, M., Withers, S. G. & McIntosh, L. P. (2012).** Nuclear magnetic resonance spectral assignments of α -1,4-galactosyltransferase LgtC from *Neisseria meningitidis*: substrate binding and multiple conformational states. *Biochemistry* **51** (41), 8278-8292.
- Chantratita, N., Wuthiekanun, V., Boonbumrung, K., Tiyawisutsri, R., Vesaratchavest, M., Limmathurotsakul, D., Chierakul, W., Wongratanacheewin, S., Pukritiyakamee, S. et al. (2007).** Biological Relevance of Colony Morphology and Phenotypic Switching by *Burkholderia pseudomallei*. *J Bacteriol* **189** (3), 807-817.
- Charuchaimontri, C., Suputtamongkol, Y., Nilakul, C., Chaowagul, W., Chetchotisakd, P., Lertpatanasuwun, N., Intaranongpai, S., Brett, P. J. & Woods, D. E. (1999).** Antilipoplysaccharide II: an antibody protective against fatal melioidosis. *Clin Infect Dis* **29**(4), 813-818.
- Chen, Y. L., Yen, Y. C., Yang, C. Y., Lee, M. S., Ho, C. K., Mena, K. D., Wang, P. Y. & Chen, P. S. (2014).** The Concentrations of Ambient *Burkholderia Pseudomallei* during Typhoon Season in Endemic Area of Melioidosis in Taiwan. *PLoS Negl Trop Dis* **8** (5), e2877.
- Chen, P. S., Chen, Y. S., Lin, H. H, Liu, P. J., Ni, W. F., Hsueh, P. T., Liang, S. H., Chen, C. & Chen, Y. L. (2015).** Airborne Transmission of Melioidosis to Humans from

- Environmental Aerosols Contaminated with *B. pseudomallei*. *PLoS Negl Trop Dis* **9** (6), e0003834.
- Cheng, A. C. & Currie, B. J. (2005)**. Melioidosis: epidemiology, pathophysiology, and management. *Clin Microbiol Rev* **18**, 383-416.
- Cheng, A. C., Wuthiekanun, V., Limmathurotsakul, D., Chierakul, W. & Peacock, S. J. (2008)**. Intensity of exposure and incidence of melioidosis in Thai children. *Trans R Soc Trop Med Hyg.* **102**, S37-39.
- Chenthamarakshan, V., Kumutha, M. V., Vadivelu, J. & Puthuchear, S. D. (2001a)**. Distribution of immunoglobulin classes and IgG subclasses against a culture filtrate antigen of *Burkholderia pseudomallei* in melioidosis patients. *J Med Microbiol.* **50** (1), 55-61.
- Chenthamarakshan, V. Vadivelu, J. & Puthuchear, S. D. (2001b)**. Detection of immunoglobulins M and G using culture filtrate antigen of *Burkholderia pseudomallei*. *Diagn Microbiol Infect Dis* **39** (1), 1-7.
- Chieng, S., Carreto, L. & Nathan, S. (2012)**. *Burkholderia pseudomallei* transcriptional adaptation in macrophages. *BMC Genomics* **13**, 328.
- Chodimella, U., Hoppes, W. L., Whalen, S., Ognibene, A. J. & Rutecki, G. W. (1997)**. Septicemia and Suppuration in a Vietnam Veteran. *Hosp Pract* **32** (5), 219-221.
- Choh, L-C., Ong, G-H., Vellasamy, K. M., Kalaiselvam, K., Kang, W-T., Al-Maleki, A. R., Mariappan, V. & Vadivelu, J. (2013)**. *Burkholderia* vaccines: are we moving forward? *Front Cell Infect Microbiol* **3** (5), 1 – 18.
- Choy, J. L., Mayo, M., Janmaat, A. & Currie, B. J. (2000)**. Animal melioidosis in Australia. *Acta Trop* **74**, 153–158.
- Chua, K. L., Chan, Y. Y. & Gan, Y. H. (2003)**. Flagella Are Virulence Determinants of *Burkholderia pseudomallei*. *Infect Immun* **71** (4), 1622-1629.
- Clements, A., Gaboriaud, F., Duval, J. F. L., Farn, J. L., Jenney, A. W., Lithgow, T., Wijburg, O. L. C., Hartland, E. L. & Strugnell, R. A. (2008)** The Major Surface-Associated Saccharides of *Klebsiella pneumoniae* Contribute to Host Cell Association. *PLoS ONE* **3** (11), e3817.
- Cochi, S. L., Hegg, L., Kaur, A., Pandak, C. & Jafari, H. (2016)**. The Global Polio Eradication Initiative: Progress, Lessons Learned, And Polio Legacy Transition Planning **35** (2), 227-283.
- Conway, T., Creecy, J. P., Maddox, S. M., Grissom, J. E., Conkle, T. L., Shadid, T. M., Teramoto, J., San Miguel, P., Shimada, T. et al. (2014)**. Unprecedented high-resolution view of bacterial operon architecture revealed by RNA sequencing. *MBio* **5** (4), e01442..

- Corcoran, A. T., Annuk, H. & Moran, A. P. (2006).** The structure of the lipid anchor of *Campylobacter jejuni* polysaccharide. *FEMS Microbiol Lett* **257**, 228–235
- Courtney, H. S., Dale, J. B. & Hastly, D. I. (1996).** Differential effects of the streptococcal fibronectin-binding protein, FBP54, on adhesion of group A streptococci to human buccal cells and HEp-2 tissue culture cells. *Infect Immun* **64** (7), 2415-2419.
- Coutinho, P. M., Deleury, E., Davies, G. J. & Henrissat, B. (2003).** An evolving Hierarchical family classification for glycosyltransferases. *J Mol Biol* **328**, 307-317.
- Cuadros, J., Gil, H., De Miguel, J., Marabé, G., Peña Gómez-Herruz, T. A., Lobo, B., Marcos, R. & Anda, P. (2011).** Melioidosis Imported from West Africa to Europe. *Am J Trop Med Hyg* **85** (2), 282-284.
- Cuccui, J., Easton, A., Chu, K. K., Bancroft, G. J., Oyston, P. C. F., Titball, R. W. & Wren, B. W. (2007).** Development of Signature-Tagged Mutagenesis in *Burkholderia pseudomallei* To Identify Genes Important in Survival and Pathogenesis. *Infect Immun* **75** (3), 1186–1195.
- Cuccui, J., Milne, T. S., Harmer, N., George, A. J., Harding, S. V., Dean, R. E., Scott, A. E., Sarkar-Tyson, M., Wren, B. W. et al. (2012).** Characterization of the *Burkholderia pseudomallei* K96243 Capsular Polysaccharide I Coding Region. *Infect Immun* **80** (3), 1209–1221.
- Cuccui, J., Thomas, R. M., Moule, M. G., D'Elia, R. V., Laws, T. R., Mills, D. C., Williamson, D., Atkins, T. P., Prior, J. L. & Wren, B. W. (2013).** Exploitation of bacterial N-linked glycosylation to develop a novel recombinant glycoconjugate vaccine against *Francisella tularensis*. *Open Biol* **3** (5), 130002.
- Cupps, T. R., Goldsmith, P. K., Volkman, D. J., Gerin, J.L., Purcell, R.H. & Fauci, A.S. (1984).** Activation of human peripheral blood B cells following immunization with hepatitis B surface antigen vaccine. *Cell Immunol* **86** (1), 145-154.
- Currie, B. J., Fisher, D. A., Anstey, N. M. & Jacups, S. P. (2000a).** Melioidosis: acute and chronic disease, relapse and re-activation. *Trans R Soc Trop Med Hyg* **94**, 301-304.
- Currie, B. J., Fisher, D. A., Howard, D. M., Burrow, J. N., Selvanayagam, S., Snelling, P. L., Anstey, N. M. & Mayo, M. J. (2000b).** The epidemiology of melioidosis in Australia and Papua New Guinea. *Acta Trop* **74**, 121-127.
- Currie, B. J., Mayo, M., Anstey, N. M., Donohoe, P., Haase, A. & Kemp, D. J. (2001).** A cluster of melioidosis cases from an endemic region is clonal and linked to the water supply using molecular typing of *Burkholderia pseudomallei* isolates. *Am J Trop Med Hyg* **65** (3), 177–179.
- Currie, B. J. (2003).** Melioidosis: an important cause of pneumonia in residents of and travellers returned from endemic regions. *Eur Respir J* **22**, 542-550.
- Currie, B. J. & Jacups, S. P. (2003).** Intensity of Rainfall and Severity of Melioidosis, Australia. *Emerg Infect Dis* **9** (12), 1538-1542.

- Currie, B. J., Jacups, S. P., Cheng, A. C., Fisher, D. A., Anstey, N. M., Huffam, S. E. & Krause, V. L. (2004). Melioidosis epidemiology and risk factors from a prospective whole-population study in northern Australia. *Trop Med Int Health* **9** (11), 1167-1174.
- Currie, B. J., Dance, D. A. & Cheng, A. C. (2008). The global distribution of *Burkholderia pseudomallei* and melioidosis: an update. *Trans R Soc Trop Med Hyg* **102**, S1-S4.
- Currie, B. J., Ward, L. & Cheng, A. C. (2010). The Epidemiology and Clinical Spectrum of Melioidosis: 540 Cases from the 20 Year Darwin Prospective Study. *PLoS Negl Trop Dis* **4** (11), e900.
- Dance, D. A. B. (1991). Melioidosis - the tip of the iceberg. *Clin Microbiol Rev* **1** (4), 52-60.
- Dance, D. A. B., Smith, M. D., Aucken, H. M. & Pitt, T. L. (1999). Imported melioidosis in England and Wales. *Lancet* **353**, 208.
- Danese, P. N., Pratt, L. A. & Kolter, R. (2000). Exopolysaccharide production is required for development of *Escherichia coli* K-12 biofilm architecture. *J Bacteriol* **182**, 3593-3596.
- Darazam, I. A., Kiani, A., Ghasemi, S., Sadeghi, H., Alavi, F., Moosavi, M. J., Akbari, A., Shahidi, M., Jalali, M. *et al.* (2011). Melioidosis: It is not far from here. *Tanaffos* **10** (4), 64-68.
- Defrance, T., Taillardet, M. & Genestier, L. (2011). T cell-independent B cell memory. *Curr Opin Immunol* **23** (3), 330-336.
- De Keulenaer, B. L., & Cheng, A. C. (2006). Severe Sepsis Due to Melioidosis. *Chest*. **130** (4), 1282.
- DeShazer, D., Brett, P. J., Carlyon, R. & Woods, D. E. (1997). Mutagenesis of *Burkholderia pseudomallei* with Tn5-OT182: isolation of motility mutants and molecular characterization of the flagellin structural gene. *J Bacteriol* **179** (7), 2116-2125.
- DeShazer, D., Brett, P. J. & Woods, D. E. (1998). The type II O-antigenic polysaccharide moiety of *Burkholderia pseudomallei* lipopolysaccharide is required for serum resistance and virulence. *Mol Microbiol* **30** (5), 1081-1100.
- DeShazer, D., Brett, P. J., Burtnick, M. N. & Woods, D. E. (1999). Molecular characterization of genetic loci required for secretion of exoproducts in *Burkholderia pseudomallei*. *J Bacteriol* **181**, 4661-4664.
- DeShazer, D., Waag, D. M., Fritz, D. L. & Woods, D. E. (2001). Identification of a *Burkholderia mallei* polysaccharide gene cluster by subtractive hybridization and demonstration that the encoded capsule is an essential virulence determinant. *Microb Pathog* **30** (5), 253-269.

- Ding, L., Yang, L., Weiss, T. M., Waring, A. J., Lehrer, R. I. & Huang, H. W. (2003). Interaction of Antimicrobial Peptides with Lipopolysaccharides. *Biochem* **42** (42), 12251-12259
- Doker, T. J., Quinn, C. L., Salehi, E. D., Sherwood, J. J., Benoit, T. J., Elrod, M. G., Gee, J. E., Shadomy, S. V., Bower, W. A. *et al.* (2014) Case Report: Fatal *Burkholderia pseudomallei* Infection Initially Reported as a *Bacillus* Species, Ohio, 2013. *Am J Trop Med Hyg* **91** (4), 743–746.
- Draper, A. D. K., Mayo, M., Harrington, G., Karp, D., Yinfoo, D., Ward, L., Haslem, A., Currie, B. J., & Kaestli, M. (2010). Association of the Melioidosis Agent *Burkholderia pseudomallei* with Water Parameters in Rural Water Supplies in Northern Australia. *Appl Environ Microbiol* **76** (15), 5305–5307.
- Druar, C., Yu, F., Barnes, J. L., Okinaka, R. T., Chantratita, N, Beg, S., Stratilo, C. W., Olive, A. J., Soltes, G. *et al.* (2008). Evaluating *Burkholderia pseudomallei* Bip proteins as vaccines and Bip antibodies as detection agents. **52** (1), 78-87.
- Easton, A., Haque, A., Chu, K., Lukaszewski, R. & Bancroft, G. J. (2007). A critical role for neutrophils in resistance to experimental infection with *Burkholderia pseudomallei*. *J Infect Dis* **195**, 99–107.
- Ellis, T. N. & Kuehn, M. J. (2010). Virulence and Immunomodulatory Roles of Bacterial Outer Membrane Vesicles. *Microbiol Mol Biol Rev* **74** (1), 81-94.
- Elschner, M. C., Hnizdo, J., Stamm, I., El-Adawy, H., Mertens, K. & Melzer, F. (2014). Isolation of the highly pathogenic and zoonotic agent *Burkholderia pseudomallei* from a pet green Iguana in Prague, Czech Republic. *BMC Vet Res* **10**, 283.
- Emiola, A., Andrews, S. S. Heller, C. & George, J. (2016). Crosstalk between the lipopolysaccharide and phospholipid pathways during outer membrane biogenesis in *Escherichia coli* *PNAS* **113** (11), 3108–3113.
- Engelthaler, D. M., Bowers, J., Schupp, J. A., Pearson, T., Ginther, J., Hornstra, H. M., Dale, J., Stewart, T., Sunenshine, R. *et al.* (2011). Molecular Investigations of a Locally Acquired Case of Melioidosis in Southern AZ, USA. *PLoS Negl Trop Dis* **5** (10), e1347.
- Engler, C., Kandzia, R. & Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647.
- Eberl, L. (2006). Quorum sensing in the genus *Burkholderia*. *Int J Med Microbiol* **296** (2-3), 103-110.
- Essex-Lopresti, A. E., Boddey, J. A., Thomas, R., Smith, M. P., Hartley, M. G., Atkins, T., Brown, N. F., Tsang, C. H., Peak, I. R. A. *et al.* (2005). A Type IV Pilin, PilA, Contributes to Adherence of *Burkholderia pseudomallei* and Virulence *In Vivo*. *Infect Immun* **73** (2), 1260–1264.

- Farasat, I., Kushwaha, M., Collens, J., Easterbrook, M., Guido, M. & Salis, H. M. (2014).** Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria. *Molecular Systems Biology*, **10 (6)**, 731.
- Fiala, G., Schamel, W. W. A. & Blumenthal, B. (2011).** Blue Native Polyacrylamide Gel Electrophoresis (BN-PAGE) for Analysis of Multiprotein Complexes from Cellular Lysates. *J Vis Exp* **48**, e2164.
- Finne, J., Bitter-Suermann, D., Goridis, C. & Finne, U. (1987).** IgG monoclonal antibody to group B meningococci cross-reacts with developmentally regulated polysialic acid units of glycoproteins in neural and extraneural tissues *J Immunol* **138 (12)**, 4402-4407.
- Fisher, N. A., Ribot, W. J., Applefeld, W. & DeShazer, D. (2012)** The Madagascar hissing cockroach as a novel surrogate host for *Burkholderia pseudomallei*, *B. mallei* and *B. thailandensis*. *BMC Microbiol* **12**, 117.
- Fransen, F., Hamstra, H. J., Boog, C. J., van Putten, J.P., van den Dobbelsteen, G. P. & van der Ley, P. (2010).** The structure of *Neisseria meningitidis* lipid A determines outcome in experimental meningococcal disease. *Infect Immun* **78**, 3177–3186.
- French, G. L. (2010).** The continuing crisis in antibiotic resistance. *Int J Antimicrob Agents* **36 (S3)**, S3-S7.
- Frey, P. A. (1996).** The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. *FASEB J* **10 (4)**, 461-470.
- Fricke, J., Neuhard, J., Kelln, R. A. & Pedersen, S. (1995).** The *cmk* Gene Encoding Cytidine Monophosphate Kinase Is Located in the *rpsA* Operon and Is Required for Normal Replication Rate in *Escherichia coli*. *J Bacteriol* **177 (3)**, 517–523.
- Galanos, C. & Freudenberg, M. A. (1993).** Mechanisms of endotoxin shock and endotoxin hypersensitivity. *Immunobiol* **187 (3-5)**, 346-356.
- Gamage, A. M., Shui, G., Wenk, M. R. & Chua, K. L. (2011).** N-Octanoylhomoserine lactone signalling mediated by the BpsI-BpsR quorum sensing system plays a major role in biofilm formation of *Burkholderia pseudomallei*. *Microbiol* **157**, 1176-1186.
- Gentry, D., Bengraj, C., Ikeharall, K. & Cashe, M. (1993).** Guanylate Kinase of *Escherichia coli* K-12. *J Biol Chem* **268 (19)**, 14316-14321.
- Gao, J. & Guo, Z. (2016).** Chemical Synthesis of the Repeating Unit of Type V Group B *Streptococcus* Capsular Polysaccharide. *Org Lett* **18 (21)**, 5552–5555.
- Gibney, K. B., Cheng, A. C. & Currie, B. J. (2008).** Cutaneous Melioidosis in the Tropical Top End of Australia: A Prospective Study and Review of the Literature. *Clin Infect Dis* **47 (5)**, 603-609.

- Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. 3rd. & Smith, H. O. (2009).** Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343–345.
- Gilad, J., Harary, I., Dushnitsky, T., Schwartz, D. & Amsalem, Y. (2007).** *Burkholderia mallei* and *Burkholderia pseudomallei* as Bioterrorism Agents: National Aspects of Emergency Preparedness. *Isr Med Assoc J* **9**, 499-503.
- Gilligan, P. H. (2002).** Therapeutic challenges posed by bacterial bioterrorism threats. *Curr Opin Microbiol* **5 (5)**, 489-95.
- Glass, M. B., Gee, J. E., Steigerwalt, A. G., Cavuoti, D., Barton, T., Hardy, R. D., Godoy, D., Spratt, B. G., Clark, T. A. & Wilkins, P. P. (2006).** Pneumonia and Septicemia Caused by *Burkholderia thailandensis* in the United States. *J Clin Microbiol* **44 (12)**, 4601–4604.
- Gonçalves, V. M., Takagi, M., Lima, R. B., Massaldi, H., Giordano, R. C. & Tanizaki, M. M. (2003).** Purification of capsular polysaccharide from *Streptococcus pneumoniae* serotype 23F by a procedure suitable for scale-up. *Biotechnol Appl Biochem* **37 (Pt 3)**, 283-287.
- Gong, L., Cullinane, M., Treerat, P., Ramm, G., Prescott, M., Adler, B., Boyce, J. D. & Devenish, R. J. (2011).** The *Burkholderia pseudomallei* Type III Secretion System and BopA Are Required for Evasion of LC3-Associated Phagocytosis. *PLoS ONE*, **6 (3)**, e17852.
- Gonzalez-Juarrero, M., Mima, N., Trunck, L. A., Schweizer, H. P., Bowen, R. A., Dascher, K., Mwangi, W. & Eckstein, T. M. (2013).** Polar Lipids of *Burkholderia pseudomallei* Induce Different Host Immune Responses. *PLoS ONE* **8 (11)**, e80368.
- Govan, B. & Ketheesan, N. (2004).** Exposure to *Burkholderia pseudomallei* induces cell-mediated immunity in healthy individuals. *Clin Microbiol Infect* **10 (6)**, 585-587.
- Green, E. R. & Mecsas, J. (2016).** Bacterial Secretion Systems – An overview. *Microbiol Spectr* **4 (1)**, e1-32.
- Green, M. R. & Sambrook, J. (2012).** *Molecular Cloning: A Laboratory Manual*. 4th ed. Cold Spring Harbor Laboratory Press.
- Gregory, A. E., Judy, B. M., Qazi, O., Blumentritt, C. A., Brown, K. A., Shaw, A. M., Torres, A. G. & Titball, R. W. (2015).** A gold nanoparticle-linked glycoconjugate vaccine against *Burkholderia mallei*. *Nanomedicine* **11 (2)**, 447-456.
- Ha, S., Chang, E., Lo, MC., Men, H., Park, P., Ge, M & Walker, S. (1999).** The Kinetic Characterization of *Escherichia coli* MurG Using Synthetic Substrate Analogues. *J Am Chem Soc* **121 (37)**, 8415 – 8426.
- Haas, A. (2007).** The phagosome: compartment with a license to kill. *Traffic* **8**, 311–330.

- Hampton, V., Kaestli, M., Mayo, M., Choy, J. L., Harrington, G., Richardson, L., Benedict, S., Noske, R., Garnet, S. T. *et al.* (2011). Melioidosis in Birds and *Burkholderia pseudomallei* Dispersal, Australia. *Emerg Infect Dis* **17** (7), 1310-1312.
- Haque, A., Easton, A., Smith, D., O'Garra, A., Van Rooijen, N., Lertmemongkolchai, G., Titball, R. W. & Bancroft, G. J. (2006a). Role of T Cells in Innate and Adaptive Immunity against Murine *Burkholderia pseudomallei* Infection. *J Infect Dis* **193**, 370-379.
- Haque, A., Chu, K., Easton, A., Stevens, M. P., Galyov, E. E., Atkins, T., Titball, R. & Bancroft, G. J. (2006b). A Live Experimental Vaccine against *Burkholderia pseudomallei* Elicits CD4+ T Cell-Mediated Immunity, Priming T Cells Specific for 2 Type III Secretion System Proteins. *J Infect Dis* **194**, 1241-1248.
- Hara, Y., Mohamed, R. & Nathan, S. (2009). Immunogenic *Burkholderia pseudomallei* Outer Membrane Proteins as Potential Candidate Vaccine Targets. *PLoS ONE* **4** (8), e6496.
- Haraga, A., West, T. E., Brittnacher, M. J., Skerrett, S. J. & Miller, S. I. (2008). *Burkholderia thailandensis* as a Model System for the Study of the Virulence-Associated Type III Secretion System of *Burkholderia pseudomallei*. *Infect Immun* **76** (11), 5402–5411.
- Harland, D. N., Chu, K., Haque, A., Nelson, M. Walker, N. J., Sarkar-Tyson, M., Atkins, T. P., Moore, B., Brown, K. A. *et al.* (2007). Identification of a LolC Homologue in *Burkholderia pseudomallei*, a Novel Protective Antigen for Melioidosis. *Infect Immun* **75** (8), 4173-4180.
- Harmer, N. J. (2010). The structure of sedoheptulose-7-phosphate isomerase from *Burkholderia pseudomallei* reveals a zinc binding site at the heart of the active site. *J Mol Biol* **400** (3), 379-392.
- Harris, P., Engler, C. & Norton, R. (2011). Comparative in vitro susceptibility of *Burkholderia pseudomallei* to doripenem, ertapenem, tigecycline and moxifloxacin. *Int J Antimicrob Agents* **37** (6), 547–549.
- Hartman, M. C. T., Jiang, S., Rush, J. S., Waechter, C. J. & Coward, J. K. (2007). Glycosyltransferase Mechanisms: Impact of a 5-Fluoro Substituent in Acceptor and Donor Substrates on Catalysis. *Biochem* **46** (41), 11630–11638.
- Hassan, M. R., Pani, S. P., Peng, N. P., Voralu, K., Vijayalakshmi, N., Mehanderkar, R., Aziz, N. A. & Michael, E. (2010). Incidence, risk factors and clinical epidemiology of melioidosis: a complex socio-ecological emerging infectious disease in the Alor Setar region of Kedah, Malaysia. *BMC Infect Dis* **10**, 302.
- Hasselbring, B. M., Patel, M. K. & Schell, M. A. (2011). *Dictyostelium discoideum* as a Model System for Identification of *Burkholderia pseudomallei* Virulence Factors. *Infect Immun* **79** (5), 2079-2088.

- Hayden, H. S., Lim, R., Brittnacher, M. J., Sims, E. H., Ramage, E. R., Fong, C., Wu, Z., Crist, E., Chang, J. *et al.* (2012). Evolution of *Burkholderia pseudomallei* in Recurrent Melioidosis. *PLoS ONE*, **7** (5), e36507.
- Heinrichs, D.E., Yethon, J. A. & Whitfield, C. (1998). Molecular basis for structural diversity in the core regions of the lipopolysaccharides of *Escherichia coli* and *Salmonella enterica*. *Mol Microbiol* **30**(2), 221-232.
- Heiss, C., Burtnick, M. N, Roberts, R. A., Black, I., Azadia, P. & Brett, P. J. (2013). Revised structures for the predominant O-polysaccharides expressed by *Burkholderia pseudomallei* and *Burkholderia mallei*. *Carbohydr Res.* **15** (381), 6–11.
- Henderson, D. A. (2011). The eradication of smallpox—an overview of the past, present, and future. *Vaccine*. **29** (Suppl 4), D7–9.
- Heng, B. H., Goh, K. T., Yap, E. H., Loh, H. & Yeo, M. (1998). Epidemiological Surveillance of Melioidosis in Singapore. *Ann Acad Med Singapore* **27**, 478-484.
- Hill, A. A., Mayo, M., Kaestli, M., Price, E. P., Richardson, L. J., Godoy, D., Spratt, B. G. & Currie, B. J. (2013). Short Report: Melioidosis as a Consequence of Sporting Activity. *Am J Trop Med Hyg* **89** (2), 365-366.
- Holden, M. T. G., Titball, R. W., Peacock, S. J., Cerdeño-Tárraga, A. M., Atkins, T., Crossman, L. C., Pitt, T., Churcher, C., Mungall, K. *et al.* (2004). Genomic plasticity of the causative agent of melioidosis, *Burkholderia pseudomallei*. *Proc Natl Acad Sci* **101**(39), 14240–14245.
- Holden, K. M. Gilbert, M., Coloe, P. J., Li, J. & Fry, B. N. (2012). The role of WlaRG, WlaTB and WlaTC in lipooligosaccharide synthesis by *Campylobacter jejuni* strain 811116. *Microb Pathog* **52** (6), 344-352.
- Hong, Y. & Reeves, P. R. (2014). Diversity of O-Antigen Repeat Unit Structures Can Account for the Substantial Sequence Variation of Wzx Translocases *J Bacteriol* **196** (9), 1713-1722.
- Houghton, R. L., Reed, D. E., Hubbard, M. E., Dillon, M. J., Chen, H., Currie, B. J., Mayo, M., Sarovich, D. S., Theobald, V. *et al.* (2014). Development of a prototype lateral flow immunoassay (LFI) for the rapid diagnosis of melioidosis. *PLoS Negl Trop Dis* **8** (3), e2727
- How, S. H., Ng, T. H., Jamalludin, A. R., Tee, H. P., Kuan, Y. C., Alex, F., Aminudin, C. A., Ortho, M. S. Sapari, S. & Quazi, M. H. (2009). Pahang Melioidosis Registry. *Med J Malaysia* **64** (1) 27-30.
- Howard, K & Inglis, T. J. J. (2005). Disinfection of *Burkholderia pseudomallei* in potable water. *Water Res* **39** (6), 1085–1092.

- Hunter, D. M. & Lim, D. V. (2010).** Rapid detection and identification of bacterial pathogens by using an ATP bioluminescence immunoassay. *J Food Prot* **73** (4), 739-746.
- Hurwitz, J. (1959).** The Enzymatic Incorporation of Ribonucleotides into Polydeoxynucleotide Material. *J Biol Chem* **234** (9), 2351-2358.
- Hyams, C., Camberlein, E., Cohen, J. M., Bax, K., & Brown, J. S. (2010).** The *Streptococcus pneumoniae* Capsule Inhibits Complement Activity and Neutrophil Phagocytosis by Multiple Mechanisms. *Infect Immun*, **78** (2), 704–715.
- Inglis, T. J. J., Garrow, S. C., Henderson, M., Clair, A., Sampson, J., O'Reilly, L. & Cameron, B. (2000a).** *Burkholderia pseudomallei* traced to water treatment plant in Australia. *Emerg Infect Dis* **6** (1), 56–59.
- Inglis, T. J. J., Rigby, P., Robertson, T. A., Dutton, N. S., Henderson, M. & Chang, B. J. (2000b).** Interaction between *Burkholderia pseudomallei* and *Acanthamoeba* Species Results in Coiling Phagocytosis, Endamebic Bacterial Survival, and Escape. *Infect Immun* **68** (3), 1681-1686.
- Inglis, T. J., Robertson, T., Woods, D. E., Dutton, N. & Chang, B. J. (2003).** Flagellummediated adhesion by *Burkholderia pseudomallei* precedes invasion of *Acanthamoeba astronyxis*. *Infect Immun* **71**, 2280-2282.
- Inglis, T.J., Rolim, D.B. & Sousa Ade, Q. (2006).** Melioidosis in the Americas. *Am J Trop Med Hyg* **75** (5), 947-54.
- Isshiki, Y., Matsuura, M., Dejsirilert, S., Ezaki, T. & Kawahara, K. (2001).** Separation of 6-deoxy-heptan from a Smooth-type Lipopolysaccharide Preparation of *Burkholderia pseudomallei*. *FEMS Microbiol. Lett* **99** (1), 21-25.
- Jesudason, M. V., Anbarasu. A. & John, T. J. (2003).** Septicaemic melioidosis in a tertiary care hospital in south India. *Indian J Med Res* **117**, 119-121.
- Jones, A. L., Beveridge, T. J. & Woods, D. E. (1996).** Intracellular survival of *Burkholderia pseudomallei*. *Infect Immun* **64**, 782-790.
- Jones, S. M., Ellis, J. F., Russell, P., Griffin, K. F. & Oyston, P. C. (2002).** Passive protection against *Burkholderia pseudomallei* infection in mice by monoclonal antibodies against capsular polysaccharide, lipopolysaccharide or proteins. *J Med Microbiol* **51**, 1055-1062.
- Jones, C. (2005).** Vaccines based on cell surface carbohydrates of pathogenic bacteria. *An Acad Bras Cienc* **77**, 293-324.
- Kaestli, M., Schmid, M., Mayo, M., Rothballer, M., Harrington, G., Richardson, L., Hill, A., Hill, J., Tuanyok, A. et al. (2012).** Out of the Ground: Aerial and Exotic Habitats of the Melioidosis Bacterium *Burkholderia pseudomallei* in Grasses in Australia. *Environ Microbiol* **14** (8), 2058–2070.

- Kämpf, M. M. Braun, M., Sirena, D., Ihssen, J., Thöny-Meyer, L. & Ren, Q. (2015).** *In vivo* production of a novel glycoconjugate vaccine against *Shigella flexneri* 2a in recombinant *Escherichia coli*: identification of stimulating factors for *in vivo* glycosylation. *Microb Cell Fact* **14** (12),
- Kang, W. T., Vellasamy, K. M., Chua, E. G. & Vadivelu, J. (2015).** Functional Characterizations of Effector Protein BipC, a Type III Secretion System Protein, in *Burkholderia pseudomallei* Pathogenesis. *J Infect Dis* **211**, 827-834.
- Karig, D. K., Iyer, S., Simpson, M. L. & Doktycz, M. J. (2012).** Expression optimization and synthetic gene networks in cell-free systems. *Nucleic Acids Res* **40** (8), 3763–3774.
- Karlyshev, A. V., Champion, O. L., Churcher, C., Brisson, J. R., Jarrell, H. C., Gilbert, M., Brochu, D., St Michael, F., Li J. *et al.* (2005).** Analysis of *Campylobacter jejuni* capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses. *Mol Microbiol* **55** (1), 90-103.
- Kastowsky, M., Gutberlet, T. & Bradaczek, H. (1992).** Molecular modelling of the three-dimensional structure and conformational flexibility of bacterial lipopolysaccharide. *J Bacteriol* **174** (14), 4798–4806.
- Kawahara, K., Dejsirilert, S. & Ezaki, T. (1998).** Characterization of three capsular polysaccharides produced by *Burkholderia pseudomallei*. *FEMS Microbiol Lett* **169** 283-287.
- Kespichayawattana, W., Rattanachetkul, S., Wanun, T., Utaisincharoen, P. & Sirisinh. (2000).** *Burkholderia pseudomallei* Induces Cell Fusion and Actin-Associated Membrane Protrusion: a Possible Mechanism for Cell-to-Cell Spreading. *Infect Immun* **68** (9), 5377-5384.
- Kespichayawattana, W., Intachote, P., Utaisincharoen, P. & Sirisinh, S. (2004).** Virulent *Burkholderia pseudomallei* is more efficient than avirulent *Burkholderia thailandensis* in invasion of and adherence to cultured human epithelial cells. *Microbial Pathogen* **36** (5), 287-292.
- Ketheesan, N., Barnes, J. L., Ulett, G. C., VanGessel, H. J., Norton, R. E., Hirst, R. G. & LaBrooy, J. T. (2002).** Demonstration of a Cell-Mediated Immune Response in Melioidosis. *J Infect Dis* **186**, 286-289.
- Khan, S. A., Everest, P., Servos, S., Foxwell, N., Zähringer, U., Brade, H., Rietschel, E. T., Dougan, G., Charles, I. G. & Maskell, D. J. (1998).** A lethal role for lipid A in *Salmonella* infections. *Mol Microbiol* **29** (2), 571-579.
- Khaskhely, N., Mosakowski, J., Thompson, R. S., Khuder, S., Smithson, S. L. & Westerink, M. A. (2012).** Phenotypic analysis of pneumococcal polysaccharide-specific B cells. *J Immunol* **188** (5), 2455-2463.

- Klugman, K. P., Koornhof, H. J., Robbins, J. B. & Le Cam, N. N. (1996). Immunogenicity, efficacy and serological correlate of protection of *Salmonella typhi* Vi capsular polysaccharide vaccine three years after immunization. *Vaccine* **14**(5), 435-438.
- Knirel, Y.A., Paramonov, N.A., Shashkov, A.S., Kochetkov, N.K., Yarullin, R.G., Farber, S.M., et al. (1992) Structure of the polysaccharide chains of *Pseudomonas pseudomallei* lipopolysaccharides. *Carbohydr Res* **233**, 185–193.
- Ko, W. C., Cheung, B. M. H., Tang, H. J., Shih, H. I., Lau, Y. J., Wang, L. R. & Chuang, Y. C. (2007). Melioidosis Outbreak after Typhoon, Southern Taiwan. *Emerg Infect Dis* **13** (6), 896-898.
- Koeller, K. M. & Wong, C. H. (2000). Complex carbohydrate synthesis tools for glyco-biologists: enzyme-based approach and programmable one-pot strategies. *Glycobiol* **10** (11), 1157-1169.
- Kovács, K., Hurst, L. D. & Papp, B. (2009). Stochasticity in Protein Levels Drives Colinearity of Gene Order in Metabolic Operons of *Escherichia coli*. *PLoS Biol* **12** **7**(5), e1000115.
- Kulkarni, H. M. & Jagannadham, M. V. (2014). Biogenesis and multifaceted roles of outer membrane vesicles from Gram-negative bacteria. *Microbiol* **160**, 2109-2121.
- Kumar, A., Chua, K. L. & Schweizer, H. P. (2006). Method for Regulated Expression of Single-Copy Efflux Pump Genes in a Surrogate *Pseudomonas aeruginosa* Strain: Identification of the BpeEF-OprC Chloramphenicol and Trimethoprim Efflux Pump of *Burkholderia pseudomallei* 1026b. *Antimicrob Agents Chemother* **50** (10), 3460–3463.
- Kurosaki, T., Kometani, K & Ise, W. (2015). Memory B cells. *Nat Rev Immunol* **15** (3), 149–159.
- Lafontaine, E. R., Balder, R., Michel, F. & Hogan, R. J. (2014). Characterization of an autotransporter adhesin protein shared by *Burkholderia mallei* and *Burkholderia pseudomallei*. *BMC Microbiol* **14**, 92.
- Lairson, L. L., Chiu, C. P. C., Ly, H. D., He, S., Wakarchuk, W. W., Strynadka, N. C. J. & Withers, S. G. (2004). Intermediate Trapping on a Mutant Retaining α -Galactosyltransferase Identifies an Unexpected Aspartate Residue *J Biol Chem* **279** (27), 28339–28344.
- Laos, R., Shaw, R., Leal, N. A., Gaucher, E. & Benner, S. (2013). Directed Evolution of Polymerases To Accept Nucleotides with Nonstandard Hydrogen Bond Patterns. *Biochem* **52** (31), 5288–5294.
- Lay, G., Poquet, Y., Salek-Peyron, P., Puissegur, M. P., Botanch, C., Bon, H., Levillain, F., Duteyrat, J. L., Emile, J. F. & Altare, F. (2006). Langhans giant cells from *M. tuberculosis*-induced human granulomas cannot mediate mycobacterial uptake. *J Pathol* **211** (1), 76-85.

- Leakey, K., Ulett, G. C & Hirst, R. G. (1998).** BALB/c and C57Bl/6 mice infected with virulent *Burkholderia pseudomallei* provide contrasting animal models for the acute and chronic forms of human melioidosis. *Microb Pathogen* **24**, 269–275.
- Lee, N., Wu, J. L., Lee, C. H. & Tsai, W. C. (1985).** *Pseudomonas pseudomallei* infection from drowning: the first reported case in Taiwan. *J Clin Microbiol* **22**, 352-354.
- Lee, Y. H., Chen, Y., Ouyang, X. & Gan, Y. H. (2010).** Identification of tomato plant as a novel host model for *Burkholderia pseudomallei*. *BMC Microbiol* **10**, 28.
- Leinonen, R., Diez, F. G., Binns, D., Fleischmann, W., Lopez, R. & Apweiler R. (2004).** UniProt Archive. *Bioinformatics* **20**, 3236-3237.
- Leloir, L. F. (1971)** Two Decades of Research on Biosynthesis of Saccharides. *Science* **172 (3990)**, 1299-1303.
- Leptihn, S, Castell, O. K., Cronin, B., Lee, E. H., Gross, L. C., Marshall, D. P., Thompson, J. R., Holden, M. & Wallace, M. I.** Constructing droplet interface bilayers from the contact of aqueous droplets in oil. *Nat Protoc* **8 (6)**, 1048-1057.
- Li, M. Z. & Elledge, S. J. (2012).** SLIC: a method for sequence- and ligation-independent cloning. *Methods Mol Biol* **852**, 51–59.
- Liebau, J., Pettersson, P., Szpryngiel, S. & Mäler, L. (2015).** Membrane Interaction of the Glycosyltransferase WaaG *Biophys J* **109 (3)**, 552-563.
- Lim, H. N., Lee, Y. & Hussein, R. (2011).** Fundamental relationship between operon organization and gene expression. *PNAS* **108 (26)**, 10626–10631.
- Limmathurotsakul, D., Wongratanacheewin, S., Teerawattanasook, N., Wongsuvan, G., Chaisuksant, S., Chetchotisakd, P., Chaowagul, W., Day, N. P. J. & Peacock, S. J. (2010).** Increasing Incidence of Human Melioidosis in Northeast Thailand. *Am J Trop Med Hyg.* **82 (6)**, 1113–1117.
- Limmathurotsakul, D., Dance, D. A.B., Wuthiekanun, V., Kaestli, M., Mayo, M., Warner, J., Wagner, D. M., Tuanyok, A., Wertheim, H. et al. (2013).** Systematic Review and Consensus Guidelines for Environmental Sampling of *Burkholderia pseudomallei*. *PLoS Negl Trop Dis* **7 (3)**, e2105.
- Limmathurotsakul, D., Wongsuvan, G., Aanensen, D., Ngamwilai, S., Saiprom, N., Rongkard, P., Thaipadungpanit, J., Kanoksil, M., Chantratita, N. et al. (2014).** Melioidosis Caused by *Burkholderia pseudomallei* in Drinking Water, Thailand, 2012 *Emerg Infect Dis* **20 (2)**, 265–268.
- Limmathurotsakul, D., Golding, N., Dance, D. A. B., Messina, J. P., Pigott, D. M., Moyes, C. L., Rolim, D. B., Bertherat, E., Day, N. P. J. et al. (2016).** Predicted global distribution of *Burkholderia pseudomallei* and burden of melioidosis. *Nat microbiol* **1**, 1-5.
- Lindberg, A. A. (1999).** Glycoprotein conjugate vaccine. *Vaccine* **17**, S28-S36.

- Link, A. J. ed. (1999).** *Methods in Molecular Biology Volume 112: 2-D Proteome Analysis Protocols*. Humana Press.
- Liston, S. D., Ovchinnikova, O. G. & Whitfield, C. (2015).** Unique lipid anchor attaches Vi antigen capsule to the surface of *Salmonella enterica* serovar Typhi *PNAS* **113 (24)**, 6719–6724.
- Liu, B., Koo, G. C., Yap, E. H., Chua, K. L. & Gan, Y. H. (2002).** Model of Differential Susceptibility to Mucosal *Burkholderia pseudomallei* Infection. *Infect Immun* **70 (2)**, 504-511.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. (2014).** The Carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**, D490–D495.
- Lukáčová, M., Baumann, M., Brade, L., Mamat, U. & Brade, H. (1994).** Lipopolysaccharide smooth-rough phase variation in bacteria of the genus *Chlamydia*. *Infect Immun* **62(6)**, 2270-2276.
- Maeda, Y. & Kinoshita, T. (2008).** Dolichol-phosphate mannose synthase: structure, function and regulation. *Biochim Biophys Acta* **1780 (6)**, 861-868.
- Marchetti, R., Dillon, M. J., Burtnick, M. N., Hubbard, M. A., Kenfack, M. T., Blériot, Y., Gauthier, C., Brett, P. J., AuCoin, D. P. et al. (2015).** *Burkholderia pseudomallei* Capsular Polysaccharide Recognition by a Monoclonal Antibody Reveals Key Details toward a Biodefense Vaccine and Diagnostics against Melioidosis. *ACS Chem Biol* **10 (10)**, 2295-2302.
- Marques, M. B., Kasper, D. L., Pangburn, M. K. & Wessels, M. R. (1992).** Prevention of C3 deposition by capsular polysaccharide is a virulence mechanism of type III group B streptococci. *Infect Immun* **60 (10)**, 3986-3993.
- Masoud, H., Ho, M., Schollaardt, T. & Perry, M. (1997).** Characterization of the capsular polysaccharide of *Burkholderia (Pseudomonas) pseudomallei* 304b. *J Bacteriol* **179**, 5663–5669.
- Mahenthalingam, E., Urban, T. A. & Goldberg, J. B. (2005).** The multifarious, multireplicon *Burkholderia cepacia* complex. *Nat Rev Microbiol* **3**, 144-156.
- McRobb, E., Kaestli, M., Price, E. P., Sarovich, D. S., Mayo, M., Warner, J., Spratt, B. G. & Currie, B. J. (2014).** Distribution of *Burkholderia pseudomallei* in Northern Australia, a Land of Diversity. *Appl Environ Microbiol* **80 (11)**, 3463–3468.
- Merino, S., Camprubi, S., Alberti, S., Benedi, V. J. & Tomas, J. M. (1992)** Mechanisms of *Klebsiella pneumoniae* resistance to complement-mediated killing. *Infect Immun* **60**, 2529–2535.

- Meumann, E. M., Cheng, A. C., Ward, L. & Currie, B. J. (2012).** Clinical Features and Epidemiology of Melioidosis Pneumonia: Results From a 21-Year Study and Review of the Literature *Clin Infect Dis* **54** (3), 362–369.
- Mott, T. M., Vijayakumar, S., Sbrana, E., Endsley, J. J. & Torres, A. G. (2015).** Characterization of the *Burkholderia mallei tonB* Mutant and Its Potential as a Backbone Strain for Vaccine Development. *PLoS Negl Trop Dis* **9** (6), e0003863.
- Mima, T., Schweizer, H. P. & Xu, Z. Q. (2011).** *In vitro* activity of cethromycin against *Burkholderia pseudomallei* and investigation of mechanism of resistance. *J Antimicrob Chemother* **66**, 73–78.
- Mitchell, R., Kelly, D. F., Pollard, A. J. & Trück, J. (2014).** Polysaccharide-specific B cell responses to vaccination in humans. *Hum Vaccin Immunother* **10** (6), 1661-1668.
- Mohanram, H. & Bhattacharjya, S. (2014).** Resurrecting Inactive Antimicrobial Peptides from the Lipopolysaccharide Trap. *Antimicrob Agents Chemother* **58** (4), 1987–1996.
- Moore, R. A., DeShazer, D., Reckseidler, S, Weissman, A. & Woods, D. E. (1999).** Efflux-Mediated Aminoglycoside and Macrolide Resistance in *Burkholderia pseudomallei*. *Antimicrob Agents Chemother* **43** (3), 465–470.
- Morosini, M. I., Quereda, C., Gil, H., Anda, P., Núñez-Murga, M., Cantón, R. & López-Vélez, R. (2013).** Melioidosis in Traveler from Africa to Spain. *Emerg Infect Dis* **19** (10), 1656–1659.
- Moule, M. G., Spink, N., Willcocks, S., Lim, J., Guerra-Assunção, J. A., Cia, F., Champion, O., Senior, N., Atkins, H. S., *et al.* (2015).** Characterization of new virulence factors involved in the intracellular growth and survival of *Burkholderia pseudomallei*. *Infect Immun* **84**(3), 701–710.
- Müller, C. M., Conejero, L., Spink, N., Wand, M. E., Bancroft, G. J. & Titball, R. W. (2012).** Role of RelA and SpoT in *Burkholderia pseudomallei* virulence and immunity. *Infect Immun* **80**, 3247-3255.
- Murphy, K. and Casey, C.** *Janeway's Immunobiology*, 9th ed. London: Garland Science, 2016.
- Nakazawa, K., Furukawa, K., Narimatsu, H., & Kobata, A. (1993).** Kinetic Study of Human β -1,4-Galactosyltransferase Expressed in *E. coli*. *J Biochem* **113**, 747-753.
- Neefjes, J., Jongasma, M. L. M., Paul, P. & Bakke, O. (2011).** Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev* **11**, 823-836.
- Nelson, M., Prior, J. L., Lever, M. S., Jones, H. E., Atkins, T. P. & Titball, R. W. (2004).** Evaluation of lipopolysaccharide and capsular polysaccharide as subunit vaccines against experimental melioidosis. *J Med Microbiol* **53**, 1177-1182.

- Nelson, M., Dean, R. E., Salguero, F. J., Taylor, C., Pearce, P. C., Simpson, A. J. H. & M. S. Lever. (2011).** Development of an acute model of inhalational melioidosis in the common marmoset (*Callithrix jacchus*). *Int J Exp Path* **94** (1), 428-435.
- Nelson, M., Salguero, F. J., Dean, R. E., Ngugi, S. A., Smither, S. J., Atkins, T. P. & Lever, M. S. (2014).** Comparative experimental subcutaneous glanders and melioidosis in the common marmoset (*Callithrix jacchus*). *Int J Exp Pathol* **95** (6), 378-391.
- Ngaui, V., Lemeshev, Y., Sadkowski, L., & Crawford, G. (2005).** Cutaneous Melioidosis in a Man Who Was Taken as a Prisoner of War by the Japanese during World War II. *J Clin Microbiol* **43** (2), 970–972.
- Ngugi, S. A., Ventura, V. V., Qazi, O., Harding, S.V., Kitto, G. B., Estes, D. M., Dell, A., Titball, R. W., Atkins, T. P. et al. (2010).** Lipopolysaccharide from *Burkholderia thailandensis* E264 provides protection in a murine model of melioidosis. *Vaccine*. **28** (47), 7551-7555.
- Nierman, W. C., Yu, Y. & Losada, L. (2015).** The *In vitro* Antibiotic Tolerant Persister Population in *Burkholderia pseudomallei* is Altered by Environmental Factors. *Front Microbiol* **6**, 1338.
- Nieves, W., Asakrah, S., Qazi, O., Brown, K. A., Kurtz, J., Aucoin, D. P., McLachlan, J. B., Roy, C. J. & Morici, L. A. (2011).** A naturally derived outer-membrane vesicle vaccine protects against lethal pulmonary *Burkholderia pseudomallei* infection. *Vaccine* **29** (46), 8381-8389.
- Nieves, W., Petersen, H., Judy, B. M., Blumentritt, C. A., Russell-Lodrigue, K., Roy, C. J., Torres, A. G. Morici, L. A. (2014).** A *Burkholderia pseudomallei* outer membrane vesicle vaccine provides protection against lethal sepsis. *Clin Vaccine Immunol* **21** (5), 747-754.
- Nimtz, M., Wray, V., Domke, T., Brenneke, B., Häussler, S. & Steinmetz, I. (1997).** Structure of an acidic exopolysaccharide of *Burkholderia pseudomallei*. *Eur J Biochem* **250**, 608-616.
- Nishizaki, T., Tsuge, K., Itaya, M., Doi, N. & Yanagawa, H. (2006).** Metabolic Engineering of Carotenoid Biosynthesis in *Escherichia coli* by Ordered Gene Assembly in *Bacillus subtilis*. *Appl Environ Microbiol* **73** (4), 1355-1361.
- Norheim, G., Tunheim, G., Næss, L. M., Kristiansen, P. A., Caugant, D. A. & Rosenqvist, E. (2012).** An outer membrane vesicle vaccine for prevention of serogroup A and W-135 meningococcal disease in the African meningitis belt. *Scand J Immunol* **76** (2), 99-107.
- Norrby, S. R., Nord, C. E., & Finch, R. (2005).** Lack of development of new antimicrobial drugs: a potential serious threat to public health. *Lancet Infect Dis* **5** (2), 115–119.

- Novak, R. T., Glass, M. B., Gee, J. E., Gal, D., Mayo, M. J., Currie, B. J. & Wilkins, P. P. (2006).** Development and evaluation of a real-time PCR assay targeting the type III secretion system of *Burkholderia pseudomallei*. *J Clin Microbiol* **44** (1), 85-90.
- Noyal, M. J., Harish, B. N., Bhat, V. & Parija, S. C. (2009).** Neonatal melioidosis: a case report from India. *Indian J Med Microbiol* **27** (3), 260-263.
- Nutt, S. L., Hodgkin, P. D., Tarlinton, D. M. & Corcoran, L. M. (2015).** The generation of antibody-secreting plasma cells. *Nat Rev Immunol* **15** (3), 160-171.
- O'Quinn, A. L., Wiegand, E. M. & Jeddloh, J. A. (2001).** *Burkholderia pseudomallei* kills the nematode *Caenorhabditis elegans* using an endotoxin-mediated paralysis. *Cell Microbiol* **3** (6), 381-393.
- O'Sullivan, B. P., Torres, B., Conidi, G., Smole, S., Gauthier, C., Stauffer, K. E., Glass, M. B., Gee, J. E., Blaney, D. & Smith, T. L. (2011).** *Burkholderia pseudomallei* Infection in a Child With Cystic Fibrosis: Acquisition in the Western Hemisphere. *Chest* **140** (1), 239-242.
- Okamoto, S., Kawabata, S., Terao, Y., Fujitaka, H., Okuno, Y. & Hamada, S. (2004).** The *Streptococcus pyogenes* Capsule Is Required for Adhesion of Bacteria to Virus-Infected Alveolar Epithelial Cells and Lethal Bacterial-Viral Superinfection. *Infect Immun* **72** (10), 6068–6075.
- Olins, P. O., Devine, C. S., Rangwala, S. H. & Kavka, K. S. (1988).** The T7 phage gene 10 leader RNA, a ribosome-binding site that dramatically enhances the expression of foreign genes in *Escherichia coli*. *Gene* **73**(1), 227-235.
- Olins, P. O. & Rangwala, S. H. (1989).** A novel sequence element derived from bacteriophage T7 mRNA acts as an enhancer of translation of the lacZ gene in *Escherichia coli*. **264** (29), 16973-16976
- Ong, C., Ooi, C. H., Wang, D. L., Chong, H. L., Ng, K. C., Rodrigues, F., Lee, M. A. & Tan, P. (2004).** Patterns of large-scale genomic variation in virulent and avirulent *Burkholderia* species. *Genome Res* **14**, 2295–2307.
- Ophir, T. & Gutnick, D. L. (1994).** A Role for Exopolysaccharides in the Protection of Microorganisms from Desiccation. *Appl Environ Microbiol* **60**, 740–745.
- Ofiteru, A., Bucurenci, N., Alexov, E., Bertrand, T., Briozzo, P., Munier-Lehmann, H & Gilles, A. M. (2007).** Structural and functional consequences of single amino acid substitutions in the pyrimidine base binding pocket of *Escherichia coli* CMP kinase. *FEBS J* **274** (13), 3363–3373.
- Paciello, I., Silipo, A., Lembo-Fazio, L., Curcurù, L., Zumsteg, A., Noël, G., Ciancarella, V., Sturiale, L., Molinaro, A. & Bernardini, M. L. (2013).** Intracellular *Shigella* remodels its LPS to dampen the innate immune recognition and evade inflammasome activation. *Proc Natl Acad Sci U S A* **110**(51), E4345–E4354.

- Pacinelli, E., Wang, L. & Reeves, P. R. (2002).** Relationship of *Yersinia pseudotuberculosis* O antigens IA, IIA, and IVB: the IIA gene cluster was derived from that of IVB. *Infect Immun* **70**, 3271–3276.
- Pan, C., Sun, P., Liu, B., Liang, H., Peng, Z., Dong, Y., Wang, D., Liu, X., Wang, B. et al. (2016).** Biosynthesis of Conjugate Vaccines Using an O-Linked Glycosylation System. *mBio* **7** (2), e00443-16.
- Papo, N. & Shai, Y. (2005).** A Molecular Mechanism for Lipopolysaccharide Protection of Gram-negative Bacteria from Antimicrobial Peptides. *J Biol Chem* **280** (11), 10378 – 10387.
- Pâquet, M.R. & Moscarello, M. A. (1984).** A kinetic comparison of partially purified rat liver Golgi and rat serum galactosyltransferases. *Biochem J* **218** (3), 745-751.
- Perrett, K. P., McVernon, J., Richmond, P. C., Marshall, H., Nissen, M., August, A., Percell, S., Toneatto, D. & Nolan, T. (2015).** Immune response to a recombinant, four-component, meningococcal serogroup B vaccine (4CMenB) in adolescents: A phase III, randomized, multicentre, lot-to-lot consistency study. *Vaccine* **33** (39), 5217-5224.
- Perry, M. B., MacLean, L. L., Schollaardt, T., Bryan, L. E. & Ho, M. (1995).** Structural characterisation of the lipopolysaccharide O antigens of *Burkholderia pseudomallei*. *Infect Immun* **63**, 3348–3352.
- Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G. & Strynadka, N. C. J. (2001).** Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs. *Nat Struct Biol* **8** (2), 166-175.
- Pichichero, M. E., Rennels, M. B., Edwards, K. M., Blatter, M. M., Marshall, G. S., Bologa, M., Wang, E. & Mills, E. (2005).** Combined tetanus, diphtheria, and 5-component pertussis vaccine for use in adolescents and adults. *J Am Med Assoc* **293** (24), 3003-3011.
- Pilatz, S., Breitbach, K., Hein, N., Fehlhaber, B., Schulze, J., Brenneke, B., Eberl, L. & Steinmetz, I. (2006).** Identification of *Burkholderia pseudomallei* genes required for the intracellular life cycle and in vivo virulence. *Infect Immun* **74**, 3576–3586.
- Pollard, A. J., Perrett, K. P. & Beverley, P. C. (2009).** Maintaining protection against invasive bacteria with protein–polysaccharide conjugate vaccines. *Nat Rev Immunol* **9**, 213-220.
- Prehm, P., Schmidt, G., Jann, B. & Jann, K. (1976).** The Cell-Wall Lipopolysaccharide of *Escherichia coli* K-12 Structure and Acceptor Site for O-Antigen and Other Substituents *Eur J Biochem* **70**, 171 – 177.
- Price, N. L., Goyette-Desjardins, G., Nothaft, H., Valguarnera, E., Szymanski, C. M., Segura, M. & Feldman, M. F. (2016).** Glycoengineered Outer Membrane Vesicles: A Novel Platform for Bacterial Vaccines. *Sci Rep* **6**, 24931.

- Propst, K. L., Mima, T., Choi, K. H., Dow, S. W. & Schweizer, H. P. (2010).** A *Burkholderia pseudomallei* Δ purM Mutant Is Avirulent in Immunocompetent and Immunodeficient Animals: Candidate Strain for Exclusion from Select-Agent Lists. *Infect Immun* **78** (7), 3136-3143.
- Pruekprasert, P. & Jitsurong, S. (1991).** Case report: septicemic melioidosis following near drowning. *Southeast Asian J Trop Med Public Health* **22**, 276-278.
- Pruksachartvuthi, S., Aswapokee, N. & Thankerngpol, K. (1990).** Survival of *Pseudomonas pseudomallei* in human phagocytes. *J Med Microbiol* **31**, 109-114.
- Puangpetch, A., Anderson, R., Huang, Y. Y., Sermswan, R. W., Chaicumpa, W., Sirisinha, S., & Wongratanacheewin, S. (2012).** Cationic Liposomes Extend the Immunostimulatory Effect of CpG Oligodeoxynucleotide against *Burkholderia pseudomallei* Infection in BALB/c Mice. *Clin Vaccine Immunol* **19** (5), 675–683.
- Qasba P. K., Ramakrishnan, B. & Boeggeman, E. (2008).** Structure and function of beta -1,4-galactosyltransferase. *Curr Drug Targets* **9** (4), 292-309.
- Qian, J., Garrett, T. A. & Raetz, C. R. H. (2014).** *In Vitro* Assembly of the Outer Core of the Lipopolysaccharide from *Escherichia coli* K-12 and *Salmonella typhimurium*. *Biochemistry* **53** (8), 1250–1262.
- Quan, J. & Tian, J. (2009).** Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS One* **4**, e6441.
- Räbinä, J., Mäki, M., Savilahti, E. M., Järvinen, N., Penttilä, L. & Renkonen, R. (2001).** Analysis of nucleotide sugars from cell lysates by ion-pair solid-phase extraction and reversed-phase high- performance liquid chromatography. *Glycoconj J* **18**, 799–805.
- Rainbow, L., Hart, C. A. & Winstanley, C. (2002).** Distribution of type III secretion gene clusters in *Burkholderia pseudomallei*, *B. thailandensis* and *B. mallei*. *J Med Microbiol* **51** (5), 374-384.
- Rajadhyaksha, A., Sonawale, A., Khare, S., Kalal, C. & Jankar, R. (2012).** Disseminated melioidosis presenting as septic arthritis. *J Assoc Physicians India* **60**, 44-45.
- Ralph, A., McBride, J. & Currie, B. J. (2004).** Transmission of *Burkholderia pseudomallei* via breast milk in northern Australia. *Pediatr Infect Dis J* **23** (12), 1169-1171.
- Ramakrishnan, B. & Qasba, P. K. (2001).** Crystal structure of lactose synthase reveals a large conformational change in its catalytic component, the beta1,4-galactosyltransferase-I. *J Mol Biol* **310** (1), 205-218.
- Randall, L. B., Dobos, K., Papp-Wallace, K. M., Bonomo, R. A. & Schweizer, H. P. (2015).** Membrane-Bound PenA β -Lactamase of *Burkholderia pseudomallei*. *Antimicrob. Agents Chemother.* **60** (3), 1509-1514.

- Rao, N. (1991).** Protecting travelers from typhoid fever. *Infect Control Hosp Epidemiol* **12** (3), 168-172.
- Rappuoli, R. & De Gregorio, E. (2011).** A sweet T cell response. *Nat Med* **17**, 1551-1552.
- Rattanavong, S., Wuthiekanun, V., Langla, S., Amornchai, P., Sirisouk, J., Phetsouvanh, R., Moore, C. E., Peacock, S. J., Buisson, Y. & Newton, P. N. (2011).** Randomized Soil Survey of the Distribution of *Burkholderia pseudomallei* in Rice Fields in Laos. *Appl Environ Microbiol* **77** (2), 532–536.
- Ray, K, Marteyn, B., Sansonetti, P. J. & Tang, C. M. (2009).** Life on the inside: the intracellular lifestyle of cytosolic bacteria. *Nat Rev Microbiol* **7**, 333-340.
- Redondo, M. C., Gómez, M., Landaeta, M. E., Ríos, H., Khalil, R., Guevara, R. N., Palavecino, S., Figuera, M., Caldera, J. et al. (2011).** Melioidosis presenting as sepsis syndrome: a case report. *Int J Infect Dis* **15** (3), e217-e218.
- Reckseidler, S. L., DeShazer, D., Sokol P, A. & Woods, D. E. (2001).** Detection of bacterial virulence genes by subtractive hybridization: identification of capsular polysaccharide of *Burkholderia pseudomallei* as a major virulence factor. *Infect Immun* **69**, 34–44
- Reckseidler-Zenteno, S., L., DeVinney, R. & Woods, D., E. (2005).** The capsular polysaccharide of *Burkholderia pseudomallei* contributes to survival in serum by reducing complement factor C3b deposition. *Infect Immun* **73** (2), 1106-1115.
- Reckseidler-Zenteno, S. L., Viteri, D. F., Moore, R., Wong, E., Tuanyok, A. & Woods, D. E. (2010).** Characterization of the type III capsular polysaccharide produced by *Burkholderia pseudomallei* *J. Med. Microbiol* **59**, 1403-1414.
- Registry of Standard Biological Parts**
[http://parts.igem.org/Main_Page?title=Main_Page] [website]
- Reeve, B., Hargest, T., Gilbert, C., & Ellis, T. (2014).** Predicting Translation Initiation Rates for Designing Synthetic Biology. *Front Bioeng Biotechnol* **2** (1), 1-6.
- Reynes, J. P., Tiraby, M., Baron, M., Drocourt, D. & Tiraby, G. (1996).** *Escherichia coli* thymidylate kinase: molecular cloning, nucleotide sequence, and genetic organization of the corresponding tmk locus. *J Bacteriol* **178** (10), 2804-2812.
- Rholl, D. A., Papp-Wallace, K. M., Tomaras, A. P., Vasil, M. L., Bonomo, R. A. & Schweizer, H. P. (2011).** Molecular investigations of PenA-mediated beta-lactam resistance in *Burkholderia pseudomallei*. *Front Microbiol* **2**, 139.
- Rizzo, R. (2003).** Exopolysaccharides produced by a clinical strain of *Burkholderia cepacia* isolated from a cystic fibrosis patient. *Carbohydrate Res* **338** (23), 2687-2695.

- Riyapa, D., Buddhisa, S., Korbsrisate, S., Cuccui, J., Wren, B. W., Stevens, M. P., Ato, M. & Lertmemongkolchai, G. (2012). Neutrophil Extracellular Traps Exhibit Antibacterial Activity against *Burkholderia pseudomallei* and Are Influenced by Bacterial and Host Factors. *Infect Immun* **80** (11), 3921-3929.
- Roantree, R. J., Kuo, T. T. & MacPhee, D. G. (1977). The effect of defined lipopolysaccharide core defects upon antibiotic resistances of *Salmonella typhimurium*. *J Gen Microbiol* **103**, 223-234.
- Roberson, E. B. & Firestone, M. K. (1992). Relationship between Desiccation and Exopolysaccharide Production in a Soil *Pseudomonas* sp. *Appl Environ Microbiol* **58** (4), 1284-1291.
- Robertson, G., Sorenson, A., Govan, B., Ketheesan, N., Houghton, R., Chen, H., AuCoin, D., Dillon, M. & Norton, R. (2015). Rapid diagnostics for melioidosis: a comparative study of a novel lateral flow antigen detection assay. *J Med Microbiol* **64**, 845-848.
- Rolim, D. B., Vilar, D. C. F. L., de Góes Cavalcanti, L. P., Freitas, L. B. N., Inglis, T. J. J., Nobre Rodrigues, J. L. & Nagao-Dias, A. T. (2011). *Burkholderia pseudomallei* Antibodies in Individuals Living in Endemic Regions in Northeastern Brazil. *Am J Trop Med Hyg* **84** (2), 302-305.
- Rotz, L. D., Khan, A. S., Lillibridge, S. R., Ostroff, S. M. & Hughes, J. M. (2002). Public health assessment of potential biological terrorism agents. *Emerg Infect Dis* **8**, 225-230.
- Roush, S. W. & Murphy, T. V. (2007). Historical Comparisons of Morbidity and Mortality for Vaccine-Preventable Diseases in the United States. *J Am Med Assoc* **298** (18), 2155-2163.
- Sakamoto, T., Ojidaa, A. & Hamachi, I. (2009). Molecular recognition, fluorescence sensing, and biological assay of phosphate anion derivatives using artificial Zn(II)-Dpa complexes. *Chem Commun* **2**, 141-152.
- Sarkar-Tyson, M., Thwaite, J. E., Harding, S. V., Smither, S. J., Oyston, P. C. F., Atkins, T. P. & Titball, R. W. (2007). Polysaccharides and virulence of *Burkholderia pseudomallei*. *J Med Microbiol* **56**, 1005-1010.
- Sarkar-Tyson, M., Smither, S. J., Harding, S. V., Atkins, T. P. & Titball, R. W. (2009). Protective efficacy of heat-inactivated *B. thailandensis*, *B. mallei* or *B. pseudomallei* against experimental melioidosis and glanders. *Vaccine* **27**, 4447-4451.
- Savitsky, P., Bray, J., Cooper, C. D. O., Marsden, B. D., Mahajan, P., Burgess-Brown, N. A. & Gileadia, O. (2010). High-throughput production of human proteins for crystallization: The SGC experience. *J Struct Biol* **172**(1), 3-13.
- Sawana, A., Adeolu, M. & Gupta, R. S. (2014). Molecular signatures and phylogenomic analysis of the genus *Burkholderia*: proposal for division of this genus into the emended

genus *Burkholderia* containing pathogenic organisms and a new genus *Paraburkholderia* gen. nov. harboring environmental species. *Front Genet* **5**, 429.

Schwechheimer, C & Kuehn, M. J. (2015). Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. *Nat Rev Microbiol* **13**, 605-619.

Schrager, H. M., Albertí, S., Cywes, C., Dougherty, G. J. & Wessels, M. R. (1998). Hyaluronic acid capsule modulates M protein-mediated adherence and acts as a ligand for attachment of group A *Streptococcus* to CD44 on human keratinocytes. *J Clin Invest* **101** (8), 1708-1716.

Schülin, T. & Steinmetz, I. (2001). Chronic Melioidosis in a Patient with Cystic Fibrosis. *J Clin Microbiol* **39** (4), 1676-1677.

Schwarz, S., Singh, P., Robertson, J. D., LeRoux, M., Skerrett, S. J., Goodlett, D. R., West, T. E. & Mougous, J. D. (2014). VgrG-5 Is a *Burkholderia* Type VI Secretion System-Exported Protein Required for Multinucleated Giant Cell Formation and Virulence. *Infect Immun* **82** (4), 1445–1452.

Schmidt, M. A. & Jann, K. (1983). Structure of the 2-Keto-3-deoxy- D-manno-octonic-acid-Containing Capsular Polysaccharide (K12 Antigen) of the Urinary-Tract-Infective *Escherichia coli* 04: K12: H⁻. *Eur J Biochem.* **131**, 509–517

Scott, A. E., Laws, T. R., D'Elia, R. V., Stokes, M. G. M., Nandi, T., Williamson, E. D., Tan, P., Prior, J. L. & Atkins, T. P. (2013). Protection against Experimental Melioidosis following Immunization with Live *Burkholderia thailandensis* Expressing a manno-Heptose Capsule. *Clin Vaccine Immunol* **20** (7), 1041-1047.

Scott, A. E., Twine, S. M., Fulton, K. M., Titball, R. W., Essex-Lopresti, A. E., Atkins, T. P. & Prior, J. L. (2011). Flagellar Glycosylation in *Burkholderia pseudomallei* and *Burkholderia thailandensis*. *J Bacteriol* **193** (14), 3577–3587.

Scott, A. E., Burtnick, M. N., Stokes, M. G. M., Whelan, A. O., Williamson, E. D., Atkins, T. P., Prior, J. L. & Brett, P. J. (2014). *Burkholderia pseudomallei* Capsular Polysaccharide Conjugates Provide Protection against Acute Melioidosis. *Infect Immun* **82** (8), 3206-3213.

Scott, A. E., Christ, W. J., George, A. J., Stokes, M. G. M., Lohman, G. J. S., Guo, Y., Jones, M., Titball, R. W., Atkins, T. P. et al. (2016). Protection against Experimental Melioidosis with a Synthetic manno-Heptopyranose Hexasaccharide Glycoconjugate. *Bioconjug Chem* **27** (6), 1435–1446.

Seeberger, P. H., Finney, N., Rabuka, D. & Bertozzi, C. R. (2009) Chemical and Enzymatic Synthesis of Glycans and Glycoconjugates. In: Cummings, V. A., Esko, R. D. et al. eds. *Essentials of Glycobiology*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press

- Seetharaman J., Kanagalaghatta R.R., Solorzano V., Kniewel R., Lima C.D., Bonanno J.B., Burley, S. K. & Swaminathan, S. (2006).** Crystal structures of two putative phosphoheptose isomerases. *Proteins Struct Funct Bioinf* **63**, 1092-1096.
- Serina, L., Blondin, C., Krin, E., Sismeiro, O., Danchin, A., Sakamoto, H., Gilles, AM. & Barzu, O. (1995).** *Escherichia coli* UMP kinase, a Member of the Aspartokinase Family, Is a Hexamer Regulated by Guanine Nucleotides and UTP. *Biochem* **34 (15)**, 5066–5074.
- Shapiro, E. D., Berg, A. T., Austrian, R., Schroeder, D., Parcels, V., Margolis, A., Adair, R. K. & Clemens, J. D. (1991).** The protective efficacy of polyvalent pneumococcal polysaccharide vaccine. *N Engl J Med* **325 (21)**, 1453-1460.
- Shetty, R. P., Endy, D. & Knight, T. F., Jr. (2008).** Engineering BioBrick vectors from BioBrick parts. *J Biol Eng* **2 (5)**, 1-12.
- Shivbalan, S., Reddy, N., Tiru, V. & Thomas, K. (2010).** Systemic melioidosis presenting as suppurative parotitis. *Indian Pediatr* **47 (9)**, 799-801.
- Silva, E. B. & Dow, S. W. (2013).** Development of *Burkholderia mallei* and *pseudomallei* vaccines. *Front Cell Infect Microbiol* **3 (10)**, 1-13.
- Silva, E. B., Goodyear, A., Sutherland, M. D., Podnecky, N. L., Gonzalez-Juarrero, M., Schweizer, H. P. & Dow, S. W. (2013).** Correlates of Immune Protection following Cutaneous Immunization with an Attenuated *Burkholderia pseudomallei* Vaccine. *Infect Immun* **81 (12)**, 4626-4634.
- Silva-Rocha, R., Martinez-Garcia, E., Calles, B., Chavarria, M., Arce-Rodriguez, A., de Las Heras, A., Paez-Espino, A. D., Durante-Rodriguez, G., Kim, J. et al. (2013).** The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res* **41**, D666–D675.
- Sim, S. H., Yu, Y., Lin, C. H., Karuturi, R. K. M., Wuthiekanun, V., Tuanyok, A., Chua, H. H., Ong, C., Paramalingam, S. S. et al. (2008).** The Core and Accessory Genomes of *Burkholderia pseudomallei*: Implications for Human Melioidosis. *PLoS pathog* **4 (10)**, e1000178.
- Sim, B. M. Q., Chantratita, N., Ooi, W. F., Nandi, T., Tewhey, R., Wuthiekanun, V., Thaipadungpanit, J., Tumapa, S., Ariyaratne, P. et al. (2010).** Genomic acquisition of a capsular polysaccharide virulence cluster by non-pathogenic *Burkholderia* isolates. *Genome Biol* **11 (8)**, R89.
- Singh, S. M. & Panda, A. K. (2005).** Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng* **99 (4)**, 303-310.
- Sleight, S. C., Bartley, B. A., Lieviant, J. A. & Sauro, H. M. (2010).** In-fusion BioBrick assembly and re-engineering. *Nucleic Acids Res* **38**, 2624–2636.

- Smanski, M. J., Bhatia, S., Zhao, D., Park, Y. B. A., Woodruff, L., Giannoukos, G., Ciulla, D., Busby, M., Calderon, J. et al. (2014).** Functional optimization of gene clusters by combinatorial design and assembly. *Nat Biotechnol* **32**(12), 1241-1249.
- Smit, P., Oberholzer, D., Hayden-Smith, S., Koornhof, H. J. & Hilleman, M. R. (1977).** Protective efficacy of pneumococcal polysaccharide vaccines. *JAMA* **238** (24), 2613-2616.
- Snyder, D. S. & McIntosh, T. J. (2000).** The Lipopolysaccharide Barrier: Correlation of Antibiotic Susceptibility with Antibiotic Permeability and Fluorescent Probe Binding Kinetics. *Biochem* **39** (38), 11777-11787.
- Srilunchang, T., Prongvitaya, T., Wongratanacheewin, S., Strugnell, R. & Homchampa, P. (2009).** Construction and characterization of an unmarked *aroC* deletion mutant of *Burkholderia pseudomallei* strain A2. *Southeast Asian J Trop Med Public Health* **40** (1), 123-130.
- Stevens, M. P., Wood, M. W., Taylor, L. A., Monaghan, P., Hawes, P., Jones, P. W., Wallis, T. S. & Galyov, E. E. (2002).** An Inv/Mxi-Spa-like type III protein secretion system in *Burkholderia pseudomallei* modulates intracellular behaviour of the pathogen. *Mol Microbiol* **46**, 649-659.
- Stevens, M. P., Haque, A., Atkins, T., Hill, J., Wood, M. W., Easton, A., Nelson, M., Underwood-Fowler, C., Titball, R. W. et al. (2004).** Attenuated virulence and protective efficacy of a *Burkholderia pseudomallei* bsa type III secretion mutant in murine models of melioidosis. *Microbiol* **150**, 2669–2676.
- Stevens, J. M., Galyov, E. E. & Stevens, M. P. (2006).** Actin-dependent movement of bacterial pathogens. *Nat Rev Microbiol* **4**, 91–101.
- Stiefel, P., Mauerhofer, S., Schneider, J., Maniura-Weber, K., Rosenberg, U. & Ren, Q. (2016).** Enzymes Enhance Biofilm Removal Efficiency of Cleaners. *Antimicrob Agents Chemother* **60** (6), 3647–3652.
- Studier, F. W. (2005).** Protein production by auto-induction in high-density shaking cultures. *Prot Exp Pur* **41**, 207–234.
- Su, Y. C., Wan, K. L., Mohamed, R. & Nathan, S. (2010).** Immunization with the recombinant *Burkholderia pseudomallei* outer membrane protein Omp85 induces protective immunity in mice. *Vaccine* **28**, 5005-5011.
- Suparak, S., Kespichayawattana, W., Haque, A., Easton, A., Damnin, S., Lertmemongkolchai, G., Bancroft, G. J. & Korbsrisate, S. (2005).** Multinucleated Giant Cell Formation and Apoptosis in Infected Host Cells Is Mediated by *Burkholderia pseudomallei* Type III Secretion Protein BipB. *J Bacteriol* **187** (18), 6556–6560.
- Suputtamongkol Y., Chaowagul, W., Chetchotisakd, P., Lertpatanasuwun, N., Intaranongpai, S., Ruchutrakool, T., Budhsarawong, D., Mootsikapun, P., Wuthiekanun, V. et al. (1999).** Risk Factors for Melioidosis and Bacteremic Melioidosis. *Clin Infect Dis* **29**, 408-413.

- Tamaki, S., Sato, T. & Matsubishi, M. (1971).** Role of Lipopolysaccharides in Antibiotic Resistance and Bacteriophage Adsorption of *Escherichia coli* K-12. *J Bacter* **105 (3)**, 968–975.
- Tandhavanant, S., Thanwisai, A., Limmathurotsakul, D., Korbsrisate, S., Day, N. P., Peacock, S. J. & Chantratita, N. (2010).** Effect of colony morphology variation of *Burkholderia pseudomallei* on intracellular survival and resistance to antimicrobial environments in human macrophages *in vitro*. *BMC Microbiol* **10**, 303
- Taylor, P. L., Blakely, K. M., De Leon, G. P., Walker, J. R., McArthur, F., Evdokimova, E., Zhang, K., Valvano, M. A., Wright, G. D. & Junop, M. S. (2008).** Structure and function of GmhA (sedoheptulose 7-phosphatase isomerase): a critical enzyme for lipopolysaccharide biosynthesis and a target for antibiotic adjuvants. *J Biol Chem* **283**, 2835-2845
- Tedaldi, L., Evitt, A., Göös, N., Jiang, J. & Wagner, G. (2014).** A practical glycosyltransferase assay for the identification of new inhibitor chemotypes. *Med Chem Commun* **5**, 1193-1201.
- Terry, J. M., Adcock, J. L., Olson, D. C., Wolcott, D. K., Schwanger, C., Hill, L. A., Barnett, N. W. & Francis, P. S. (2008).** Chemiluminescence Detector with a Serpentine Flow Cell *Anal Chem* **80**, 9817–9821.
- Thaipadungpanit, J., Chierakul, W., Pattanaporkrattana, W., Phoodaeng, A., Wongsuvan, G., Huntrakun, V., Amornchai, P., Chatchen, S., Kitphati, R., et al. (2014).** *Burkholderia pseudomallei* in water supplies, southern Thailand. *Emerg Infect Dis* **20 (11)**, 1947-1949.
- Timmons, S. C. & Jakeman, D. L. (2008).** Stereospecific synthesis of sugar-1-phosphates and their conversion to sugar nucleotides. *Carbohydr Res* **343 (5)**, 865-674.
- Tippayawat, P., Saenwongsa, W., Mahawantung, J., Suwannasaen, D., Chetchotisakd, P., Limmathurotsakul, D., Peacock, S. J., Felgner, P. L. Atkins, H. S. et al. (2009).** Phenotypic and Functional Characterization of Human Memory T Cell Responses to *Burkholderia pseudomallei*. *PLoS Negl Trop Dis* **3 (4)**, e407.
- Titball, R. W. & Williamson, E. D. (2004).** *Yersinia pestis* (plague) vaccines. *Expert Opin Biol Ther* **4 (6)**, 965-973.
- Tomas, J. M., Benedi, V. J., Ciurana, B. & Jofre, J. (1986).** Role of capsule and O antigen in resistance of *Klebsiella pneumoniae* to serum bactericidal activity. *Infect Immun* **54**, 85–89.
- Truong, K. K., Moghaddam, S., Al Saghbini, S. & Saatian, B. (2015).** Case of a Lung Mass due to Melioidosis in Mexico. *Am J Case Rep* **16**, 272-275

- Tuanyok, A., Tom, M., Dunbar, J. & Woods, D. E. (2006).** Genome-wide expression analysis of *Burkholderia pseudomallei* infection in a hamster model of acute melioidosis. *Infect Immun* **74**, 5465-5476.
- Tuanyok, A., Stone, J. K., Mayo, M., Kaestli, M., Gruendike, J., Georgia, S., Warrington, S., Mullins, T., Allender, C. J. et al. (2012).** The Genetic and Molecular Basis of O-Antigenic Diversity in *Burkholderia pseudomallei* Lipopolysaccharide. *PLoS Negl Trop Dis* **6** (1), e1453.
- Valentine, J. L., Chen, L., Perregaux, E. C., Weyant, K. B., Rosenthal, J. A., Heiss, C., Azadi, P., Fisher, A. C., Putnam, D. et al. (2016).** Immunization with outer membrane vesicles displaying designer glycotopes yields class-switched, glycan-specific antibodies. *Cell Chem Biol* **23** (6), 655–665.
- Valenzuela, J. G., Chuffe, O. M. & Ribeiro, J. M. C. (1996).** Apyrase Salivary and Anti-Platelet Activities from the Glands of the Bed Bug *Cimex lectularius*. *Insect Biochem Molec Biol* **21** (6), 557-562.
- Valenzuela, J. G., Charla, R., Galperin, M. Y. & Ribeiro, J. M. C. (1998).** Purification, Cloning, and Expression of an Apyrase from the Bed Bug *Cimex lectularius*. *J Biol Chem* **273** (46), 30583–30590.
- Van Rompay, A. R., Johansson, M. & Karlsson, A. (1999).** Phosphorylation of deoxycytidine analog monophosphates by UMP-CMP kinase: molecular characterization of the human enzyme. *Mol Pharmacol* **56** (3), 562-569.
- Van Schaik, E. J., Tom, M. & Woods, D. E. (2009).** *Burkholderia pseudomallei* isocitrate lyase is a persistence factor in pulmonary melioidosis: implications for the development of isocitrate lyase inhibitors as novel antimicrobials. *Infect Immun* **77**, 4275-4283.
- Van Summeren-Wesenhagen, P. V., Voges, R., Dennig, A., Sokolowsky, S., Noack, S., Schwaneberg, U., & Marienhagen, J. (2015).** Combinatorial optimization of synthetic operons for the microbial production of p-coumaryl alcohol with *Escherichia coli*. *Microb Cell Fact* **14** (79) 1-10.
- Vandamme, P., Holmes, B., Van Canneyt, M., Coenye, T., Hoste, B., Coopman, R., Revets, H., Lauwers, S., Gillis, M. et al. (1997).** Occurrence of Multiple Genomovars of *Burkholderia cepacia* in Cystic Fibrosis Patients and Proposal of *Burkholderia multivorans* sp. nov. *Int J Syst Bacter*, **47** (4), 1188-1200.
- Vasu, C., Vadivelu, J. & Puthuchery, S. D. (2003).** The Humoral Immune Response in Melioidosis Patients during Therapy. *Infect* **31** (1), 24-30.
- Vercauteren, R., Dom, P. & Haesebrouck, F. (1993).** Virulence of bacteria in relation to their hydrophobicity, adhesiveness and phagocytosis. *Vlaams Diergeneeskd Tijdschr* **62**, 29–34.
- Vidyalakshmia, K., Lipikab, S., Vishalc, S., Damodard, S. & Chakrapanid, M. (2012).** Emerging clinico-epidemiological trends in melioidosis: analysis of 95 cases from Western coastal India. *Int J Infect Dis* **16** (7), e491-e497.

- Vidyalakshmi, K., Chakrapani, M., Shrikala, B., Damodar, S., Lipika, S. & Vishal, S. (2008).** Tuberculosis mimicked by melioidosis. *Int J Tuberc Lung Dis* **12** (10), 1209–1215.
- Visca, P., Cazzola, G., Petrucca, A. & Braggion, C. (2001).** Travel-Associated *Burkholderia pseudomallei* Infection (Melioidosis) in a Patient with Cystic Fibrosis: A Case Report. *Clin Infect Dis* **32** (1), e15-e16.
- Vivoli, M., Isupov, M. N., Nicholas, R., Hill, A., Scott, A. E., Kosma, P., Prior, J. L. & Harmer, N. J. (2015).** Unraveling the *B. pseudomallei* Heptokinase WcbL: From Structure to Drug Discovery. *Chem Biol* **22** (12), 1622–1632.
- Vollmar, P., Zange, S., Fieser, N., Riehm, J., Zöller, L. & Thoma B., (2014).** Osteomyelitis caused by *Burkholderia pseudomallei*: relapse six years after pulmonary infection. *Clin Lab* **60** (9), 1565-1567.
- Vongphayloth, K., Rattanavong, S., Moore C. E., Phetsouvanh, R., Wuthiekanun, V., Sengdouangphachanh, A., Phouminh, P., Newton, P. N. & Buisson, Y. (2012).** *Burkholderia pseudomallei* detection in surface water in southern Laos using Moore's swabs. *Am J Trop Med Hyg* **86** (5), 872-877.
- Wakarchuk, W. W., Cunningham, A., Watson, D. C. & Young, N. M. (1998).** Role of paired basic residues in the expression of active recombinant galactosyltransferases from the bacterial pathogen *Neisseria meningitidis*. *Protein Eng* **11** (4), 295–302.
- Wand, M. E., Müller, C. M., Titball, R. W. & Michell, S. L. (2011).** Macrophage and *Galleria mellonella* infection models reflect the virulence of naturally occurring isolates of *B. pseudomallei*, *B. thailandensis* and *B. oklahomensis*. *BMC Microbiol* **11** (1), e1.
- Warawa, J. & Woods, D. E. (2005).** Type III secretion system cluster 3 is required for maximal virulence of *Burkholderia pseudomallei* in a hamster infection model. *FEMS Microbiol Lett* **242**, 101–108.
- Warawa, J. M., Long, D., Rosenke, R., Gardner, D. & Gherardini, F. C. (2009).** Role for the *Burkholderia pseudomallei* capsular polysaccharide encoded by the wcb operon in acute disseminated melioidosis. *Infect Immun* **77**, 5252-5261.
- Wayllace, N. Z., Valdez, H. A., Merás, A., Ugalde, R. A., Busi, M. V. & Gomez-Casati, D. F. (2012).** An enzyme-coupled continuous spectrophotometric assay for glycogen synthases. *Mol Biol Rep* **39**, 585–591.
- West, T. E., Ernst, R. K., JanssonHutson, M. J. & Skerrett, S. J. (2008).** Activation of Toll-like receptors by *Burkholderia pseudomallei*. *BMC Immunol* **9**, 46.
- White, N. J., Dance, D. A., Chaowagul, W., Wattanagoon, Y., Wuthiekanun, V. & Pitakwatchara, N. (1989).** Halving of mortality of severe melioidosis by ceftazidime. *Lancet* **23** (2), 697-701.

- Whitfield, C. & Roberts, I. S. (1999).** Structure, assembly and regulation of expression of capsules in *Escherichia coli*. *Mol Microbiol* **31** (5), 1307-1319.
- Whitfield, C. (2006).** Biosynthesis and Assembly of Capsular Polysaccharides in *Escherichia coli*. *Annu Rev Biochem* **75**, 39-68.
- Whitlock, G. C., Estes, M. & Torres, A. G. (2007).** Glanders: off to the races with *Burkholderia mallei* *FEMS Microbiol Lett* **277**, 115-122.
- Whitmore, A. (1913).** An Account of a Glanders-like Disease occurring in Rangoon. *J Hyg (Lond)* **13** (1), 1–34.
- Wiersinga, W. J., van der Poll, T., White, N. J., Day, N. P. & Peacock, S. J. (2006).** Melioidosis: insights into the pathogenicity of *Burkholderia pseudomallei*. *Nature Rev* **4**, 272-282.
- Wiersinga, W. J., Wieland, C. W., Dessing, M. C., Chantratita, N., Cheng, A. C., Limmathurotsakul, D., Chierakul, W., Leendertse, M., Florquin, S. et al. (2007).** Toll-Like Receptor 2 Impairs Host Defense in Gram-Negative Sepsis Caused by *Burkholderia pseudomallei* (Melioidosis). *PLoS Med* **4** (7), e248.
- Wiersinga, W. J., Currie, B. J. & Peacock, S. J. (2012).** Melioidosis. *N Engl J Med* **367**, 1035-1044.
- Wiersinga, W. J., Birnie, E., Weehuizen, T. A. F., Alabi, A. S., Huson, M. A. M., Huis in 't Veld, R. A. G., Mabala, H. K., Adzoda, G. K., Raczynski-Henk, Y. et al. (2015).** Clinical, Environmental, and Serologic Surveillance Studies of Melioidosis in Gabon, 2012–2013. *Emerg Infect Dis* **21** (1), 40–47.
- Wilde, H & Suankratay, C. (2007).** There is Need for Antigen-Based Rapid Diagnostic Tests to Identify Common Acute Tropical Illnesses. *J Travel Med* **14** (4), 254–258.
- Willcocks, S. J., Denman, C. C., Atkins, H. S. & Wren, B. W. (2016).** Intracellular replication of the well-armed pathogen *Burkholderia pseudomallei*. *Curr Opin Microbiol* **29**, 94-103.
- Wikraiphat, C., Charoensap, J., Utaisinchaoen, P., Wongratanacheewin, S., Taweechaisupapong, S., Woods, D. E., Bolscher, J. G. & Sirisinha, S. (2009).** Comparative in vivo and in vitro analyses of putative virulence factors of *Burkholderia pseudomallei* using lipopolysaccharide, capsule and flagellin mutants. *FEMS Immunol Med Microbiol* **56**, 253-259.
- Wikraiphat, C., Saiprom, N., Tandhavanant, S., Heiss, C., Azadi, P., Wongsuvan, G., Tuanyok, A., Holden, M. T., Burtnick M. N. et al. (2015).** Colony morphology variation of *Burkholderia pseudomallei* is associated with antigenic variation and O-polysaccharide modification. *Infect Immun* **83** (5), 2127-2138.

- Winstanley, C., Hales, B. A. & Hart, C. A. (1999).** Evidence for the presence in *Burkholderia pseudomallei* of a type III secretion system-associated gene cluster. *J Med Microbiol* **48** (7), 649-656.
- Wongkongkatep, J., Miyahara, Y., Ojida, A. & Hamachi, I. (2006).** Label-Free, Real-Time Glycosyltransferase Assay Based on a Fluorescent Artificial Chemosensor. *Angew Chem* **118** (4), 681-684.
- Wu, Z. L., Ethen, C. M., Prather, B., Machacek, M. & Jiang, W. (2011).** Universal phosphatase-coupled glycosyltransferase assay. *Glycobiol* **21** (6) 727–733.
- Wuthiekanun, V., Dance, D. A., Wattanagoon, Y., Supputtamongkol, Y., Chaowagul, W. & White, N. J. (1990).** The use of selective media for the isolation of *Pseudomonas pseudomallei* in clinical practice. *J Med Microbiol* **33**, 121-126.
- Wuthiekanun, V., Limmathurotsakul, D., Chantratita, N., Feil, E. J., Day, N. P. J & Peacock, S. J. (2009).** *Burkholderia pseudomallei* Is Genetically Diverse in Agricultural Land in Northeast Thailand. *PLoS Negl Trop Dis* **3** (8), e496.
- Wuthiekanun, V., Amornchai, P., Saiprom, N., Chantratita, N., Chierakul, W., Koh, G. C. K. W., Chaowagul, W., Day, N. P. J., Limmathurotsakul, D. & Peacock, S. J. (2011).** Survey of Antimicrobial Resistance in Clinical *Burkholderia pseudomallei* Isolates over Two Decades in Northeast Thailand. *Antimicrob Agents Chemother* **55** (11), 5388-5391.
- Xu, Y., Smith, R., Vivoli, M., Ema, M., Goos, N., Gehrke, S., Harmer, N. J., Wagner, G. K.** Covalent inhibitors of LgtC: A blueprint for the discovery of non-substrate-like inhibitors for bacterial glycosyltransferases. *Bioorg Med Chem* [Under Review.]
- Yabuuchi, E., Kosako, Y., Oyaizu, H., Yano, I., Hotta, H., Hashimoto, Y., Ezaki, T., & Arakawa, M. (1992).** Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. *Microbio Immunol* **36** (12),1251-1275.
- Yates, R. M., Hermetter, A. & Russell, D. G. (2005).** The kinetics of phagosome maturation as a function of phagosome/lysosome fusion and acquisition of hydrolytic activity. *Traffic* **6**, 413–420.
- Yi W1, Zhu L, Guo H, Li M, Li J, Wang PG. (2006).** Formation of a new O-polysaccharide in *Escherichia coli* O86 via disruption of a glycosyltransferase gene involved in O-unit assembly. *Carbohydr Res* **341** (13), 2254-2260.
- Yoona, Y. K., Alia, M. A., Weia, A. C., Choona, T. A. & Ismaild, R. (2015).** Synthesis and evaluation of antimycobacterial activity of new benzimidazole aminoesters. *Eur J Med Chem* **93**, 614–624.
- Yu, Y., Kim, H. S., Chua, H. H., Linm, C. H., Sim, S. H., Lin, D., Derr, A., Engels, R., DeShazer, D. et al. (2006).** Genomic patterns of pathogen evolution revealed by

comparison of *Burkholderia pseudomallei*, the causative agent of melioidosis, to avirulent *Burkholderia thailandensis*. *BMC Microbiol* **6**, 46.

Zanetti, F., De Luca, G. & Stampi, S. (2000). Recovery of *Burkholderia pseudomallei* and *B. cepacia* from drinking water. *Int J Food Microbiol* **59 (1-2)**, 67-72.

Zehnder, A. M., Hawkins, M. G., Koski, M. A., Lifland, B., Byrne, B. A., Swanson, A. A., Rood, M. P., Gee, J. E., Elrod, M. G. et al. (2014). *Burkholderia pseudomallei* isolates in 2 pet iguanas, California, USA. *Emerg Infect Dis* **20 (2)**, 304-306.

Zhang, N. & Bevan, M. J. (2011). CD8+ T Cells: Foot Soldiers of the Immune System. *Immun* **35 (2)**, 161-168.

Zhang, S., Feng, S. H., Li, B., Kim, H. Y., Rodriguez, J., Tsai, S. & Lo, S. C. (2011). *In Vitro* and *In Vivo* studies of monoclonal antibodies with prominent bactericidal activity against *Burkholderia pseudomallei* and *Burkholderia mallei*. *Clin Vaccine Immunol* **18 (5)**, 825-834.

Zhang, Y., Werling, U. & Edelman, W. (2012). SLiCE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res* **40 (8)**, e55.

Zhou, D., Utkina, N., Li, D., Dong, C., Druzhinina, T., Veselovsky, V., Liu, B. (2013). Biochemical characterization of a new beta-1,3-galactosyltransferase WbuP from *Escherichia coli* O114 that catalyzes the second step in O-antigen repeating-unit. *Carbohydr Res* **381**, 43–50.

Zhao, G., Perepelov, A. V., Senchenkova, S. N., Shashkov, A. S., Feng, L., Li, X., Knirel, Y. A. & Wang, L. (2007). Structural relation of the antigenic polysaccharides of *Escherichia coli* O40, *Shigella dysenteriae* type 9, and *E. coli* K47. *Carbohydr Res* **342 (9)**, 1275-1279.

Zhou, D., Chen, C., Xua, L., Utkina, N., Danilov, L., Torgov, V., Veselovsky, V., Liu, B. & Feng, L. (2016). Mass spectrometric characterization of a two-glycosyltransferase tandem reaction for assembly of tetrasaccharide repeating unit of *Escherichia coli* O77 O-antigen. *Carbohydr Res* **424 (7)**, 24–29.

Appendices

Appendix 1 – Primer sequences

Target	Direction ¹	Sequence	Notes
<i>apy</i> (Apyrase)	F	CCATGGGCCATCATCACCATCACCATAGC	
	R	AAGCTTTTAGACAAAGTCCACGCCTTC	
<i>wbaD</i>	F	<u>TACTTCCAATCCATGAGCAACATCGACATCAACAAGAAACAG</u>	2
	R	<u>TATCCACCTTTACTGTCATTACGCATAGTCGGGCACATCATAACGGG</u>	2
<i>wbnI</i>	F	<u>TACTTCCAATCCATGGTCATTAACATCTTTTATATCTGTAAGGG</u>	2
	R	<u>TATCCACCTTTACTGTCATTATTTATCGTCGTCATCCTTGTAATCTTTC</u>	2
<i>wbuP</i>	F	<u>TACTTCCAATCCATGCGCTGCAAACATGCCGTGATTATGTGC</u>	2
	R	<u>TATCCACCTTTACTGTCATTAATGGTGATGATGGTGATGCCGCAGTTTC</u>	2
<i>wbbK</i>	F	<u>TACTTCCAATCCATGGGAAAATCCATTGTCGTTGTTCCGCTG</u>	2
	R	<u>TATCCACCTTTACTGTCATTAGGCGTAATCGGGAACATCATAACGGATAG</u>	2
<i>waaB</i>	F	<u>TACTTCCAATCCATGAAAATCGCCTTCATTGGCGAAGCAGTG</u>	2
	R	<u>TATCCACCTTTACTGTCATTATTTATCATCATCGTCTTTGTAGTCTTTACG</u>	2
<i>wffQ</i>	F	<u>TACTTCCAATCCATGAAAATCATTCTGTGGCATCTGGGTCTG</u>	2
	R	<u>TATCCACCTTTACTGTCATTAATGATGGTGGTGATGATGCTGGTTG</u>	2
<i>wbyB</i>	F	<u>TACTTCCAATCCATGAATATCCTTATCGTCTATTTGGGGCG</u>	2
	R	<u>TATCCACCTTTACTGTCATTACTTGTCGTCATCGTCTTTGTAATCTGACCG</u>	2
<i>wcbB</i>	F	<u>TACTTCCAATCCATGTCAATCATCTACCTAGACGTCACGCGTCTCG</u>	2
	R	<u>TATCCACCTTTACTGTCAGCGCAGCGTGTCAAGAAAGGCATCGACG</u>	2
	F	ATGTCCATCATCTACCTGGACGTCACACG	3
	R	TTATTTGTCATCATCGTCTTTGTAATCG	3
<i>wcbE</i>	F	<u>TACTTCCAATCCATGAAGAATAAAGTGCGTGTATTGGTTGAACTTCG</u>	2
	R	<u>TATCCACCTTTACTGTCAGGGGCGCAACCAACGCTTTGTATAAACTCATCC</u>	2
	F	ATGAAGAACAAGTACGCGTCCTGGTCTG	3
	R	TTACGCATAATCCGGTACGTCGTACGGG	3

Mesophilic <i>cmk</i>	F	TACTTCCAATCCATGACGGCAATTGCCCGGTTATTACCATTGATGGCC	
	R	TATCCACCTTTACTGTCATTATGCGAGAGCCAATTTCTGGCGCGCGTATTGTAGC	
Mesophilic <i>gmk</i>	F	TACTTCCAATCCATGGCTCAAGGCACGCTTTATATTGTTTCTGC	
	R	TATCCACCTTTACTGTCATCAGTCTGCCAACAATTTGCTGATTAAAGC	
pSB1C3 F1	F	TCCGGCAAAAAAGGGCAAG	7
pSB1C3 R1	R	GAAAGGCGGACAGGTATCC	7
pSB1C3 F2	F	TTCTCACCAATAAAAAACGCCCG	7
pSB1C3 R2	R	GTGAAAACCTGGCCTATTTCCC	7
<i>wcbJ</i> R2	R	GCGCCATAAACGGTATCGG	7
<i>wcbK</i> R2	R	GGATGAAAGGTGCACTCTTCG	7
<i>wcbM</i> R2	R	GCTACAATGGTCGAAAACCG	7
<i>wcbL</i> F2	F	CATGATTCGCAACTACAATCATGG	7
<i>wcbL</i> R2	R	ATTGTCGCTCTGATCCTGG	7
<i>gmhA</i> F2	F	TTGATCGTCCTGGTCTGCC	7
<i>wcbN</i> R2	R	GGACAAAGAATCGGACATTCAGG	7

Notes

- 1: F = 5' to 3'; R = 3' to 5'
- 2: overhanging sequences (underlined) for LIC cloning
- 3: for use with synthetic genes flanked by restriction sites for BioBricks™ 3A assembly
- 4: pSB4X5 = pSB4A5/pSB4C5/pSB4K5
- 5: BioBricks™ constitutive promoter
- 6: BioBricks™ double terminator
- 7: GDP-6dBHep biosynthesis plasmid sequencing primer

Appendix 2 – Protein List

Gene name	Protein	Organism of origin	Vector	Antibiotic Resistance	Notes*
APY	Apyrase	<i>Cimex lectularius</i>	pET-Duet-1	Ampicillin	1
WbaD		<i>Escherichia coli</i>	pEX-K2	Kanamycin	2,3,4,5&7
WbnI		<i>Escherichia coli</i>	pEX-A2	Ampicillin	2,3,4,5&6
WbuP		<i>Escherichia coli</i>	pEX-A2	Ampicillin	2,3,4,5&8
WbbK		<i>Escherichia coli</i>	pEX-K2	Kanamycin	2,3,4,5&7
WaaB		<i>Escherichia coli</i>	pEX-K2	Kanamycin	2,3,4,5&6
WffQ		<i>Escherichia coli</i>	pEX-K2	Kanamycin	2,3,4,5&8
WbyB	Putative 6-deoxy-D-manno-heptose transferase	<i>Yersinia pseudotuberculosis</i>	pEX-K2	Kanamycin	2,3,4,5&6
WcbB	Putative CPS glycosyltransferase biosynthesis	<i>Burkholderia pseudomallei</i>	pNIC28-Bsa4	Kanamycin	
			pEX-K4	Kanamycin	2,3,4,5 &
WcbE	Putative CPS glycosyltransferase biosynthesis	<i>Burkholderia pseudomallei</i>	pNIC28-Bsa4	Kanamycin	
			pEX-K4	Kanamycin	2,3,4,5 &
WcbH	Putative CPS glycosyltransferase biosynthesis	<i>Burkholderia pseudomallei</i>	pNIC28-Bsa4	Kanamycin	
			pUC57	Ampicillin	2,3,4,5 &
GmhA		<i>Burkholderia pseudomallei</i>	pEX-A	Ampicillin	2,3,4,5&9
WcbL		<i>Burkholderia pseudomallei</i>	pEX-K	Kanamycin	2,3,4,5&9
WcbN		<i>Burkholderia pseudomallei</i>	pEX-A	Ampicillin	2,3,4,5&9

WcbM		<i>Burkholderia pseudomallei</i>	pEX-A	Ampicillin	2,3,4,5&9
WcbJ		<i>Burkholderia pseudomallei</i>	pEX-A	Ampicillin	2,3,4,5&9
WcbK		<i>Burkholderia pseudomallei</i>	pEX-A	Ampicillin	2,3,4,5&9
WcbI		<i>Burkholderia pseudomallei</i>	pEX-K4	Kanamycin	2,3,4,5 & 9
Cmk^m	Cytidylate kinase	<i>Eschericia coli</i>			
Gmk^m	Guanylate kinase	<i>Eschericia coli</i>			
Cmk^t	Cytidylate kinase	<i>Thermus thermophilus</i>			
Gmk^t	Guanylate kinase	<i>Thermus thermophilus</i>			
Tmk^t	Thymidylate kinase	<i>Thermus thermophilus</i>			
Umpk^t	Uridylate kinase	<i>Thermus thermophilus</i>			

Notes

1: N-terminus His-tag in sequence

2: With BioBricks™ 5' and 3' restriction sites

3: Stop codon

4: pBAT4 RBS

5: Spacer

6: C-terminus FLAG-tag

7: C-terminus HA-tag

8: C-terminus His-tag

9: C-terminus Myc-tag

m: mesophilic

t: thermophilic

*see Appendix 5 for sequences

Appendix 3 – Protein Amino Acid Sequences

Gene name	Amino acid sequence
APY	MGHHHHHSSGVDLGTENLYFQSMRSSYRVGNPIRFQPTNVVGLLLLVLVLSFMLVQSYELGHASGETNANSKYPLTTPVEENLKVRFKIGVISD DDKNAVSKDESNTWVSTYLTGTLEWEKSTDKITVQWDKGNKVKVSKYSYGGRMELSELVTFNGNLLTFDDRTGLVYILKDDKVYPWVVLADG DGKNSKGFKSEWATEKAGNLYVGSSGKEWTTKEGTIENYNPMWVKMINKNGEVTSLNWQTNYEKIRSSMNITFPGYMWHEAACWSDKYNKWF FLPRALSQEAYDSKKFETQGANVIISCDDKFEKCEPTQIQGKTEDKRGFSNFKFVPTSEDKIIVGLKTVEADTTETTYFTAFLDLEGKVLLEETKIDDH KYEGVDFV
WbaD	MSNIDINKKQILVLTFRFPFVIGDRLRIYKICKELSKHYDLTLLSLCDKREELNYEYDREVFSSVHRVYLSKKKSILNVIFSLFSNTPLQIGYYKSKE FEDKLKQLPEHSATLSHLIRVGDYVKENKDINFLEMTDAISLNYKRVEKASLLSLKTFVYSFEQKRLERYERTINNKFSLLTTLVSQVDSYLYPDR PNNVLVCGNGVDAVSLPFSERKIAKDKKITLVFIGNLYSLQNMDGVRWFTKEVLPFLNKHGNFEFKVIGRITDKDKSWLESQPGVVVTGEVDSITYA AADGHIGVCPIRLGAGIQNKVLEYMALGLPCISSTVGFEGLGAEEGKEIYVANTKEEYLRVLNYFITNLDKYTETALVAKKFIGENFSWEAKLSPYIQ KIKESVK
WbnI	MVINIFYICTGEYKRFFDKFYLSCEDKFIPEFGKKYVFTDSDRIYFSKYLNVVEVINVEKNCWPLNLTLLRFSYFLKVIDKLQTNSTYFFFNANAVIVKEI PFSTFMESDLIGVIHPGYKNRISILYPWERRKNATCYLGYLKKGIYYQGCENGKGTASFKRLIQICNMMTMADLKKNLIAKVHDESYLNNYYNKP LLLSELYSWPEKYGENKDAKIIMRDKERESWYGNIKK
WbuP	MRCKHAVIMCIYNGDNVNNVLAVDLSILNQTVNNHLYIYVDGEVNPFLNTYLNKEESNHQIFIFRSEFNLGLAYGLNFLIDKIQSKSYEYISRMDGDDI SLDRLERQELYNKHDPDVIDGSYCTEFGSDCSLPVKKLPQYHDELVRFSVLRCPHPTVMFRAKVFKHGLRYPINTYLSDELALWYEMIYGLK FGNVPDVLKYLRTDSTISRKGLKAGNEIALRFSYMKMKMLITLSNLILLSFKFISHFMPLIITKIMYKCLR
WbbK	MGKSIVVSAVNFTTGGPFTILKKFLAATNNKENVSFIALVHSAKELKESYPWVKFIEFPEVKGSWLKRLHFEYVVCCKLSKELNATHWICLHDITAN VVTKKRYVYCHNPAPFYKILFREILMEPSFFLFKMLYGLIYKINIKKNTAVFVQQFWMKEKFIKKYSINNIIIVSRPEIKLSDKSQLTDDDSQFKNNPSE LTIFYPAVPRVFKNYELIISAARKLKEQSNIKFLLTISGTENAYAKYIISLAEGLDNVHFLGYLDKEKIDHCYNISDIVCFPSRLETWGLPLSEAKERGK WVLASDFPFTRETLGSYEKKAFFDSNNDMLVKLIIDFKGNLKKDISDANFIYRNENLVGFDELVNFITEH
WaaB	MKIAFIGEAVSGFGGMETVIRDVITFRQQHIQSEMFFFCRNDKMDKGWLEGIKYSCSFSNIRLGLFRRAKHIHALSKWLQEYQPDIVICIDVISCLFA AKARKKSGIDMPVFSWPHFSLDHKKHAEYITCADYHLAISSGIKQQMINRGVAESTINVIFNPVETKDSVIPAPEEGETATFIYVGRMKFEGQKRVK DLLDGLSQAQGNWKLHVLGDGSDFEKCAQYGRELNIDDRIVWYGWQQYPWELVQQDIEKVSALLTSSFEFGFMTLLEALSWGIPCISADCVSG PADIIQPDVNGHLYQPGDIAGFVTLN KYIAGEIHIEHEKIPASIDEFYQSKYYDRHLHKVIISAISRRK
WffQ	MKIILWHLGRKGAGPKYNLEMALALKEGYDVYACISKQAENANDYKKNINHLNNTYTGVSALINSLRPLWLINEFKDYLKINRITHVYCTMPHI WNFFFSKRFKLDVSYLLTIHDAELHSGEENKLIDILMKDRKNADGIVVLTKHVKEQLKNLGSIFIIHPAFSNNNKNRAKSIVPDQKIKLLFFGRI HHYKGLDILIDAFEILNKKNNRYSLSIYSGNISPYQKKILELSNITVHNWIDWVEIPSIISSHDVCVIPYRDASQSGVIPTVMTCGVPLVVTVPVAGLRE QVVHLQTAIFAEVSSDSVARAIDSLVKDISLYEKISKNALLYAQEEMGWEKAAYRSVEALISITHNSKKNYQ

WbyB	MNILIVYLGRKGGGAHYTYELVNELSKKTNVSLLVSKQAENIKSYESLNISLNTINTFNGMFSFILSSFKLPFIKNKIKKIILDEKIDIVISTMTHVWNGFIF TPNFMKIPYILTIHDAVPHSGEENFILNLLNNKDIKNANAYIALTEHVKNQFINIYGDKKDICILPLGSFNFKSGEIKSWCGVRPLRILFFGRIHEYKGI DLLLDSMILQEKYDLELSIYGQDLSYVDKIDVIKINIENRWIGDSEVGDILQAHDVCVIPYKDAQSGVIPVAQSAGLPTIVNPVDGLIEQIEKDKT SLVTDRVDSVSLALSIEKIIKNPELVNKL SIGSLEYASRKL SWVPIV NKLITYSKETLNSSDRS
WcbB	MSIYLDVTRLVTRLYQGLLPTGVDRVGLQYIRHYGSRARAVLSERGFFAILSERDSALVFAWLTSSIGNKNAIFRLAAHAACLSIFNTSFQNGILLHT SHSGMEFPRYYKKLASLGIKSVFLIHDLIPLTHAEYTRPGVEHTRRRRIHTALGYASGLIANSRSTLES LAEAATRAALPLPPCAIAHLASGVEPQPP RQRLLDAPYFVMLGTIEPRKNHWFILHVWRR LIEQLGNAAPKLVVIGRRGWECENVIDMLERCASLPGTVIEEANCSDERLHAWLQHARALLFPSPF VEGYGMPLVEALGLGVPVLASDLDFREVAADIPDYLDPLDGPAAARIRDYARDDSHERTAQLARIEHFREPTWVDHFERVDAFLDTRLR
WcbE	MKNKVRVLELRLPALDGYSGIPQETRLLFRTLLGLKKMAVTGLLQQGGRDITAKMVYAAESERPAAQIRAMSKIVVSFTDDGRRGVAGAISRAISIV MRGISLRASVIAGLSEKLGFFAPVEFEDFVWRRFFAKTLPVDDRPHVLAQAQYRVLPLAMATLHRLKISLSGLLRVKRFKKIETRDFDVFISQTPFPGV LSSNTKLIVRYHDAIPIFLPHTIKDRAFHHGAHYDALARNVRSGAYFACVSAATRNDLLKIFPEIEDRAVVIHNTVSP IYSEIDSPKELVKDIVSSRLFR PLAKKYKALSSSHVDVFRSRLNADQQDFEYLLMVSTVEPRKNHTSLVAAWELLRSRHNEALKLVLVGDLGWDYAEFVDAIVPWVERGELFLLSN VAAPDLRDLYKHAKVTICPSYAEFGDYSGVEAMMSGGVVAASDIPVHREVYEDACIYFN PYSVESIVGGIREAISLRNASETFAEYRRC AKRISARY TAEAISPRWDEFIQSVGSAP
WcbH	MKLAFFTFPHFACGPGFMSKAVVEHLKDYFDVTVFYQWHALHQPYEVEDVSVRIIEDDMDMEPVAREFD AVIYNLGNNEENHYSIFQCLKRLPGI VVLHDFVMQHGIVNDFNRKQKSDLYWLLAALYGKDGVRACHASALSYAGGVRGGWDSVSVASFPMFEVFSGLASACVVHSKFFENQLKPHF AGPILMTRNPYDLKRVPTEDVIETFYKVRGKAGGKVVFAAFGHMSPSKCIDRVLRVFGHSELRKEARLILCGGASPEYDQYLRRLCDAGDLGHC VEFRGSVSDTELYALQVEADVFNLRQPNTREGQSGSLIEQLAAGKPVICYDTGCGYDLS EDVAYKVHSTTDFNELQRVFERLLNGVAERREVG LR GRAAALEYDCARYARDMF EFVFENREYLNAFQRRNFDGLRPLGVSGTTIDTRHPENYSRTQSIWGVLDG FVHHASSEALSLIPGDARHHYLSFLT QPYPRLVDSLLELSASDVPEVELEADGTSLYSTASAPVLNTSFWTKLKCPVSEHFAYIGYKLLCRDPDLA GLRDYAEKLASETISRKEMILAMLRS TEFAERFGDIRKSSDYVSLIRWAGAQA
GmhA	MENRELYITNSIAEAQRVMAAMLADERLLATVQKVADACIASIAQGGKVLLAGNGGSAADAQHIAGEFVSRFAFDRPGLPAVALTTDTSILT AIGN DYGYEKLF SRQVQALGNKGDVLIGYSTSGKSPNIIAAFREAKAKGMTCVGFTGNRGGEMRELCDLLELVPSADTPKIQEGHLVLGHIVCGLVEHSI FGKQ
WcbL	MNPTIIRARAPRLRLGLAGGGTDVAPYADTFGGYVLNATIDRYAYAVIKTLTIPAVRFVST DQQVEKHQLISEPLELNGTLNLHKAVYNHMIRNYNHGK PIALELSTFCDAPAGSGLGSSSTLVVVMIKAFVELLNLPLDDY AIAQLAYRIERVDCGLAGGRDQYSATFGGFNFMEFYEEERTIVNPLRIKNWVL CELEASLVFYTGVSRSAKIIQDQSDNVVSHKTA AIEAMHGKREALVMKEALLKGDFKAFVASMRLGWDNKKNSARTVSN AHIDEIYDAAIRAGA QAGKVS GAGGGFMLFFVPTEKRMDLIRTLGEYDGGVSNCHFTKNGTQAWRIAN
WcbN	MKNRALFLDRDGVINRDDGYVFEIEKFVFLDGIFELAGA AAKALGYLSIVVTNQAGIGRGYSEDDFFRLSDWMKGVFATEGAPIDGVYFCPTHPEH GIGRYKVESRFRKPNPGMILAAQHDFDLDLGASLLVGDKESDIQAGSTAGVGTLLICDRDASRVATAASAVVRNPRDVIPFLTGP GPDAGSF
WcbM	MREAILLAGGFTRLRTVSDVPKPMAPIAGRPFLEILLTRLSEK KFSRVVLSVGFMAEKIMSHFGDRFAGIDLAYSVESDPLGTGGALKATLPYCE GDHAFVFN GDTYLDLEVDELDDGWQTGGFPTIVARQVPDTGRYGR LVVDGGRVTGFAEKGVSGPGLINAGCYVLPKDILAGETAETFSFETDFM SSAVQSRRFDV FVTRGQFIDIGVPEDFYRAQDELSGICK

WcbJ	MKVFLVGSTGYIGKTLFDACSRWRWTLGTSTRDGADIVFSLARAEAFPYEQVSAGDVVAVAAAISPPDACAKDYETAFQVNVGTGLTLIRGVVARG ARVIFSSDTVYGASEQLLSEEAELTPAGAYGAMKRRVEAELGENAAVKVIRLSYVFSLRDRFTQYLLGCAKEGKRADIFKPFRCVVYLSDVVEG VVSLIERWDAIDERVINFVGPPELVAREDFVEKIRNLAPELDYGFSEPEGDFVNRPRIINVSSARFEKLLGRRPRSLGEAINLELSN
WcbK	MAKRVLITGITGMVGSHLADFLLENTDWEIYGLCRWRSPLDNVSHLLPRINEKNRIRLVYGDLDYLSIHEAVKQSTPDFVFHLLAAQSYPKTSFDSP LDTLETNVQGTANVLEALRKNNDIVTHVCASSEVFGRVPREKLPIDEECTFHPASPYAISKVGTDLIGRYAAEAYNMTVMTTRMFTHTGPRRGDV FAESTFAKQIAMIERGLIPPVVKTGNLDSLRTFADVRDAVRAYYMLVTINPIPGAYYNIGGTYSCTVGQMLDTLISMSTSKDVIRVETDPERLRPIDAD LQVPNTRKFEAVTGWKPEISFEKTMEDLLNYWRARISAGEKFLTR
Wcbl	MMAQRKKYSVYGSCQAPALAKMLNSCPTFARDWELVEMEPCFVASEEQIDRHLAETIPKLDLFLYQPVSEGYRGEKYSSVFLRNSMPPGGNALS VQYMHWEGYHPTVNSPYGLPPHPEGYVDALIAGAVVMDVDKETYLRHLEEIGASLRIDIDEIESWCDELKTREVGENDGGKQIDISVTDFILANC RQKRLFYTMNHPTAALMREIAARCMLALGYTYSDISFDQNLDPDVTKMSLYPIYRDCDFDFSELNRMNEYQVLYKKKAYEPYLLEQFEWFERSPK ADVSAFFDRVAANRRWVRTALRRAFES
Cmk^m	MTAIAPVITIDGPSGAGKGTLCAMAEALQWHLLDSGAIYRVLALAALHHHVDVASEDALVPLASHLDVRFVSTNGNLEVILEGEDVSGEIRTQEVA NAASQVAAPFRVREALLRRQRAFRELPLIADGRDMGTVPDAPVKIFLDASSEERAHRRMLQLQEKGFSVNFERLLAEIKERDDDRDRNRAVAP LVPAADALVLDSTTLSIEQVIEKALQYARQKLALA
Gmk^m	MAQGTLYIVSAPSGAGKSSLIQALLKTQPLYDTQVSVSHTTRQPRPGEVHGEHYFFVNHDEFKEMISRDAFLEHAEVFGNYGTSREAIEQVLATG VDVFLDIDWQGAQQIRQKMPHARSIFILPPSKIELDRRLRGRGQDSEEVIAKRMAQAVAEMSHYAEYDYLIVNDDFDALTDLKTIIRAERLRMSRQ KQRHDALISKLLAD
Cmk^t	MRGIVTIDGPSASGKSSVARRVAAALGVPYLSSGLLYRAAFLALRAGVDPGDEEGLLALLEGLGVRLLAQAEGNRVLADGEDLTSFLHTPEVDRV VSAVARLPGVRAWVNRRLKEVPPPFVAEGRDMGTAVFPEAAHKFYLTASPEVRAWRRARERPAQAYEEVLRDLLRRDERDKAQSAPAPDALVLD TGGMTLDEVVAWVLAHIRR
Gmk^t	MKGMRGRLFVMTGASGVGKGTVRAKVLERTRLFYSISMTRPPRPGEVDGVDYFVDRPTFEALVREDGFLEYAEYVGHLYGTTPRAPVERALSR GEDVLEIEVQGALQVKRAVPEAVLIFLLPPSLSELKRRLVYRGKDSPEKIQRLEQAEWEIRNAHLFDYVVVNDVLEEAVADFLAILTAERRRSGR MGEALEMALRRDLALEAELDEILRRRYGGTGH
Tmk^t	MPGLFLTLEGLDGSKTTQARRLAFLAQGRPVLLTREPGGGLPEVRSLLLTQELSPEAEYLLFSADRAEHVRKVILPGLAAGKVVISDRYLDSSL AYQGYGRGLPLPWLREVAREATRGLKPRLTFLLDLPEAALRRVRRPDRLEGLGLEFFRRVREGYLALARAEPGRFVVLDTLPEEEIARAIQAHL RPLL
Umpk^t	MKYKRVLLKLSGEFLTRNGFGIEPEATQALAREIKAAYDTGVQLAIVIGAGNLWRGARQGVGMDRATADYIGMLATIMNALALQDALES LGVPTRV QTALTITQVAEPYIRRRALRHLEKERIVIFGGGTGNPFFSTDTAAALRALEVGAEVVLMKNKVDGVYSDDPRKNPEAVRFDELTYLEVLNRGLQV MDTTAITLCMEAGLPVVFDIFKPGALVGIQGEKVGTLIH

Notes

1 = from synthesised gene m = mesophilic t = thermophilic

Appendix 4 – Protein Nucleic Acid Sequences

Gene name	Nucleic acid sequence
<i>Apy</i>	CCATGGGCCATCATCACCATCACCATAGCAGTGGTGTGGATCTGGGTAAGTAAACCTGTATTTTCAGTCTATGCGTTTCGAGTTATCGCGTAGGAAATCCGATTTCGCTTTCAACCTACGAACGTCGTTGGGCTGTTGCTGCTGAGTCTGGTGTGAGCTTTATGTTAGTACAGTCTTACGAGCTCGGGCATGCTTCAGGAGAAACGAACGCGAACTCGAAGTATCCCCTTACCACTCCTGTTGAGGAGAACCTGAAAGTGCCTTTAAAATTGGCGTCATTTCCGATGACGATAAGAATGCTGTTTCTAAAGACGAAAGCAATACATGGGTGTCCACCTATCTTACGGGTACACTGGAATGGGAGAAATCGACTGATAAAATCACCGTCCAATGGGATAAAGGCAACGAAAAGAAGGTGAAAAGCAAATACAGCTATGGTGGTCGTGGCATGGAACCTTCAGAACTGGTAACCTTCAATGGGAATTTACTGACCTTTGATGATCGTACCGGCCTGGTGTACATTCTGAAAGACGATAAAGTGTACCCGTGGTTGTGTTAGCGGATGGTGACGGCAAGAATAGCAAAGGATTCAAATCAGAATGGGCCACTGAGAAAGCAGGCAACCTGTATGTAGGCAGTTCTGGTAAAGAATGGACCACGAAAGAAGGCACGATTGAAAACCTACAATCCGATGTGGGTCAAATGATCAACAAGAATGGGGAAGTGACCTCATTGAATTGGCAGACGAACTATGAGAAAATTCGCAGCAGTATGAACATTACGTTTCCGGGTTATATGTGGCACGAAGCAGCGTGTGGTCCGACAAATACAACAAATGGTTCTTTCTCCCACGGGCGTTATCGCAAGAGGCGTATGACTCCAAGAAATTCGAAACCCAGGGAGCCAATGTCATCATCTCGTGCGACGATAAATTCGAGAAATGCGAACCGACTCAGATTCAGGGTAAAACCGAAGATAAACGCGGTTTCTCCAACCTCAAATTCGTGCCAACAAGCGAAGATAAAATCATTGTTGGCCTGAAAACCGTTGAAAGCAGACGATAACCACGGAGACATACTTTACCGCCTTTGATCTGGAAGGCAAGGTTCTCTTGGAGGAAACCAAATCGATGATCACAAAGTATGAAGGCGTGGACTTTGTCTAAAAGCTTC
<i>wbnI</i>	GAATTCTCTAGAGAAGGAGATATATCCATGGTCATTAACATCTTTTATATCTGTACTGGGGAATACAAACGTTTCTTCGACAAGTTTTACCTGCTTTGTGAGGACAAGTTTATCCCCGAGTTTGGTAAGAAATATTACGTCTTTACGGACTCAGATCGCATCTATTTCTCCAAATATCTGAATGTGGAAGTCATCAACGTTGAGAAGAATTGTTGGCCACTGAATACCCTGTTACGCTTTAGCTACTTTCTTAAGGTCATTGACAAGTTGCAGACGAACTCGTATACCTTCTTCTTCAATGCGAATGCGGTAATTGTGAAAGAGATTCCGTTCTCGACCTTCATGGAATCGGATCTGATCGGTGTGATTCATCCAGGCTACAAGAATCGGATTTCCATTCTGTATCCGTGGGAACGTCGCAAGAATGCTACCTGCTATTTAGGCTATCTCAAGAAAGGCATCTACTATCAGGGTTGCTTAAATGGTGGGAAAACCTGCCAGTTTTAAGCGTCTGATTCAGATCTGCAACATGATGACGATGGCTGATCTCAAGAAGAACCTCATTGCGAAAAGTTCACGATGAAAGCTATCTGAATTACTACTATTACTACAACAAACCGTTACTGCTGAGCGAATTGTACAGTTGGCCAGAGAAATATGGCGAGAATAAAGACGCGAAAATCATCATGCGTGATAAGGAACGTGAAAGCTGGTATGGGAACATTAAGAAAGATTACAAGGATGACGACGATAAATAAAGTCTGACG
<i>wbaD</i>	GAATTCTCTAGAGAAGGAGATATATCCATGAGCAACATCGACATCAACAAGAAACAGATTCTGGTCTTGACGCCACGTTTTCTTTTCCCGTATTGGCGGTGATCGCTTACGCATTTACAAAATCTGCAAAGAAGTGTGCAAGCATTATGATCTGACCCTGTTAAGTCTTTGCGATAAACGCGAAGAATTGAATTACGAATACGATCGTGAAGTGTCTCTAGTGTCCATCGCGTTTATCTGAGCAAGAAGAAATCCATCCTGAACGTCATCTTTCTCTGTTCTCGAATACTCCGCTTCAAATCGGCTACTATAAGTCGAAAGAATTTGAGGACAACTGAAGCAGTTGCTGCCTGAACACTCAGCCACGCTGTCTCATCTCATCCGTGTTGGAGACTACGTGAAAGAGAACAAAGACATTAACCTTTCTGAAATGACCGATGCGATTTCACTGAATTACAAACGGGTGAAAGAGAAAGCATCTCTGTTGAGTCTGAAAACCTTCGTTTATTCGTTGAGCAGAAACGCTTAGAACGCTATGAACGTACGATCA

	ACAACAAATTCTCCCTCACAACCCTGGTTTCACAAGTGGATAGTGATTATCTCTACCCAGATCGTCCGAATAACGTCTTAGTTTGCGGCAATG GGGTTGATGCTGTGAGTCTGCCGTTTAGCGAACGCAAATTGCCAAAGATAAGAAAATTACACTGGTCTTTATCGGCAACCTGTATTCCTC AGAACATGGATGGCGTACGCTGGTTTACGAAAGAGGTGCTGCCGTTCTTGAATAAACACGGAAATTTGAGTTTAAAGTGATTGGTCGTATT ACCGACAAAGACAAAAGCTGGCTGGAAAGCCAACCAGGTGTGGTAGTAAGTGGCGAGGTAGACAGCATCACGTATGCTGCCGCAGATGGT CACATTGGTGTCTGTCCGATTCCGGCTGGGTGCGGGTATTGAGAATAAAGTGCTGGAATACATGGCTCTGGGGTTACCGTGTATTAGCTCCA CTGTGGGGTTTCAAGGCTTAGGTGCCGAAGAAGGCAAAGAAATCTATGTGGCAAATACCAAAGAGGAATATTTGCGCGTCTCAACTATTTT ATCACC AATCTGGATAAATACACCGAAACAGCGCTGGTTGCCAAGAAATTCATTGGCGAAAACTTTTCTGGGAGGCGAAAACTTTTCGCCGTA TATTCAGAAAATTAAGGAAAGCGTTAAATACCCGTATGATGTGCCCGACTATGCGTAAACTAGTCTGCAG
<i>wbuP</i>	GAATTCTCTAGAGAAGGAGATATATCCATGCGCTGCAAACATGCCGTGATTATGTGCATTTATAACGGCGATAATGTCAATAACGTGAAATTA GCGGTGATAGTATTCTGAATCAGACCGTGAATAACCACCTGTACATCTATGTAGACGGTGAAGTTAACCCGTTTCTCAATACCTATCTGAAC AAAGAAGAATCGAACCATCAGATCTTCATCTTTGAGCGAGTTTAACTTAGGTCTTGCATGATGGTCTGAACTTTCTGATTGACAAAATTCAGT CGAAAAGCTACGAATACATCTCACGCATGGATGGCGATGATATTAGCTTGTAGATCGCCTTGAACGCCAAGAGCTGTACCTGAATAACAT CCTGACGTTGACGTAATCGGCTCGTATTGCACCGAATTTGGGAGTGACTGTTCACTGCCAGTGAAGAAATTACCGCAATACCACGATGAACT GGTTCGCTTCTGTGCTCCGTTGTCGCTGATTACCCAACGGTGTGTTTCGTGCCAAGGTATTCAAACATGGACTTCGCTATCCGATTA ACACCTATCTGTCCGAAGATTTGGCACTCTGGTATGAGATGATCTATCTCGGTCTGAAATTTGGCAATGTTCCCGATGTCCTGTTGAAATATC GCCTGACTGACAGCACTATTTCCCGTCTGTAAGGGCTGAAGAAAGCGGGCAATGAGATTGCTCTTCGGTTTTCTACATGAAGAAAATGAAA CTGATTACACTGAGTAACCTGATTCTGTTGAGCTTCAAATTCATCTCTCACTTCATGCCGTTAATTATCACGAAAATCATGTACAAGAACTGC GGCATCACCATCATCACCATTAACTAGTCTGCAG
<i>wbbK</i>	GAATTCTCTAGAGAAGGAGATATATCCATGGGAAAATCCATTGTGCTGTTGTTCCGCTGTGAATTTACCACAGGAGGACCGTTTACGATCCTT AAGAAATTCCTGGCAGCGACTAACACAAGGAAAATGTGAGCTTTATTGCACTGGTTCATAGCGCAAAGAAGTCAAGGAAAGCTATCCGTG GGTAAAATTCATCGAGTTTCCGGAAGTCAAAGGGAGTTGGTTAAAACGTCTGCATTTTGTAGTATGTGGTCTGTAAGAAGCTGTGAAAGAGT TAAATGCCACCCATTGGATCTGTCTCCATGACATTACAGCGAATGTAGTAACGAAAAAGCGGTATGTGTACTGTCACAATCCAGCCCCGTTTT ATAAAGGCATTCTTTCCGTGAAATCCTCATGGAACCGTCTTCTTTTCAAGATGCTGTATGGACTGATCTACAAAATCAACATCAAGAA GAACACCGCAGTCTTTGTGCAACAGTTCTGGATGAAGGAGAAATTCATCAAGAAATACTCGATCAACAACATTATTGTGTCTCGTCCGAAAT CAAACGTGCGGACAAAAGCCAATTAACGGATGACGACAGTCAGTTCAAGAATAATCCGTCCGAGTTGACCATCTTCTATCCGGCAGTTCCGC GCGTGTCAAGAATTACGAACTGATCATTTAGCTGCTCGGAAACTGAAGGAACAGAGTAACATTAAGTTCCTGTTGACCATTTAGGGACT GAAAATGCCTATGCCAAGTACATCATCTCATTAGCCGAAGTCTGGATAATGTCCACTTCTGGGTTATCTGGACAAAAGAAAAGATTGACCAT TGCTACAACATTTCCGATATTGTTTGTCTTCAAGCCGTTTGGAAACATGGGGACTTCCCTGAGTGAAGCGAAAAGACGTGGGAAATGGGT ACTTGCTTCAGACTTCCATTTACACGCGAAAACCTTAGGCTCTTATGAGAAGAAAGCGTTCTTCGATAGCAACAACGATGACATGCTGGTCAA GCTGATCATTGATTTAAGAAAGGCAACCTGAAGAAAGACATCTCTGACGCTAATTTATCTACCGCAACGAGAATGACTTGTAGGCTTCGA TGAGCTTGTTAACTTTATTACGGAGGAACACTATCCGTATGATGTTCCCGATTACGCCTAACTAGTCTGCAG
<i>waaB</i>	GAATTCTCTAGAGAAGGAGATATATCCATGGGAAAATCCATTGTGCTGTTGTTCCGCTGTGAATTTACCACAGGAGGACCGTTTACGATCCTT AAGAAATTCCTGGCAGCGACTAACACAAGGAAAATGTGAGCTTTATTGCACTGGTTCATAGCGCAAAGAAGTCAAGGAAAGCTATCCGTG

	GGTAAAATTCATCGAGTTTCCGGAAGTCAAAGGGAGTTGGTTAAAACGTCTGCATTTTGAGTATGTGGTCTGTAAGAAGCTGTCGAAAGAGT TAAATGCCACCCATTGGATCTGTCTCCATGACATTACAGCGAATGTAGTAACGAAAAAGCGGTATGTGTACTGTCACAATCCAGCCCCGTTTT ATAAAGGCATTCTTTTCCGTGAAATCCTCATGGAACCGTCCTTCTTTCTTTTCAAGATGCTGTATGGACTGATCTACAAAATCAACATCAAGAA GAACACCGCAGTCTTTGTGCAACAGTTCTGGATGAAGGAGAAAATTCATCAAGAAATACTCGATCAACAACATTATTGTGTCTCGTCCTGAAAT CAAAGTGTCCGACAAAAGCCAATTAACGGATGACGACAGTCAGTTCAAGAATAATCCGTCCGAGTTGACCATCTTCTATCCGGCAGTTCCGC GCGTGTCAAGAATTACGAACTGATCATTTCAGCTGCTCGGAACTGAAGGAACAGAGTAACATTAAGTTCCTGTTGACCATTTTCAGGGACT GAAAATGCCTATGCCAAGTACATCATCTCATTAGCCGAAGGTCTGGATAATGTCCACTTCTGGGTTATCTGGACAAAAGAAAAGATTGACCAT TGCTACAACATTTTCGGATATTGTTTGCTTTCCAAGCCGTTTGGAAACATGGGGACTTCCCCTGAGTGAAGCGAAAAGAACGTGGGAAATGGGT ACTTGCTTCAGACTTCCCATTTACACGCGAAACTTTAGGCTCTTATGAGAAGAAAGCGTTCTTCGATAGCAACAACGATGACATGCTGGTCAA GCTGATCATTGATTTTAAGAAAGGCAACCTGAAGAAAGACATCTCTGACGCTAATTTTCATCTACCGCAACGAGAATGACTTGTAGGCTTCGA TGAGCTTGTTAACTTTATTACGGAGGAACACTATCCGTATGATGTTCCCGATTACGCCATAACTAGTCTGCAG
wffQ	GAATTCTCTAGAGAAGGAGATATATCCATGAAAATCATTCTGTGGCATCTGGGTCGTAAAGGAGCAGGCCCAATACAACCTGGAAATGGC CTTAGCTCTGAAGAAAGAGGGATACGACGTGTATGCATGCATTTTGAAGCAGGCAGAAAACGCCAATGACTACAAGAAGCATAACATCAACC ACTTGTCATTGAATACCTATACAGGTGGTGTATCAGCCTTGATTAATAGCTTACGCTTACCGTGGCTGATTAATGAGTTCAAAGACTATCTGA AAATCAACCGCATTACCCACGTGTATTGCACAATGCCGCATATCTGGAATTTCTTCTTCAGTAAGCGCTTTAAGAACTTGATGTCAGCTATC TGCTGACCATTGATGATGCTGAGCTGCATAGCGGTGAGGAAAATAAACTGATCGACATCCTGATGAAGAAAAGATCGGAAGAATGCAGATGG GATTGTGGTGCTGACCAAACATGTCAAAGAGCAGCTCAAGAATTTAGGGAAAGTCGATCTTCATTATTCCCATCCAGCGTTTTTCATTTAACAA CAACAAGAATCGTGCCAAGTCGATCGTGCCAGATCAGAAGATTAACCTCTTCTTTGGTCGGATTGATCACTATAAAGGCCTTGATATCTT AATTGACGCGTTTGAGATTCTGAACAAAAAAAACAATCGCTATTCTTATCGATCTACGGTTCTGGCAACATTAGTCCGTATCAGAAGAAAATT TTGGAAGTGTCCAATATTACGGTCCACAATCGTTGGATTGATGACGTTGAGATCCCTAGTATCATTAGCTCACACGATGTGTGTGTCATTCCG TATCGTGATGCTAGCCAATCTGGGGTAATTCCGACTGTCATGACTTGTGGAGTACCCTTAGTCGTCACACCCGTTGCAGTTTTACGCGAACA AGTGGTACATCTGCAAACCGCGATCTTTGCGGAGGAAGTGTCTAGTGATTGAGTGGCTCGCGCATTGATAGCCTGGTGAAAGACATTTCCC TGTACGAGAAGATCAGCAAGAACGCTCTTCTGTATGCCAAGAGGAAATGGGCTGGGAGAAAAGCGGCTTATCGCTCTGTTGAAGCGCTGAT CAGCATTACGCACAACCTCAAAGAAAATACAACCAGCATCATCACCACCATCATTAACCTAGTCTGCAG
wbyB	GAATTCTCTAGAGAAGGAGATATATCCATGAATATCCTTATCGTCTATTTGGGGCGTAAAGGTGGTGGCGCACACTACACTTACGAACTGGT TAATGAGCTGTCAAAGAAAACGAACGTTTCCCTTCTGGTTAGTAAACAGGCTGAAAACATCAAGTCCTATGAATCTTTAAACATCTCTCTTAAC ACCATTAACACGTTTAATGGAATGTTTAGCTTTATTCTGTCTCCTTCAAAGTCCGTTTATCAAGAACAAAATCAAGAAAATCATTCTGGACG AGAAAATCGACATCGTGATTAGCACAATGACCCATGTGTGGAATGGCTTATCTTTACGCCGAATTTTATGATGAAAATCCGTACATTTTGA CCATTCACGATGCCGTACCACATTCTGGCGAAGAAAATTTTATTCTGAATCTGCTGAACAACAAGACATCAAGAACGCGAATGCGTATATTG CGTTAACTGAACACGTGAAGAATCAGTTCATCAACATTTATGGGGACAAGAAGGATATCTGCATTTTACCTCTGGGCTCGTTTAATTTCAAAT CAGGCCGAAATCAAAGCTGGTGTGGAGTACGTCCATTGCGGATTTTATTCTTTGGTCGCATTCACGAGTATAAAGGCATCGACTTACTTTTG GACAGTATGATTATTCTGCAAGAGAAAATATGATCTTGAAGTACGATTTATGGGCAAGGGGATCTCAGCGACTATGTCGACAAGATCGATGT GATTAAGAAGATTAACATTGAAAATCGCTGGATTGGCGATTCCGAGGTGGGTGATATCTTACAGGCGCATGACGTATGCGTAATTCGGTACA

	AAGATGCCTCCCAATCAGGCGTAATTCCTGTTGCACAATCAGCAGGACTCCCCACTATTGTCAACCCTGTTGACGGGTTAATCGAACAGATT GAAAAGGACAAGACCAGCTTAGTCACAGATCGTGTGATTCCGGTGAGTCTGGCTCTCTCTATTGAGAAAATCATTAAAGAATCCGGAAGTGGT GAATAAACTGAGCATTGGCAGTCTGGAATATGCCAGTCGCAAAGTGTGATGGGTGCCGATTGTGAACAACTGATCACCTATTCAAAGAAA CCCTCAACAGCAGTGATCGGTCAGATTACAAAGACGATGACGACAAGTAACTAGTCTGCAG
<i>gmhA</i>	GAATTCTCTAGAGAAGGAGATATATCCATGGAAAATCGCGAGCTGACGTACATCACGAACAGCATCGCTGAAGCCCAACGGGTAATGGCCG CTATGCTTGCCGATGAACGCTTACTGGCAACTGTTGAGAAAGTCGCTGATGCGTGCAATTGCGTCCATTGCCCAAGGTGGTAAAGTGCTCTTA GCGGGCAATGGAGGCTCTGCCGCAGATGCCAGCATATTGCAGGCGAGTTTGTCTCGCGCTTTGCGTTTGATCGTCTGGTCTGCCAGCC GTGGCGTTAACCAGTACACCAGCATTCTGACAGCGATTGGCAACGACTATGGGTATGAGAACTGTTGAGTCGCCAAGTTCAGGCATTGG GGAACAAAGGCGACGTGCTTATTGGCTACTCAACGAGTGGGAAATCGCCGAACATTCTGGCTGCATTCCGTGAAGCGAAAGCGAAAGGCAT GACCTGTGTCGGCTTTACCGGTAATCGCGGTGGTGAATGCGTGAGTTGTGCGATCTGTTGCTGGAAGTTCGTCTGCCGATACCCCGAAG ATCCAGGAAGTTCATCTGGTACTCGGTCACATTGTGTGGACTGGTGAACACAGCATCTTCGGCAAACAGGAACAGAAGCTGATCTCCG AAGAGGACCTGTAACTAGTCTGCAG
<i>wcbL</i>	GAATTCTCTAGAGAAGGAGATATATCCATGAACCCGACGATCATTGCGGCACGTGCACCCCTGCGTCTGGGCCTTGTGGCGGTGGCACC GATGTAGCCCCTTATGCCGATACGTTTGGCGGGTACGTCTGAATGCCACGATTGATCGCTATGCCTACGCGGTGATCAAAACCCCTCACTAT CCCGGCAGTACGCTTTGTGTCCACAGATCAGCAAGTCGAGAAACATCAGCTGATCAGCGAGCCGTTGGAAGTGAACGGCACTCTCAATCTG CATAAAGCGGTGTATAACCACATGATTGCAACTACAATCATGAAAACCGATTGCCCTGGAATTATCCACTTTTGTGATGCGCCAGCTGG GAGTGGTCTGGGGTCATCGTCTACGCTCGTTGTGGTGATGATCAAAGCATTTCGTTGAGCTTTTGAATCTGCCTTTGATGACTATGCCATTG CGCAACTTGCTTATCGCATCGAACGTGTTGACTGTGGCTTGGCGGGTGGCCGTCAGGATCAGTACTCGGCAACCTTTGGTGGCTTTAACTT CATGGAGTTTTATGAAGAAGAACGGACGATTGTCAATCCACTCCGTATTAAGAAGTGGGTGCTGTGCGAATTAGAGGCCAGTCTGGTGCTGT TTTATACCGGCGTAAGCCGCGAAAAGTGCGAAAATTATCCAGGATCAGAGCGACAATGTGGTTTCGCACAAAACAGCGGCTATTGAAGCAAT GCATGGGATTAACCGCGAAGCGCTGGTCATGAAAGAAGCCTTACTGAAGGGAGACTTCAAAGCCTTTGTGGCGAGCATGCGCTTAGGCTGG GACAACAAGAAAACTCTGCTCGCACCGTATCCAACGCACATATTGACGAGATCTACGATGCAGCCATTTCGTGCGGGTGCGCAAGCTGGCA AAGTCTCTGGTGCGGGTGGCGGTGGTTTCATGCTGTTCTTCGTTCCGACCGAGAAACGTATGGACCTGATTCCGACTCTGGGAGAATACGA TGGTCAGGTTAGCAACTGCCACTTTACCAAGAATGGGACACAAGCGTGCCGCATCGCGAATGAACAGAACTGATCTCAGAAGAAGATTTG TAACTAGTCTGCAG
<i>wcbN</i>	GAATTCTCTAGAGAAGGAGATATATCCATGAAGAACCGTGCCTTATTTCTGGATCGTGATGGGGTGATTAACCGCGATGACGGCTATGTGTT CGAGATTGAGAAGTTCGTATTTCTCGATGGCATCTTTGAACTTGCCGGTGCAGCCAAAGCCTTGGGCTATCTCTCGATCGTCGTAACCAACC AAGCAGGCATTGGTCGTGGCTATTACAGCGAAGATGATTTCTTCCGCTTAAGCGATTGGATGAAAGGAGTGGTTCGACCGAAGGTGCGCC AATTGACGGTGTGTACTTTTGTCCGACCCATCCGGAACACGGTATTGGCCGCTACAAAGTCGAATCCCGCTTTGCGAAACCCAATCCGGGC ATGATCTTGGCTGCGCAGCATGACTTCGACCTGGATCTTGGCGCCTCACTGCTGGTTGGGACAAAGAATCGGACATTCAGGCGGGAAGTA CGGCTGGTGGTGGGACCACGCTGCTGATTTGCGATCGTGATGCCAGCCGTGTTGCGACAGCAGCATCTGCGGTGGTTCCGGAATCCACGCG ATGTCATCCCCTTTCTGACTGGTCTGGCCCGGATGCTGGTTCTTCGAACAGAACTGATCAGTGAAGAGGACCTGTAACTAGTCTGCAG

wcbM	<p>GAATTCTCTAGAGAAGGAGATATATCCATGCGTGAAGCGATCATTCTGGCAGGCGGCTTTGGCACACGCCTGCGTACGGTCGTGAGCGATG TCCCTAAACCGATGGCACCGATTGCTGGTCGCCATTCTGGAGATTCTGCTGACGCGCTCTCCGAGAAGAAATTCTCACGCGTTGTCCT GTCTGTGGGCTTCATGGCTGAAAAGATCATGTCCCACTTTGGCGACCGTTTTGCGGGCATTGATCTTGCGTACAGCGTGAAAGTGACCCA TTAGGCACTGGTGGTGCATTGAAAGCGACGTTACCGTACTGCGAAGGTGATCATGCCTTTGTTTTCAACGGTGATACGTACTTGGACTTGGA AGTGGATGAACTGGATGATGGCTGGCAAACCTGGCGGTTTTCCGACCATTGTAGCCCGCCAAGTTCCGGATACAGGACGGTATGGTCGGTTA GTGGTGGATGGTGGTCGCGTAACCGGTTTTGCCGAGAAAGGGTAAGTGGCCCTGGGCTCATTAAATGCGGGGTGTTATGTGCTGCCGAAA GACATCCTTGACAGGAGAAACCGCGGAAACCTTTTTCGTTTGAACCGATTTTCATGAGCTCAGCCGTTACGTCGCGTCGCTTTGACGTCTTTGT GACCCGTGGACAGTTCATCGATATCGGCGTTCCCGAGGATTTCTATCGCGCTCAGGACGAACTGTCTGGGATTTGCAAAGAACAGAAACTG ATTAGCGAAGAGGACCTGTAACCTAGTCTGCAG</p>
wcbK	<p>GAATTCTCTAGAGAAGGAGATATATCCATGGCGAAACGCGTCTGATCACGGGTATTACAGGCATGGTAGGCTCACATCTCGCCGATTTCT GTTGGAAAACACCGATTGGGAGATCTATGGACTGTGTCGTTGGCGCTCTCCACTTGATAACGTGTCTCACTTATTGCCGCGCATCAATGAGA AGAATCGCATTTCGGCTGGTTTATGGGGATCTCCGCGACTATCTGAGTATCCACGAAGCGGTCAAGCAAAGTACCCCGGATTTCTGTGTTTAC CTGGCGGCTCAGAGCTATCCGAAAACAGCTTCGATTCGCCATTAGATACTCTGGAAACGAACGTCCAGGGCACAGCCAATGTTCTGGAAG CTCTGCGCAAGAACAACATTGACGCGGTTACCCATGTATGCGCTTCGTCTGAGGTCTTTGGTCGCGTACCTCGTGAAAAGCTGCCCATTA CGAAGAGTGCACCTTTCATCCGGCATCGCCTTATGCGATTAGCAAAGTGGGTACTGATCTGATTGGCCGCTATTACGCCGAAGCGTACAATA TGACCGTGATGACGACTCGGATGTTCAACCATACTGGTCCACGTCGTGGGGACGTTTTCGCCGAAAGTACCTTTGCGAAACAGATTGCCAT GATCGAACGTGGGTTAATCCCGCCTGTGGTAAAACCGGCAATCTGGATTCACTGCGCACTTTTGCAGATGTTGCGGATGCGGTTGCGCGG TACTACATGCTGGTGACCATTAACCCGATTCCGGGTGCTTACTACAACATTGGCGGCACATATTCCTGTACGGTCGGACAGATGTTGGACAC CCTTATCAGCATGAGCACGAGCAAAGACGTTATTCGCGTGGAAACGGATCCGGAACGTTTACGCCGATTGATGCAGACTTGCAAGTGCCC AATACCCGTAAATTTGAAGCCGTAACGGGTTGGAAACCGGAAATCTCCTTTGAGAAAACGATGGAAGATCTGCTCAACTATTGGCGTGCACG TATTTACAGCAGGCGAGAAATTCCTGACACGTGAACAGAACTTATCTCCGAAGAGGACCTGTAACCTAGTCTGCAG</p>
wcbJ	<p>GAATTCTCTAGAGAAGGAGATATATCCATGAAGGTGTTCTGGTCGGGAGTACCGGCTATATTGGCAAACCCCTGTTTGACGCCTGTTACAG CCGTTGGCGTACCCTTGGCACGAGCACACGCGATGGTGCGGATATTGTGTTACAGTTGGCTCGCGCGGAGGCGTTTTCCCTACGAACAAGT CTCAGCGGGCGATGTGGTTGCAGTGGCAGCAGCCATTTCTAGTCCGGATGCCTGCGCTAAAGACTACGAAACGGCCTTTCAGGTTAACGTC ACTGGTACCCTGACTCTGATTTCGCGGGGTAGTTGCGCGTGGTGCGCGCGTATCTTCTTTTCGTCCGATACCGTTTATGGCGCGTCAGAAC AACTGCTGTGCGAAGAAGCGGAACTGACACCAGCGGGTCTTATGGAGCCATGAAACGTCGGGTTGAGGCAGAATTGGGCGAAAATGCAG CAGTCAAGGTCATTGCGCTCAGTTACGTGTTTTCTCTCCGTGATCGTTTCACGCAGTACTTACTTGGTTGCGCGAAAAGAAGGCAAACGCGCC GACATCTTTAAGCCGTTTAGCCGCTGTGTGGTCTATCTGTCCGATGTGGTTGAAGGGGTGGTAAGCCTGATTGAACGCTGGGACGCCATTG ACGAGCGTGTGATTAACCTTTGTAGGCCCGGAACTGGTAGCACGCGAAGATTTTCGTGGAGAAAATTGCGAACTTAGCCGCTCCAGAGCTCGA TTATGGCTTTTCTGAACCGGAAAGGTGACTTCTTCGTTAATCGCCCGCGTATCATCAACGTTAGCTCCGCTCGCTTTGAGAACTTCTGGGTC GTCGTCCTCGCTCGTTAGGAGAAGCGATCAATCTGGAACCTGAGCAATGAGCAGAACTGATCTCGGAAGAGGATTTGTAACCTAGTCTGCA G</p>

wcbI	<p>GAATTCTCTAGAGAAGGAGATATATCCATGATGGCGCAACGCAAGAAATACAGCGTCTACGGCTCGTGTCAAGCTCCTGCACTTGCCAAAAT GCTGAACTCTTGCCCTACCTTTGCCCGTGATTGGGAACTGGTGGAAATGGAACCGTGCTTTGTGGCGAGCGAGGAACAGATTGATCGCCAT CTGGCTGAAACCATTCTAAATTGGACTTATTCTCTACCAACCGTTTTCCGAAGGTTATCGCGGGGAGAAATACAGCTCAGTTTTCTGCG CAATAGTATGCCGCCAGGTGGCAACGCGTTATCTGTCCAGTATATGCACTGGGAAGGCTATCACCCAACCGTCAATAGCCCCTATGGTCTG CCACCGCATCCGGAAGGTTACGTGGACGCATTGATTGCTGGAGCAGTGGTCATGGATGTGGACAAAGAAACCTATTTACGCCATCTGGAGG AAATTGGTGCATCTCTGCGCATTGACATCGATGAAATCGAGAGCTGGTGCCTAGATGAGCTCAAACCTCGCGAAGTTGGCGAAAACGATGG CGGAAAACAGATCGACATCTCGGTGACTGATTTTATTCTGGCCAATTGTCCGGCAGAAACGTCTGTTTTACACCATGAACCACCCAACAGCCG CCCTGATGCGCGAAATTGCGGCACGCTGTATGCTGGCTTTGGGGTACACGTATCCGATATCTCCTTTGATCAGAACCTTGATCCGCTTGAC GTGACGAAAATGAGCCTGTATCCCATTTATCGCGACTGCTTTGACTTCAGTGAACCTGAACCGGATGAATGAGTACCAGGTACTGTACAAGAA GAAAGCCTATGAGCCGTATCTCCTGGAACAGTTTGAATGGTTTGAAGCCTTACCGAAAGCGGATGTGAGTGCCTTCTCGATCGTGTGCA GCGAATCGTCTGGTTCGACACGGCCTTACGTCTGCGTTTGAATCGTATCCGTATGATGTACCGGATTATGCGTAAACTAGTCTGCAG</p>
wcbB¹	<p>ATGTCCATCATCTACCTGGACGTCACACGGCTGGTTACTCGTCTGTATCAGGGCCTGTTACCGACAGGTGTAGATCGTGTGGGACTCCAGT ACATCCGCCACTATGGTTCACGTGCACGCGCCGTTCTGTGAGAACGTGGCTTCTTTGCTATCCTGAGCGAACGTGATTCCGGCCTTGGTCTTT GCGTGGCTTACCAGCAGCATTGGCAACAAGAATGCCATTTTCCGCCCTTCCGCGCATGCGTGTCTCCGCTCGATCTTCAACACCAGCTTTC AGAACGGCATTCTGTTGCACACGAGCCATTCTGGAATGGAGTTCCCACGCTACTACAAGAAACTGGCCTCCCTTGGCATCAAAAGTGTGTTT CTGATTCACGACTTAATTCCGCTGACGCATGCCGAGTACGCGTCTGGTGTGGAACATAACCATCGTCGCCGCATTACATACCGCATTGG GTTATGCGTCTGGTCTGATTGCGAACAGTCGCTCTACGCTGGAGAGTCTTCCCGCAGAAGCGACTCGTGCGGCCCTGCCGTTACCGCCTT GCGCTATTGCGCATCTGGCAAGCGGCGTTGAACCCCAACCGCCTCGCCAACGCTTGGTGGATGCGCCATACTTCGTGATGCTGGGCACCA TTGAGCCGCGCAAAAACCACTGGTTTATCTTACACGTCTGGCGTGCCTGATCGAACAGCTCGGCAATGCCGCCCGAAACTGGTGGTAAT CGGACGCCGTGGGTGGGAGTGCAGAAATGTGATCGATATGCTGGAACGGTGTGCATCATTGCCCGGTACCGTGATTGAAGAAGCGAATTG CTCGGATGAACGCTTACACGCTTGGCTGCAACATGCTCGCGCACTGCTGTTTCCGTGCTTTGTTGAAGGGTATGGTATGCCACTGGTGGAA GCTCTCGGTCTGGGCGTACCCGTCTTAGCGTCCGATCTTGACGTGTTTCCGTGAAGTTGCGGCAGATATTCCGGATTATCTCGATCCACTGG ATGGGCCGGCTTGGGCCGACGCATTCTGACTATGCACGTGATGACAGTCATGAACGGACTGCGCAGTTAGCGCGGATTGAGCATTTC GCGAACCGACCTGGGTGACCACTTTGAACGCGTTGATGCGTTCTTGGACACGCTGCGC</p>
wcbE¹	<p>ATGAAGAACAAGTACGCGTCTGGTTCGAACTGCGTCCGGCGCTGGATGGCTATTCCGGTATTCCGCAAGAAACCCGTCTGCTGTTCCGCA CTCTGCTTGGGCTGAAAAAATGGCCGTTACAGGGCTTCTCCAGCAAGGTGGGCGTGACATTACGGCCAAAATGGTCTATGCGGCCGAAAG CGAACGCCAGCAGCTCAGATCCGCGCGATGAGCAAGATCGTAGTGAGTTTTACGGACGATGGTCGTGCGGGTGTGGCCGGTGCGATTTT GCGTGCGATCAGCATCGTCATGCGTGGCATTAGCCTCCGCGCAAGCGTGATTGCGGGCTTGTGAGAGAACTTGGCTTCTTTGCTCCGGTT GAATTTGAGGATTTCTGCTGCGCGCTTTCTTTGCCAAAACCTGCCAGTGGATGACCGTCCGATGTGCTGGCCGCGCAGTATCGCGTGC TGCCCTCTGGCAATGGCTACCTTGCATCGCCTGAAGATCTCACTGAGCGGCTTATTACGTGTAACCGCTTCAAGAAAATTGAAACCCGCGAT TTGATGTGTTTATCAGTACAGACACCGTTTCCGGGTGTGCTGTGCTCCAACACCAAACCTCATTGTTGCTATCACGACGCCATTCTATCTTC CTGCCCCATACGATCAAAGACCGCGCTTTTACCACGGTGCACATTATGATGCATTAGCGCGCAATGTCCGGTCTGGCGCGTATTTTGCCT GTGTTTCTGCGGCTACTCGGAACGATCTGCTGAAAATCTTCCGGAAATCGAGGATCGTGCGGTGGTGAATCACAAACCGGTGTACCGAT CTACAGCGAAATCGATTCACCTAAGAAGTGGTGAAGACATTGTGAGCAGCCGTTTGTCCGCCATTGGCGAAGAAATACAAAGCATTGT CCTCGTGCATGTAGACGTCTTACGCGCTCTCGTCTGAATGCCGATCAGCAGGATTTGCAATATCTGCTGATGGTTTCTACCGTCAACCC CGCAAGAACCACACCTCATTGGTTGCCGCTGGGAACTTCTGCGCAGTCGCCACAATGAAGCATTAAACTGGTTCTCGTGGGCGACCTCG GTTGGGATTACGCTGAGTTTGTGGACGCGATTGTTCCGTGGGTGCAACGGGGAGAGCTGTTTCTGCTGAGCAATGTTGCAGCTCCTGATT</p>

	ACGCGATCTTTACAAACATGCCAAAGTGACAATCTGCCCATCCTATGCAGAGGGCTTTGACTATTCTGGGGTGGAAAGCGATGATGTCCGGA GGCGTAGTTGCAGCAAGCGATATTCCGGTCCATCGTGAAGTGTACGAAGATGCTTGCATTTACTTTAACCCGTACAGCGTTGAATCGATTGT AGGTGGCATTCTGTGAGGCGATTAGTTTACGCAATGCGTCCGAAACCTTTGCCGAGTATCGTCGTTGTGCCAAACGCATTAGTGCCCGCTATA CTGCGGAAGCGATTTCCGCCACGGTGGGATGAGTTCATTCAATCTGTGGGAAGTGCTCCC
<i>wcbH¹</i>	ATGAAACTTGCTTTCTTCACCCCGTTTACCCGGCATGTGGCCCGGGCTTCATGAGCAAAGCGGTTGTCGAACATCTCAAAGACTATTTTCGA TGTGACCGTGTCTACCAGTGGCATGCGCTTCATCAGCCGTATGAAGTTGAAGATGTGTCCGTGCGGATCATCGAAGATGACATGGATATG GAGCCCGTTGCCCGCGAGTTCGATGCCGTTATCTACAATCTGGGCAACAATGAGGAAAACCACTACAGCATTTTCCAGTGCCTGAAACGTCT GCCGGGTATCGTGGTGTTCACGATTTTGTATGCAGCACGGCATTGTTAACGACCTGTTAACCGCAAACAGAAAAGCGATCTGTATTACT GGCTGCTGGCGGCCCTCTATGGCAAAGACGGTGTTCGGGCCTGCCATGCCTCTGCTCTGAGTTATGCCGGCGGAGTACGTGGTGGCTGG GATTCGAGCAGTGTTCATCCTTTCCGATGTTTGAGGTATTTAGCGGTTAGCCTCCGCCTGTGTTGTCCATAGCAAATTTGAAAATCAG CTGAAACCTCATTTTGCAGGGCCGATTCTGATGACGCGTAATCCATACGACCTTAAACGCGTGCCAACAGAAGATGTGATTGAAACGTTCTA TAAGGTACGCGGGAAAGCCGGCGGGAAAGTGGTATTCGCGGCGTTTGGTACATGAGCCCTAGCAAAGTGCATTGATCGCGTTCTCCGCGT ATTTGGGCACAGCGAACGCCTGCGCAAAGAAGCGCGCTTAATCCTTTGCGGCGGTGCTTCTCCGGAATACGACCAGTATTTACGCCGCTTA TGTGCGGATGGAGATTTAGGTCAATGCGTCAATTTGCGGCGAGTGTCTCTGATACCGAACTGTATGCGCTCCAAGTGGAAAGCTGATGTGT CGTCAATCTGCGTCAACCGAATACGGAAGGCCAATCCGGTCTTTGATTGAACAGCTCGCGGCGGGCAAACCCGTCATTTGCTACGATAAC GGCTGTTACGGTGACCTGTGAGAGGATGTGGCCTATAAAGTGCACCTGACCACCGACTTTAACGAGTTGACGCGCGTATTTGAACGCCTGC TGAATGGCGTGGCAGAGCGTCTGAAGTTGGACTGCGTGGGCGTGCCGCAGCGTTGGAGTACGACTGTGCTCGGTATGCGCGTGATATGT TTGAGTTTGTGTTTAAAACCGCGAATACCTGAACGCGTTTCAACGCCGCAACTTCGATGGCCTTCGCCACTGGGTGTTTCAGGTACGACA ATTGATACTCGTCATCCGGAGAATTCGCGCACGCAAAGCATTTGGGGCGTATTAGACGGTTTTGTGCATCACGCATCCAGTGAGGCGC TGTCTTGATTCCGGGAGATGCACGTCACTATCTCTCGTTCCTGACTCAACCCTACCGTCCACTGGTTCGATTTCGCTGCTGGAACGTCA GCGTCGGACGTCCCTGAAGTCGAACTGGAAGCAGATGGTACATCTCTGTACAGCACCGCAAAGTGCGCCTGTTCTGAATACCTCGTTCTGGA CCAAGCTGAAATGTCCGGTGAGTGAACATTTGCGGTATATCGGGTATAAGAACTGTTGTGCCGTGATCCGGACTTGGCTGGTCTGCGTGA CTATGCAGAGAAGTTGGCCTCAGAAACGATCTCTCGTAAGGAGATGATTTTGGCGATGTTACGGTCAACTGAATTTGCCGAACGTTTTGGCG ATATCCGCAAATCCAGCGACTATGTCAGTTTAAATTCGCTGGGCTGGAGCACAGGTGGCT
<i>cmk^m</i>	ATGACGGCAATTGCCCGGTTATTACCATTGATGGCCCAAGCGGTGCAGGGAAAGGCACCTTGTGTAAGGCTATGGCGGAAGCGTTGCAAT GGCATCTGCTGGACTCGGGTGCAATTTATCGCGTACTGGCATTGGCGGCATTACATCACCATGTTGATGTTGCGTCGGAAGATGCGCTGGT ACCGCTGGCATCCCATCTGGATGTACGTTTTGTGTCGACCAATGGCAATCTGGAAGTGATCCTCGAAGGGGAAGATGTCAGCGGCGAAATT CGTACTCAGGAAGTGGCGAATGCAGCTTACAAGTCGCGGCATTCCACGCGTTCGTGAAGCATTATTGCGTCGCCAACGCGGTTTTCGCG AATTACCAGGTCTGATTGCCGATGGCCGCGACATGGGAACGGTGGTATTCCCTGATGCACCAGTGAAAATTTTCTTGACGCCTCCTCGGA AGAACGTGCGCATCGCCGCATGCTACAGTTGCAGGAGAAGGGCTTTAGTGTTAACTTTGAGCGCCTTTTGGCCGAGATCAAAGAACGCGAC GACCGCGATCGTAACCGAGCGGTAGCGCCACTGGTTCGGCGAGCCGATGCTTTAGTGTTGGATTCCACCACCTTAAAGCATTGAGCAAGTGA TTGAAAAAGCGCTACAATACGCGCGCCAGAAATTGGCTCTCGCATAA
<i>gmk^m</i>	ATGGCTCAAGGCACGCTTTATATTGTTTCTGCCCCAGTGGCGCGGGTAAATCCAGCCTGATTGAGGCTTTATTA AAAACCCAACCGTTGTA TGACACCCAGGTTTCTGTTTACACACCACACGCCAACC GCGTCCTGGTGAAGTCCACGGTGAACATTATTTCTTTGTTAATCATGATGAATT

	TAAAGAAATGATTAGCAGAGATGCGTTCCTCGAACACGCAGAAGTTTTGGTAATTACTATGGCACTTCGCGTGAGGCCATTGAGCAAGTAC TGGCGACCGGTGTCGATGTTTTCTCGATATCGACTGGCAGGGCGCGCAGCAAATTCGCCAGAAGATGCCGCACGCGCGGAGTATCTTTAT TTTACCGCCGTCCAAAATTGAACTGGACCGCGTCTACGCGGTGCGGTCAGGACAGCGAAGAGGTCATTGCAAAGCGTATGGCGCAAGC TGTTGCAGAAATGAGCCATTACGCCGAATATGATTATCTGATTGTGAATGATGACTTCGATACCGCGTTGACCGATTTGAAGACCATTATTCG CGCCGAACGTCTGCGCATGAGCCGCCAAAAGCAGCGTCATGACGCTTTAATCAGCAAATTGTTGGCAGACTGA
cmk	GAATTCTCTAGATAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCTCTAGGAATAATTTGTTAACTTTAAGAAGGAGAT ATACATATGCACCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCATGCGTGGTATCGTGACCATT GATGGCCCGTCAGCCAGCGGGAAATCCTCGGTTGCACGCCGTGTAGCGGCAGCACTGGGAGTCCCCTATCTGAGCAGTGGCCTCCTCTAT CGCGCTGCGGCGTTTCTGGCCTTACGCGCCGGTGTGATCCGGGTGACGAAGAGGGCTTGTGGCGTTACTGGAGGGTCTGGGAGTCCG CTTGCTGGCGCAAGCCGAAGGCAACCGCGTTCTGGCGGATGGTGAAGGATCTGACCTCGTTCCTGCATACGCCGGAGGTTGATCGGGTCG TAGCGCTGTGGCACGCTTACCGGGCGTACGTGCCTGGGTGAATCGGCGCCTTAAGAAGTGCCTCCACCCTTTGTCGCAGAAGGTCGCGA CATGGGCACTGCCGTGTTCCGGAAGCGGCTCACAAGTTCTACCTCACGGCTAGTCCGGAAGTTCGCGCGTGGCGTGTGCACGCGAAGC TCCTCAGGCCTACGAGGAAGTACTGCGTGATCTTCTGCGCCGTGATGAACGCGACAAAGCGCAGTCTGCACCAGCTCCGGATGCGCTGGT GCTGGACACAGGCGGGATGACCCTGGACGAAGTGGTGGCTGGGTCCTTGCGCATATTCGCCGTTAATAAACTAGTCTGCAG
gmk	GAATTCTCTAGATAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCTCTAGGAATAATTTGTTAACTTTAAGAAGGAGAT ATACATATGCACCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCATGAAAGGCATGCGTGGTCCG CCTGTTTGTATGACCGGGGCAAGTGGTGTGGCAAAGGCACGGTTCGCGCAAGTTTTGGAGCGTACTCGCCTGTTTTACTCCATTAGC ATGACCACTCGTCTCCACGTCCGGGTGAAGTCGATGGCGTCACTACTTCTGGATCGCCCTACGTTTGAGGCCTTAGTTCGGGAAG ATGGCTTTCTCGAATATGCTGAGTATGTGGGTGATCTGTACGGTACCCACGCGCTCCGGTAGAGCGGGCACTGTACGTGGAGAGGATGT GTTGCTTGAATCGAAGTGCAAGGGGCGTTACAGGTCAAACGTGCGGTACCGGAAGCGGTTCTGATCTTCTGTTACCGCCGTCGTTAGC GAACTGAAACGTCGCTTAGTGTATCGCGGCAAAGATTCGCCGAAAAGATCCAGAAACGGCTGGAACAGGCAGAATGGGAGATTGCAATG CCCCTTGTTCGACTATGTCGTGGTGAACGACGTAATCGAAGAAGCCGTTGCCGACTTTCTGGCGATTCTCACCGCAGAACGCCGCCGTTT TGGACGCATGGGTGAAGCGCTGGAATGGCGCTTCGTCGCGATCTGGCTCTGGAAGCGGAACTGGATGAGATTCTGCGCCGCCGTTATGG CGGGACAGGCCATTAATAAACTAGTCTGCAG
tmk	GAATTCTCTAGATAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCTCTAGGAATAATTTGTTAACTTTAAGAAGGAGAT ATACATATGCACCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTCCAATCCATGCCAGGCCTGTTTCTGACC CTGGAAGGACTGGACGGTTCGGGCAAACCACCCAAGCTCGCCGTTTGGCGGCGTTTCTGGAAGCGCAAGGTCGGCCTGTTTTGCTGACT CGCGAGCCAGGTGGTGGTCTGCCGGAAGTTCGCAGTCTCTGCTTACCCAGGAGCTCTCTCCGGAAGCGGAGTACCTCCTGTTAGCGCG GATCGTGCAGAGCACGTTCCGAAGGTGATCTTCCGGGCTTAGCTGCCGGGAAAGTGGTCATCTCCGACCGCTACCTGGACAGCTCACTG GCCTATCAGGGCTATGGACGCGGCTTACCGCTGCCGTGGCTTCGGGAAGTAGCGCGTGAAGCCACACGCGGGTTGAAACCCCGCTTAACG TTCCTGCTGGATCTGCCTCCCGAAGCAGCCTTACGTCGTGTACGTCGCTGATCGCCTTGAAGGTCTGGGCCTTGAGTTCTTCCGTCGCG TGCGCGAAGGCTATCTCGCGCTGGCACGTGCCGAACCAGGGCGCTTTGTCGTGCTGGATGCCACGCTGCCGGAAGAAGAGATTGCACGT GCGATTCAAGGCTCATCTGCGCCGTTGTTACCGTAATAAACTAGTCTGCAG

<i>umpk^t</i>	GAATTCTCTAGATAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGGAATAATTTTGTTTAACTTTAAGAAGGAGAT ATACATATGCACCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTA CTCTCCAATCCATGAAATACAAACGGGTCCTG CTGAAACTGAGCGGAGAGTTTCTGACGCGGAATGGCTTTGGCATCGAACC GGAAGCTACACAGGCACTTGCCCGCGAAATCAAAGCAGCG TATGATACCGGCGTT CAGCTGGCGATTGTCATTGGAGCGGGTAACTTATGGCGCGGTGCACGTCAAGGCGTTGGGATGGATCGCGCTACA GCGGACTACATTGGCATGCTTGCCACCATTATGAACGCCTTGGCGCTCCAGGATGCGCTGGAGAGTCTCGGCGTACCAACCCGTGTGCAG ACTGCACTGACCATTACCCAGGTTGCGGAACCGTACATTCGCCGTCGTGCCTTGCGCCATCTGGAGAAAGAACGCATCGTGATCTTTGGTG GTGGA ACTGGCAATCCGTTCTTTTCCACGGATACTGCTGCCGCATTACGTGCGCTTGAAGTTGGGGCAGAAGTGGTCCTGATGGCCAAGAA CAAAGTTGACGGTGTGTATTCGGATGACCCTCGCAAGAATCCC GAAGCTGTGCGCTTCGATGAACTCACCTATCTGGAGGTA CTGAACCGT GGCCTGCAAGTCATGGATACGACGGCGATTACGCTGTGCATGGAAGCCGGTTTACCGATTGTGGTGT TTGACATCTTCAAACCAGGTGCGT TGGTGGGGATTATCCAAGGCGAGAAAGTAGGTACCCTGATCCACTAATAAACTAGTCTGCAG
--------------------------------	--

Notes

1 = synthesised gene

m = mesophilic

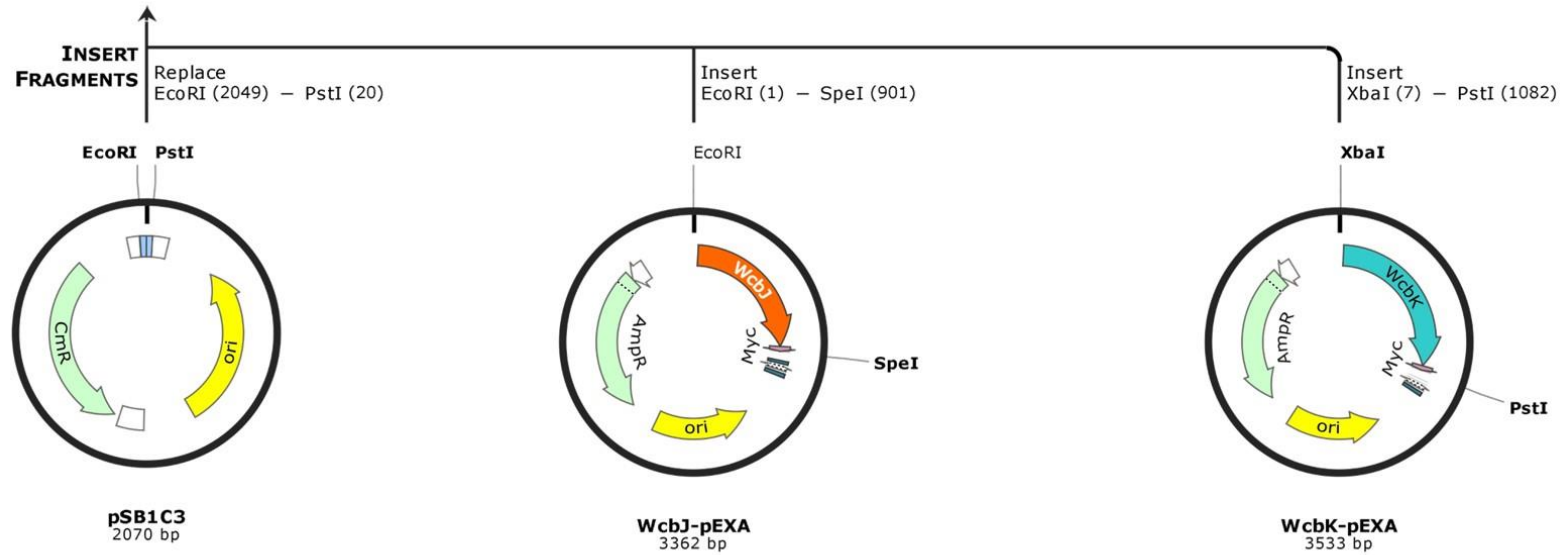
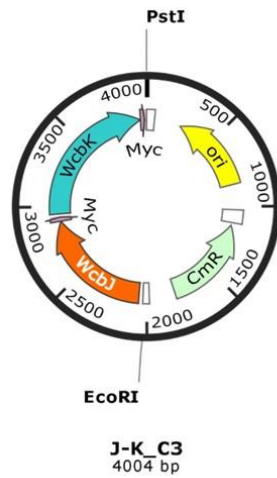
t = thermophilic

Appendix 5 – Purification tags and cloning sequences

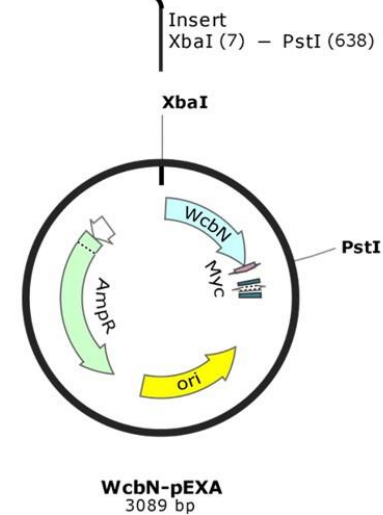
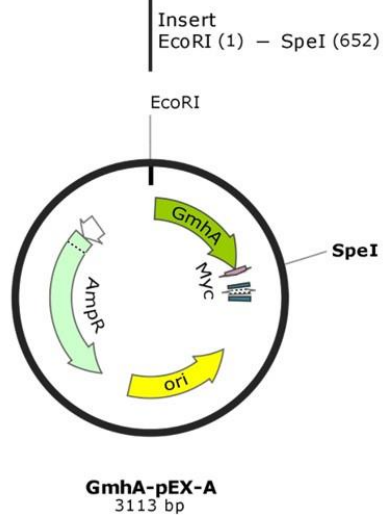
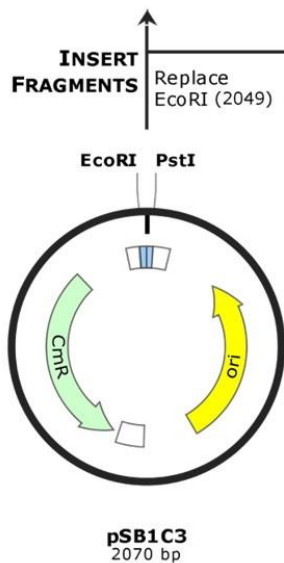
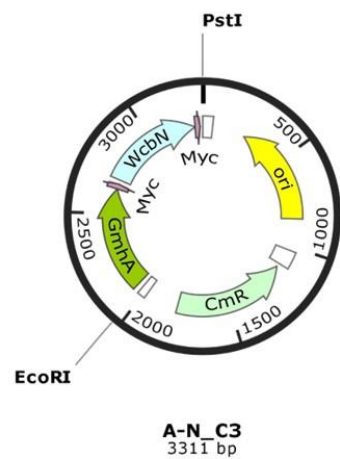
Name	Amino acid sequence	Nucleic acid sequence	Function
His-tag	HHHHHH	-	For protein purification / antibody detection
Myc-tag	EQKLISEEDL	-	For protein purification / antibody detection
FLAG-tag	DYKDDDDK	-	For protein purification / antibody detection
HA-tag	YPYDVPDYA	-	For protein purification / antibody detection
BioBricks™ 5' sites	-	GAATTC TCTAGA	<i>EcoRI</i> and <i>XbaI</i> restriction sites for BioBricks 3A assembly
BioBricks™ 3' sites	-	ACTAGT CTGCAG	<i>SpeI</i> and <i>PstI</i> restriction sites for BioBricks 3A assembly
Stop	-	TAA	stop codon
Spacer	-	ATATATCC	spacer between RBS and start codon
pBAT4 RBS	-	GAAGGAG	ribosome binding site

Appendix 6 – BioBricks™ 3A assembly of GDP-6dbHep biosynthesis plasmid

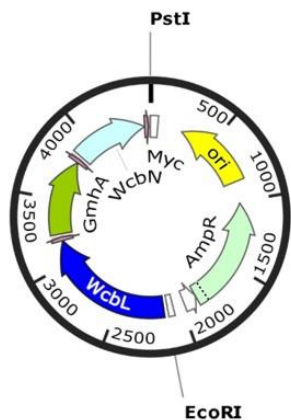
Created with SnapGene®



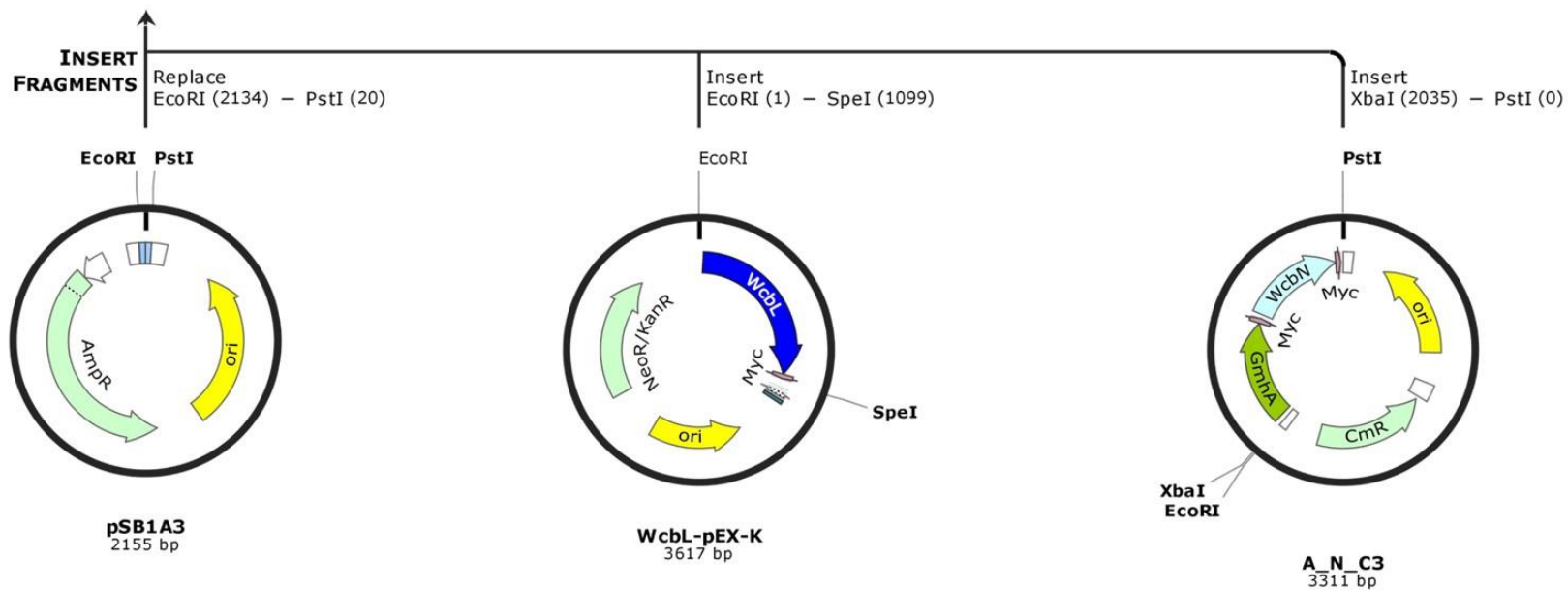
Created with SnapGene®



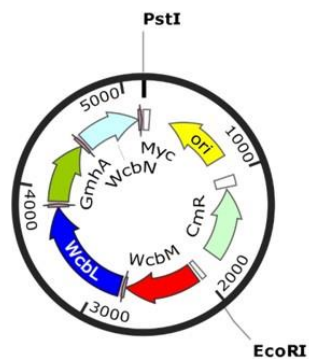
Created with SnapGene®



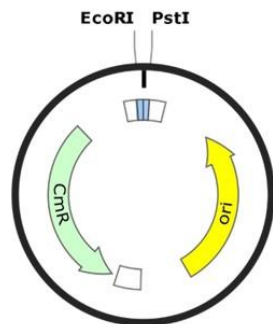
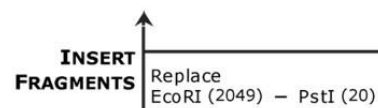
L-A-N_A3
4488 bp



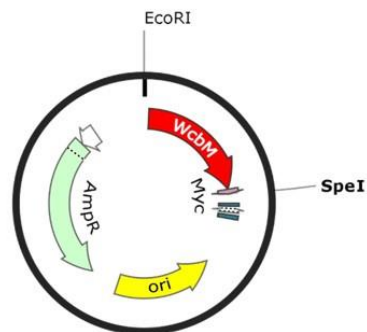
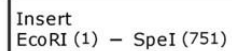
Created with SnapGene®



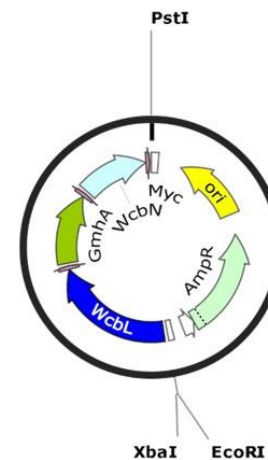
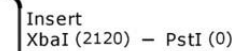
M-L-A-N_C3
5147 bp



pSB1C3
2070 bp

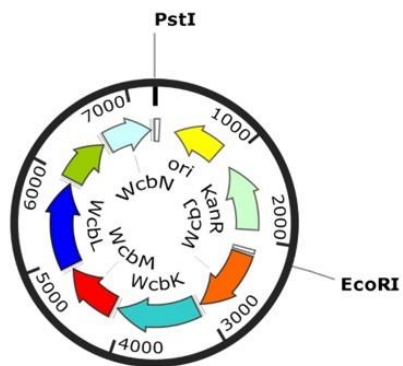


WcbM-pEXA
3212 bp

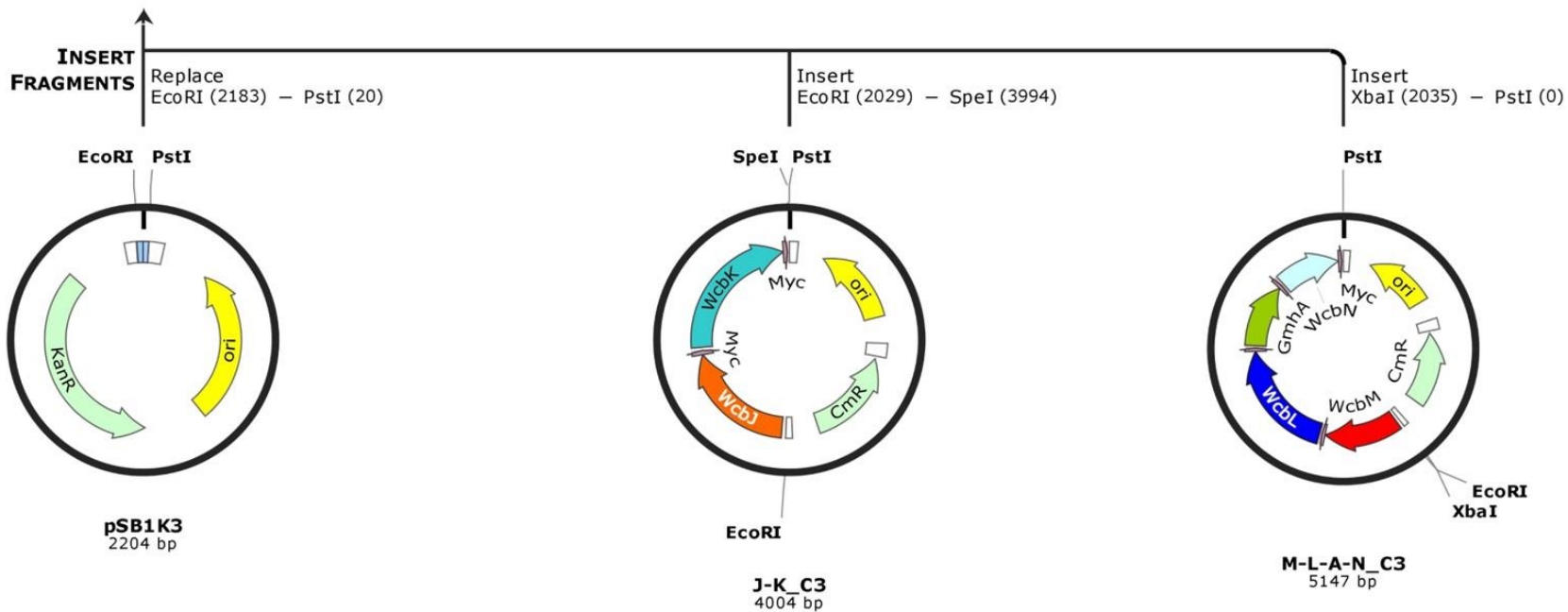


L-A-N_A3
4488 bp

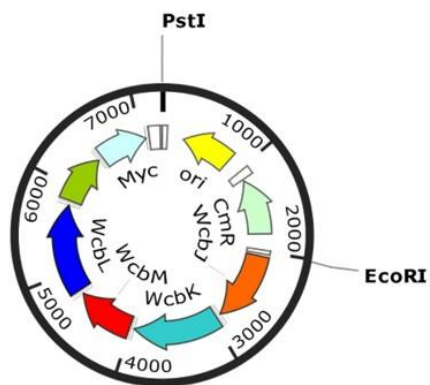
Created with SnapGene®



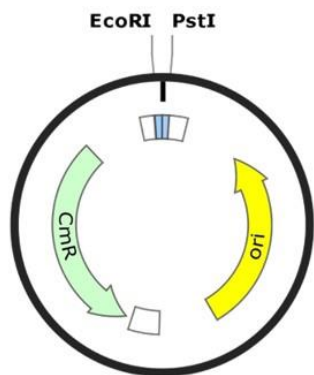
J-K-M-L-A-N_K3
7240 bp



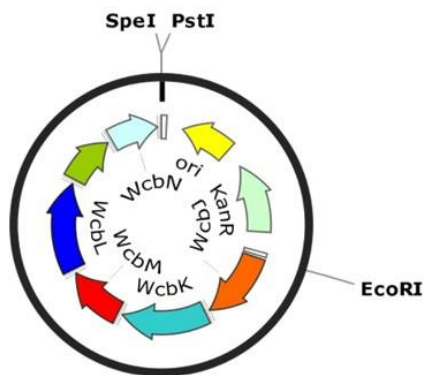
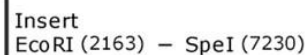
Created with SnapGene®



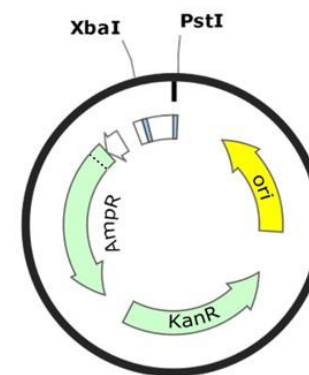
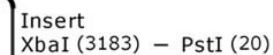
JKMLAN-tt_C3
7251 bp



pSB1C3
2070 bp

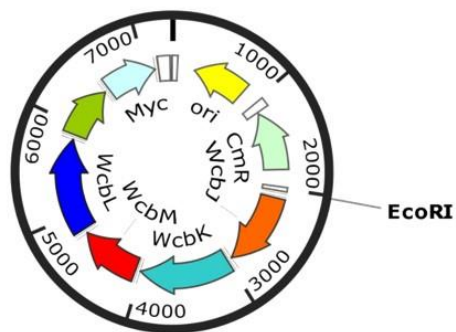


J-K-M-L-A-N_K3
7240 bp

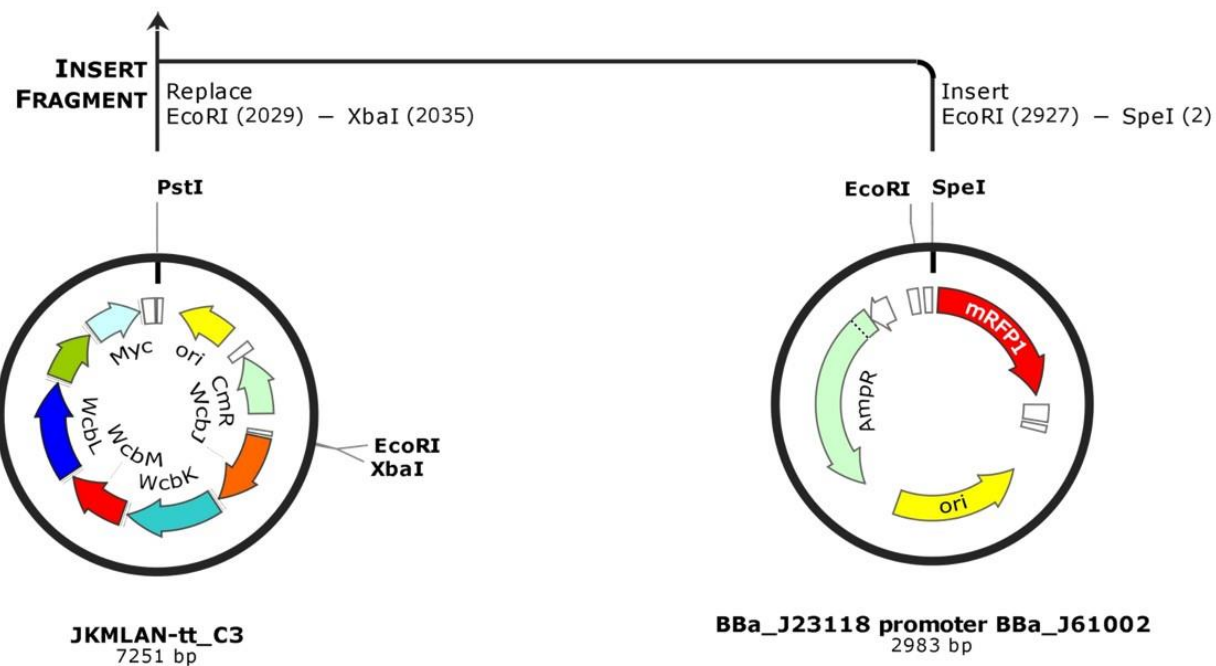


BBa_B0015 terminator pSB1AK3
3318 bp

Created with SnapGene®



p-JKMLAN-tt
7303 bp



JKMLAN-tt_C3
7251 bp

BBa_J23118 promoter BBa_J61002
2983 bp

