

Glycan Utilisation and Resource Allocation by Prominent Members of the Human Gut Microbiota

A thesis submitted for the degree of Doctor of Philosophy

2012-2016

Jonathon Alexander Briggs

Institute for Cell and Molecular Biosciences

Faculty of Medical Sciences

Newcastle University

Abstract

The human gut microbiota (HGM) represents a diverse community of bacteria with cell numbers in the trillions. The HGM is presented with dietary glycans, primarily from plant material consumed in the human diet, and host glycans which allow the bacterial community to proliferate in the gut. As such carbohydrate active enzymes (CAZymes) and multi-protein glycan utilisation systems are commonly expressed by members of the HGM. The human genome encodes few CAZymes rendering the vast majority of carbohydrates consumed accessible to the HGM. During bacterial fermentation of glycans short chain fatty acids (SCFAs), acetate, butyrate and propionate, are released into the gut lumen and are utilised for energy by colonocytes. SCFAs have also been shown to repress genes relating to proliferation in cancerous gut cells and activation of Treg cells.

Bacteroidetes have been shown to possess a unique glycan utilisation strategy involving degradation of the target glycan at the cell surface before transport into the periplasm where degradation is completed before fermentable substrates are transported into the cytoplasm. *Bacteroides ovatus* deploys a xylan utilisation system expressed from two loci which is capable of targeting the relatively simple glucuronoxylans (GX) and arabinoxylans (AXs), but also the highly complex glucuronoarabinoxylans (GAXs) from corn. Work presented in this thesis has contributed to our understanding of xylan degradation. The data showed that xylan degradation begins at the cell surface with endo-acting xylanases from family GH10, targeting the backbone of simple xylans, while a GH98 xylanase targets specific structures in complex GAXs generating oligosaccharides that are transported into the periplasm. Co-culturing *B. ovatus* alongside *Bifidobacterium adolescentis* on GX and AX demonstrated cross-feeding of xylooligosaccharides, allowing *Bi. adolescentis* access to previously inaccessible substrate. This interaction provides further evidence for use of xylans and xylooligosaccharides as prebiotic supplements to the human diet with the potential to enrich for both butyrate-producing *Bifidobacterium* and propionate producing *Bacteroides*.

Bacteroides thetaiotaomicron is regarded as a glycan generalist with upwards of 88 polysaccharide utilisation loci (PULs), the majority of which target different glycans in the gut. Among these target glycans are pectic polysaccharides, homogalacturonan, galactan, arabinan, rhamnogalacturonan I (RGI) and rhamnogalacturonan II (RGII). Galactan, which represents a relatively simple polysaccharide, required only two glycoside hydrolases for complete degradation, a surface galactanase (GH53) and a periplasmic galactosidase (GH2). Arabinan required two surface endo-arabinanases (GH43), a periplasmic α -1,2-arabinofuranosidase (GH43) and α -1,3-arabinosidase (GH51) for the removal of sidechains and a final periplasmic α -1,5-arabinosidase (GH51) to complete degradation of the arabinan backbone. Remaining galactooligosaccharide sidechains appended to rhamnose of RGI were degraded by three galactosidases (two GH2s and a GH35), each displaying different substrate preferences. During *B. thetaiotaomicron* pectin utilisation a degree of surface degradation was shown to be required prior to import; each of the resulting oligosaccharides can be scavenged by a *B. thetaiotaomicron* mutant lacking the specific surface enzymes required for polysaccharide degradation. These data demonstrate cross-feeding is prevalent during *B. thetaiotaomicron* utilisation of pectic polysaccharides in the HGM.

Fructans are particularly well described prebiotics found to be a bifidogenic growth substrate. *B. ovatus* and *B. thetaiotaomicron* are capable of utilising inulin- and levan-type fructans, respectively. Co-culture of *B. ovatus* and *Bi. adolescentis* on inulin demonstrates that the *Bifidobacterium sp* out competes the *Bacteroides sp*, despite *B. ovatus* being capable of utilising long and short inulin chains. Cross-feeding also occurs when *Bi. longum* or *Bacteroides vulgatus* is the oligosaccharide recipient. *B. thetaiotaomicron* growth on levan is able to support growth of *Bi. adolescentis* indicating levan oligosaccharides have probiotic potential. Cross-feeding has the potential to have a dramatic influence on how probiotics are processed, utilised and distributed in the gut and their contribution to the HGM food web must be considered when designing dietary supplements.

Acknowledgements

I would like to thank my supervisor Prof. Harry Gilbert for giving me the opportunity to study within his lab. His advice, support and patience throughout my studies was invaluable in the completion of my PhD.

I would also like to thank all members, past and present, of the Gilbert and Bolam labs for advice and friendship. This work would not have been possible without the collaboration and invaluable discussions with, in no particular order, Artur Rogowski, Amy Glenwright, Dave Bolam, Carl Morland, Elisabeth Lowe, Fiona Cuskin, Jose Munos, Adam Jackson, Ana Luis, Max Temple, Didier Ndeh, Sarah Shapiro, Alan Cartmell, Xiaoyang Zhang, Aurore Laborel, Lucy Crouch, Justina Briiliute and Immacolata Venditto.

A large portion of this work relied on the ability to generate *Bacteroides* genetic mutations, as such, I would like to thank Dr Eric Martens for providing the plasmids and *Bacteroides* strains required to create the mutants.

The work in Chapter 3 was done in close collaboration with Dr Artur Rogowski. The results in Chapter 4 build on the previous work of Dr Lauren Mckee, Dr Xiaoyang Zhang and Dr Elisabeth Lowe and was done in collaboration with Dr Didier Ndeh, Dr Artur Rogowski and Ana Luis. Work in Chapter 5 was done in collaboration with Dr Sarah Shapiro and Dr Dave Bolam.

Lastly, I would like to thank my family and friends for love and support over the last four years.

Contents

Title Page	1
Abstract	3
Acknowledgements	4
Contents	5
List of Figures	11
List of Tables	15
Abbreviations	17
Chapter 1: General Introduction	19
1.1 Introduction	19
1.2 Glycans available in the Human Gut	21
1.2.1 Host Derived Glycans	22
1.2.2 Dietary Glycans	22
1.2.2.1 Cellulose	23
1.2.2.2 Hemicellulose	25
1.2.2.2.1 Xylan	25
1.2.2.2.2 Xyloglucan	27
1.2.2.2.3 Mannan	28
1.2.2.2.4 β 1,3, β 1,4 mixed linked glucans	29
1.2.2.3 Pectin	29
1.2.2.3.1 Homogalacturonan	29
1.2.2.3.2 Rhamnogalacturonan II	30
1.2.2.3.3 Rhamnogalacturonan I	31
1.2.2.4 Fructans	32
1.2.2.4.1 Inulin	32
1.2.2.4.2 Levan	35
1.3 Human Gut Microbiota	36
1.3.1 Microbiota Composition and Diversity	36
1.3.2 Short Chain Fatty Acid production by the HGM	41
1.3.3 Interactions between members of the HGM	43
1.3.4 Interactions between the HGM and Pathogens	45
1.3.5 Effect of Prebiotics on the HGM	46
1.4 Carbohydrate Active Enzymes	48
1.4.1 Classification of Carbohydrate Active Enzymes (CAZymes)	49
1.4.2 Catalytic Mechanisms	50
1.4.2.1 Retaining Mechanism	50
1.4.2.2 Inverting Mechanism	52
1.4.3 Subsite Nomenclature	54
1.5 Glycan Utilisation Systems	54
1.5.1 <i>Bacteroides</i> Glycan Utilisation	55
1.5.1.1 Extracellular Glycan Degradation	55
1.5.1.2 Surface Glycan Binding Proteins and the SusCD-homologue Complex	56
1.5.1.3 Periplasmic Glycan Degradation	57
1.5.1.4 PUL Regulation	57
1.5.2 <i>Bifidobacterium</i> Glycan Utilisation	60
1.5.2.1 Extracellular Solute Binding Proteins and ABC-Transporter	61
1.5.2.2 <i>Bifidobacterium</i> Glycoside Hydrolases	62
1.5.2.3 Regulation	64

1.5.3 Glycan Utilisation by other Members of the HGM	65
1.6 Objectives of this Study	67
Chapter 2: Materials and Methods	69
2.1 Bacterial Strains and Mutants	69
2.2 Vectors	73
2.3 Bacterial Growth Conditions	74
2.3.1 Growth Media Composition	74
2.3.2 <i>E. coli</i> Growth Conditions	75
2.3.3 <i>Bacteroides</i> Growth Conditions	75
2.3.4 <i>Bifidobacterium</i> Growth Conditions	75
2.3.5 Selective Media	76
2.4 Basic Lab Methods	76
2.4.1 Storage of DNA and Bacteria	76
2.4.2 Sterilisation	76
2.4.3 Centrifugation	77
2.4.4 Plating Bacteria	77
2.4.5 Chemically Competent <i>E. coli</i>	77
2.4.6 Transformation of Chemically Competent <i>E. coli</i>	78
2.4.7 Small Scale Rapid Purification of Plasmid DNA from <i>E. coli</i>	78
2.4.8 Restriction Digest of DNA	78
2.4.9 Measuring DNA Concentration	79
2.4.10 Agarose Gel Electrophoresis	79
2.4.11 Visualisation and Photography of Agarose Gels	80
2.4.12 Determination of DNA Fragment Size	80
2.5 Purification of DNA fragments and cloning into vectors	80
2.5.1 Purification of Vector DNA (Gel extraction)	80
2.5.2 Purification of inserts and PCR products	80
2.5.3 Ligation of Insert and Vector DNA	80
2.5.4 Ligation Independent Cloning	81
2.5.5 Polymerase Chain Reaction	82
2.5.6 Site-Directed Mutagenesis	83
2.5.7 PCR Overlap Extension	84
2.5.8 Automated DNA Sequencing	86
2.6 Protein Expression and Purification	86
2.6.1 Protein Expression	86
2.6.2 Immobilised Metal Affinity Chromatography (IMAC)	87
2.6.3 SDS-PAGE	87
2.6.4 Determination of Protein Concentration	89
2.6.5 Protein Concentration	89
2.7 Bioinformatics	89
2.7.1 Alignments	89
2.7.2 Prediction of Prokaryotic Signal Proteins	89
2.7.3 Identification of Loci in <i>Bacteroides</i> and <i>Bifidobacterium</i>	90
2.8 Biochemistry	90
2.8.1 Enzyme Assays	90
2.8.1.1 3,5-Dinitrosalicylic acid (DNSA) Reducing Sugar Assay	90
2.8.1.2 PNP Substrates	91
2.8.1.3 Galactose/Arabinose Detection Kit	91
2.8.1.4 Acetic acid Detection Kit	92
2.8.2 High Pressure Liquid Chromatography/associated assays	92

2.8.3 Isothermal Titration Calorimetry	93
2.8.4 Thin Layer Chromatography	94
2.8.5 Acid Hydrolysis/Oligosaccharide Purification	94
2.9 Microbiology	94
2.9.1 Culture Preparation and Monitoring	94
2.9.2 <i>Bacteroides</i> and <i>Bifidobacterium</i> Co-culture	95
2.9.3 Genomic DNA Extraction	96
2.9.4 <i>Bacteroides</i> Mutagenesis and Genomic Insertions	96
2.9.5 Whole Cell Assay	99
2.9.6 Real-time Quantative PCR (qPCR)	100

Chapter 3: Xylan Utilisation and Cross-feeding by *Bacteroides ovatus* **102**

3.1 Introduction	102
3.1.1 Background	102
3.1.2 Aims	106
3.2 Results	107
3.2.1 Expression and purification of BACOVA_04390, BACOVA_04387, BACOVA_03432 and BACOVA_03433	107
3.2.2 BACOVA_04387 and BACOVA_04390	109
3.2.2.1 GH10 Product Profiles	109
3.2.2.2 BACOVA_04390 and BACOVA_04387 Activity on BX	110
3.2.2.3 HPAEC analysis of GH10 activity on Xylooligosaccharides	111
3.2.2.4 Activity of BACOVA_04390 and BACOVA_04387 on Xylooligosaccharides	112
3.2.2.5 Product profiles of BACOVA_04390 and BACOVA_04387 during Xylooligosaccharide hydrolysis	114
3.2.3 BACOVA_03432	115
3.2.3.1 BACOVA_03432 Activity	116
3.2.4 BACOVA_03433	116
3.2.4.1 BACOVA_03433 Activity	116
3.2.5 <i>B. ovatus</i> Xylan Cross-feeding	118
3.2.5.1 Supernatant Oligosaccharides	118
3.2.5.2 <i>Bifidobacterium adolescentis</i> (<i>Bi. adolescentis</i>) growth on xylans and Xylooligosaccharides	120
3.2.5.3 <i>B. ovatus</i> – <i>Bi. adolescentis</i> Crossfeeding on xylans	122
3.2.4.4 <i>B. ovatus</i> -GH98 Mutant Crossfeeding	121
3.3 Discussion	124
3.3.1 GH10 Xylanases from the Small Xylan PUL	129
3.3.1.1 Investigating number of subsites and preferred productive binding modes	129
3.3.2 Two surface endo-xylanases demonstrating different specificities from the large Xylan PUL	132
3.3.3 <i>Bi. adolescentis</i> growth on xylans and xylan derived oligosaccharides	133
3.3.4 <i>B. ovatus</i> generates oligosaccharides at the cell surface during growth on Xylans	134
3.4 Conclusion	137

Chapter 4: Pectin Utilisation by *Bacteroides thetaiotaomicron* **139**

4.1 Introduction	139
4.1.1 Background	139
4.1.2 Aims	142
4.2 Results	143
4.2.1 Cloning, Expression and Purification of Pectin Utilisation Proteins	143
4.2.2 Galactan Utilisation by <i>B. thetaiotaomicron</i>	145

4.2.2.1 Identification of the Galactan PUL	145
4.2.2.2 Purification of Galactooligosaccharides	145
4.2.2.3 Activity of BT_4668 on Galactan and Galactooligosaccharides	147
4.2.2.4 Location of BT_4668	156
4.2.2.5 Galactan PUL SusD-homologue, BT_4670, and Surface Glycan Binding Protein, BT_4669	159
4.2.2.6 Galactooligosaccharide Degradation by BT_4667	161
4.2.2.7 Binding of BT_4673 Sensor Domain to Activating Ligand	163
4.2.2.8 BACOVA_05493, an unassigned glycoside hydrolase in the <i>B. ovatus</i> galactan PUL	166
4.2.2.9 Galactan PUL Mutants	169
4.2.3 RGI Galactosidases	171
4.2.3.1 BT_4158	177
4.2.4 Arabinan PUL Characterisation	178
4.2.4.1 Surface Binding Proteins	180
4.2.4.2 Periplasmic GH51s	182
4.2.4.3 Arabinan PUL Mutants	184
4.2.5 Pectin Cross-Feeding	187
4.2.5.1 Supernatant Oligosaccharides	187
4.2.5.2 Mutant growth on oligosaccharides and polysaccharides	189
4.2.5.3 Galactan Cross-Feeding	193
4.2.5.4 Arabinan Cross-Feeding	193
4.2.5.5 Mucilage Cross-Feeding	194
4.2.5.6 RGII Cross-Feeding	194
4.3 Discussion	196
4.3.1 Utilisation of Galactan	196
4.3.1.1 BACOVA_05493	198
4.3.2 Degradation of Galactooligosaccharide Side-Chains on RGI	200
4.3.3 Arabinan Utilisation	202
4.3.4 Cross-Feeding with Pectic Polysaccharides	204
4.3.4.1 Galactooligosaccharide Cross-Feeding	204
4.3.4.2 Arabinan Cross-Feeding	205
4.3.4.3 RGI backbone (Mucilage) Cross-Feeding	206
4.3.4.4 RGII Cross-Feeding	207
4.4 Conclusion	207
4.5 Future Work	208
Chapter 5: Cross-feeding during <i>Bacteroides</i> utilisation of Fructans	210
5.1 Introduction	210
5.1.1 Background	210
5.1.2 Aims	214
5.2 Results	215
5.2.1 Growth on Fructans	215
5.2.1.1 <i>B. ovatus</i> , <i>B. vulagatus</i> , <i>Bi. adolescentis</i> and <i>Bi. longum</i> Growth on Fructans	215
5.2.1.2 Growth Supernatants	217
5.2.1.3 Fructooligosaccharide usage by <i>Bi. adolescentis</i> and <i>Bi. longum</i>	220
5.2.2 Analysis of inulin and levan utilisation loci	223
5.2.3 Cross-Feeding with Fructans	224
5.2.3.1 Inulin Cross-Feeding	224
5.2.3.1.1 <i>B. ovatus</i> Co-culture with <i>Bi. adolescentis</i>	225
5.2.3.1.2 Δ GH91 <i>B. ovatus</i> mutant Cross-Feeding with <i>Bi. adolescentis</i>	227

5.2.3.1.3 <i>B. ovatus</i> Cross-Feeding with <i>Bi. longum</i>	227
5.2.3.1.4 <i>B. ovatus</i> Cross-Feeding with <i>B. vulgatus</i>	228
5.2.3.2 Levan Cross-Feeding	228
5.2.4 Growth on Conditioned Media	228
5.3 Discussion	229
5.3.1 <i>Bacteroides</i> and <i>Bifidobacterium</i> Fructan and FOS Utilisation	229
5.3.2 <i>B. ovatus</i> shows restricted growth in co-culture with <i>Bi. adolescentis</i> on inulin but not with <i>Bi. longum</i> .	231
5.3.3 Levan Cross-feeding	233
5.4 Conclusion	234
5.5 Future Work	234
Chapter 6: Final Discussion	236
References	241
Appendix A	260
A.1 Protein Localisation Based on LipoP 1.0 Analysis	260
A.2 Validation of GH10 Kinetic analysis of XOS Hydrolysis	262
A.3 <i>Bacteroides</i> qPCR Tag Insertion	263
A.4 Recombinant Pectin PUL Protein Expression	264
A.5 Validation of BT_4668 (GH53) kinetic analysis of galactooligosaccharide hydrolysis	268
Appendix B: Primer List	269
Appendix C: Chemicals, Media, Enzymes and Substrate Suppliers	272
C.1 Chemicals	272
C.2 Media	275
C.3 Enzymes	276
C.4 Kits	276

List of Figures

Chapter 1: General Introduction

Figure 1.1 Overview of the composition of the human microbiota.	20
Figure 1.2 Diagram of mucin domain structure.	22
Figure 1.3 Schematic diagram of Plant cell wall structure.	23
Figure 1.4 Diagram of cellulose structure.	24
Figure 1.5 Diagram of xylan structures.	26
Figure 1.6 Diagram of Xyloglucan repeating unit structure.	27
Figure 1.7 Schematic diagram of hemicellulose mannan.	28
Figure 1.8 Diagram of RGII pectin.	30
Figure 1.9 Diagram of RGI pectin.	32
Figure 1.10 Fructan structures.	34
Figure 1.11 An overview of gut microbe metabolism from dietary fibre to effects on the host.	35
Figure 1.12 Diversity of phyla and genera within the HGM of healthy individuals.	39
Figure 1.13 Effects of microbe produced SCFAs on host cells.	42
Figure 1.14 Diagram of glycan cross-feeding interactions between members of the HGM.	44
Figure 1.15 Expansion of glycoside hydrolases (GHs) in gut microbiota.	49
Figure 1.16 Retaining enzymes mechanism.	51
Figure 1.17 Inverting enzymes mechanism.	53
Figure 1.18 Schematic diagram showing subsite nomenclature of glycoside hydrolases.	54
Figure 1.19 Simplified <i>Bacteroides</i> utilisation system.	59
Figure 1.20 Simplified <i>Bifidobacterium</i> glycan utilisation system.	63
Figure 1.21 Model of the <i>R. bromii</i> multi-enzyme amylosome complex.	66
Figure 1.22 Model of <i>E. rectale</i> starch utilisation.	67

Chapter 2: Materials and Methods

Figure 2.1 Sewing PCR was used to Generate Gene Knockout Fragments. Sewing PCR was	85
--	----

used to remove a gene from a fragment of DNA.

Figure 2.2 Generating Knockout Strains of <i>Bacteroides ovatus</i> and <i>Bacteroides thetaiotaomicron</i> .	98
---	----

Chapter 3: Xylan Utilisation and Cross-feeding by *Bacteroides ovatus*

Figure 3.1 Schematic structure of the main classes of xylan used in this chapter.	103
Figure 3.2 Schematic of the <i>B. ovatus</i> PULs.	104
Figure 3.3 SDS-PAGE Protein gels of Xylanase expression and TALON purification elution fraction.	108
Figure 3.4 Modular structure of BACOVA_04387 and BACOVA_04390.	109
Figure 3.5 Activity of BACOVA_04387 and BACOVA_04390.	110
Figure 3.6 HPAEC product profiles and rates of activity of BACOVA_04390 and BACOVA_04387 on xylooligosaccharides.	113
Figure 3.7 Preferential binding modes of oligosaccharides in BACOVA_04387 active site.	115
Figure 3.8 Activity of BACOVA_03432.	116
Figure 3.9 Activity of BACOVA_03433 and BACOVA_03433 inactive mutant on CX.	117
Figure 3.10 TLC showing oligosaccharides released by <i>B. ovatus</i> during growth on xylans.	119
Figure 3.11 TLC of xylan oligosaccharides and resulting <i>Bi. adolescentis</i> growth.	121
Figure 3.12 Co-culture of <i>B. ovatus</i> and <i>Bi. adolescentis</i> on xylose.	122
Figure 3.13 Selection of <i>Bi. adolescentis</i> and <i>B. ovatus</i> during colony counts.	123
Figure 3.14 Growth curves of <i>Bi. adolescentis</i> and <i>B. ovatus</i> in xylan co-cultures.	124
Figure 3.15 <i>B. ovatus</i> mutant generation procedure.	127
Figure 3.16 Analysis of qPCR data using Roche lightcycler96 software.	128
Figure 3.17 Growth of $\Delta GH98$ <i>B. ovatus</i> on CX in mono-culture and in co-culture with wild type <i>B. ovatus</i> .	130

Chapter 4: Pectin Utilisation by *Bacteroides thetaiotaomicron*

Figure 4.1 Pectin structure.	140
Figure 4.2 Pectin PUL organisation.	141
Figure 4.3 Examples of and purification of recombinant proteins involved in pectin degradation.	144

Figure 4.4 Galactooligosaccharide purification.	147
Figure 4.5 Characterisation of BT_4668.	149
Figure 4.6 Generation of an inactive catalytic residue mutant of BT_4668 by site directed mutagenesis.	152
Figure 4.7 Products of oligosaccharide hydrolysis and potential binding modes of Bt4668.	155
Figure 4.8 Whole cell assays using wild type <i>B. thetaiotaomicron</i> and the Δbt_4668 mutant.	156
Figure 4.9 BT_4669 binding to ligands measured by isothermal titration calorimetry (ITC).	158
Figure 4.10 BT_4670 binding to ligands measured by ITC.	161
Figure 4.11 BT_4667 activity on galactosides.	162
Figure 4.12 BT_4670 binding to ligands measured by ITC.	165
Figure 4.13 Discovery of an ORF present in <i>B. ovatus</i> PUL-Gal.	168
Figure 4.14 BACOVA_05493 activity on galactooligosaccharides.	169
Figure 4.15 Growth of galactan PUL mutants on galactan.	170
Figure 4.16 BT_4151, BT_4156, BT_4160 and BT_4181 activity screening TLCs.	172
Figure 4.17 BT_4151 and BT_4160 activity.	173
Figure 4.18 Characterisation of BT_4156.	174
Figure 4.19 Characterisation of BT_4181.	176
Figure 4.20 BT_4158 activity on acetylated substrates.	178
Figure 4.21 BT_0365 binding to ligands measured by ITC.	179
Figure 4.22 BT_0361 and BT_0363 binding to ligands measured by ITC.	181
Figure 4.23 Activity of GH51 enzymes BT_0348 and BT_0368.	183
Figure 4.24 Growth of <i>Bacteroides</i> with mutations in the arabinan PUL.	186
Figure 4.25 TLC of oligosaccharides released by <i>B. thetaiotaomicron</i> during growth on pectins.	187
Figure 4.26 <i>B. thetaiotaomicron</i> mutants lacking surface endo-acting enzymes encoded by pectin PULs grown on oligosaccharides and polysaccharides.	190
Figure 4.27 Galactan cross-feeding experiment.	192
Figure 4.28 Arabinan cross-feeding experiments.	194
Figure 4.29 Mucilage and RGII cross-feeding experiments.	195
Figure 4.30 <i>B. thetaiotaomicron</i> galactan utilisation.	197
Figure 4.31 RGI galactooligosaccharide sidechain breakdown.	200

Figure 4.32 Arabinan utilisation.	202
-----------------------------------	-----

Chapter 5: Cross-feeding during *Bacteroides* utilisation of Fructans

Figure 5.1 Schematic diagram of fructan structures.	211
Figure 5.2 Simplified diagram of glycan utilisation.	212
Figure 5.3 Growth of <i>B. ovatus</i> , <i>B. vulgatus</i> , <i>Bi. adolescentis</i> and <i>Bi. longum</i> on fructans, FOS-I, FOS-L and fructose.	215
Figure 5.4 TLC images of fructan growth supernatants.	218
Figure 5.5 Fructooligosaccharide utilisation by <i>Bi. adolescentis</i> and <i>Bi. longum</i> .	220
Figure 5.6 Schematic representation of predicted fructan utilisation loci of <i>Bi. adolescentis</i> and <i>Bi. longum</i> .	222
Figure 5.7 Inulin cross-feeding with <i>Bi. adolescentis</i> and <i>Bi. longum</i> .	225
Figure 5.8 Supernatant glycan content of inulin co-cultures.	226
Figure 5.9 Inulin Cross-feeding with <i>B. vulgatus</i> .	227
Figure 5.10 Crossfeeding of levan with <i>Bi. adolescentis</i> .	228
Figure 5.11 Growth of <i>B. ovatus</i> on conditioned media.	230

Appendix

Figure A.1 LipoP 1.0 results of sequence analysis of BACOVA_04390 and BACOVA_04387.	261
Figure A.2 Examples of agarose (0.8 %) DNA gel showing conformation of <i>tag1/11</i> insertion into <i>Bacteroides</i> genome.	264
Figure A.3 Galactan PUL protein expression SDS-PAGE.	265
Figure A.4 Arabinan PUL protein expression SDS-PAGE.	266
Figure A.5 RGI PUL protein expression SDS-PAGE.	267

List of Tables

Chapter 1: General Introduction

Table 1.1 Reported probiotic effects of various glycans in human studies.	47
---	----

Chapter 2: Materials and Methods

Table 2.1 Bacterial strains and mutants used in this study.	69
Table 2.2 Vectors used in this study	73
Table 2.3 Composition of bacterial growth media used in this study	74
Table 2.4 Antibiotic stocks used in this study	76
Table 2.5 Buffers for agarose gel electrophoresis (x10 stocks)	79
Table 2.6 Ligation mixture	81
Table 2.7 Ligation independent cloning mixture	82
Table 2.8 PCR reaction mixture	83
Table 2.9 PCR program	83
Table 2.10 Site directed mutagenesis thermocycler program	84
Table 2.11 SDS-PAGE Gel Running buffer	88
Table 2.12 SDS-PAGE Gel preparation	88
Table 2.13 Composition of typical galactose dehydrogenase linked assay	91
Table 2.14 qPCR program and parameters used throughout this project.	101

Chapter 3: Xylan Utilisation and Cross-feeding by *Bacteroides ovatus*

Table 3.1 BACOVA_04390 and BACOVA_04387 activity on BX.	111
Table 3.2 BACOVA_04390 and BACOVA_04387 activity on xylooligosaccharides	112

Chapter 4: Pectin utilisation by *Bacteroides thetaiotaomicron*

Table 4.1 Proteins expressed and use in this chapter.	145
Table 4.2 BT_4668 activity on galactan and galactooligosaccharides	150
Table 4.3 Thermodynamic parameters of the binding of BT_4669, BT_4670 and BT_4673 to ligands determined by ITC.	160

Table 4.4 BT_4667 activity on galactooligosaccharides, RGI and lactose.	163
Table 4.5 BACOVA_05493 activity on galactooligosaccharides, galactan and RGI	166
Table 4.6 BT_4151, BT_4156 and BT_4160 activity.	175
Table 4.7 BT_0365 binding to arabinooligosaccharides and arabinan	182
Table 4.8 GH51 activity on arabinooligosaccharides and arabinan.	184

Appendix

Table A.1 Prediction of protein localisation discussed in this project	262
Table A.2 Rate of substrate depletion of different concentrations of X6 by BACOVA_04390 and BACOVA_04387.	263
Table A.3 Rate of substrate depletion of different concentrations of Gal6 by BT_4668.	268
Table B.1 Primers used in this Project.	269

Abbreviations

$A_{340\text{nm}}$	Absorbance at wavelength of 340 nm
$A_{405\text{nm}}$	Absorbance at wavelength of 405 nm
$A_{600\text{nm}}$	Absorbance at wavelength of 600 nm
Amp	Ampicillin
Ara2	Arabinobiose
Ara3	Arabinotriose
Ara4	Arabinotetraose
Ara5	Arabinopentaose
Ara6	Arabinohexaose
Ara7	Arabinoheptaose
Ara8	Arabinoocataose
AX	Arabinoxylan
AXOS	Arabinoxyloligosaccharides
BiMM	Bifidobacterial Minimal Media
BLAST	Basic Local Alignment Search Tool
BX	Birchwood Xylan
CAZy	Carbohydrate active enzyme database
CAZymes	Carbohydrate active enzyme
CBM	Carbohydrate Binding Module
CE	Carbohydrate Esterase
CL	Cell Lysate
CM	Clostridial Media
d.p	Degree of Polymerisation
CX	Corn Arabinoxylan
DNSA	3,5-Dinitrosalicylic Acid
FOS-I	Inulin-Type fructooligosaccharides
FOS-L	Levan-Type fructooligosaccharides
Gal2	Galactobiose
Gal3	Galactotriose

Gal4	Galactotetraose
Gal5	Galactopentaose
Gal6	Galactohexaose
Gal7	Galactoheptaose
GAX	Glucuronoarabinoxylan
Gent	Gentamycin
GH	Glycoside Hydrolase
GOS	Galactooligosaccharides
GX	Glucuronoxyylan
HG	Homogalacturonan
HGM	Human Gut Microbiota
HPAEC	High Pressure Anion Exchange Chromatography
HTCS	Hybrid Two-Component System
IMAC	Immobilised Metal Affinity Chromatography
IPTG	Isopropyl β -D-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
Kan	Kanamycin
MM	Minimal Media
PL	Polysaccharide Lyase
PUL	Polysaccharide Utilisation Loci
RGI	Rhamnogalacturonan I
RGII	Rhamnogalacturonan II
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
TCS	Two-Component System
Tet	Tetracycline
TLC	Thin Layer Chromatography
TYG	Tryptone-Yeast extract-glucose media
WX	Wheat Arabinoxylan
XOS	Xylooligosaccharides

Chapter 1: General introduction

1.1 Introduction

Carbohydrates are diverse and biologically important molecules that are composed primarily of three elements, carbon, hydrogen and oxygen. The simplest sugar, monosaccharides, exist as a chain of five or more carbon atoms with differing arrangements of hydrogen or hydroxyls on each carbon, along with either an aldehyde or a ketone group. These linear monosaccharides can adopt an open chain or a cyclic five or six membered ring structures, termed furanose and pyranose conformations, respectively. Monosaccharide type is determined by the stereochemistry at each carbon, the presence of additional groups to H and OH, such as N-acetyl, and whether it is a ketone or aldehyde (Berg *et al.*, 2002).

Linked through glycosidic bonds between the anomeric carbon (the aldehyde or ketone group) of one sugar and any carbon of the other sugar, monosaccharides can form long polymers referred to as polysaccharides. Nomenclature of glycosidic bonds is based on its stereochemistry and which carbons in each ring are linked. Lactose, for example is a disaccharide of galactose and glucose with a β 1,4-linkage, indicating the glycosidic bond is between the anomeric carbon, carbon-1 (C-1), of galactose and C-4 of glucose; the linkage is equatorial with the plane of the sugar rings. The β -linkage of a D-sugar allows for linear polysaccharides, whereas the α -linkage of D-sugars usually leads to helical polysaccharides, starch for example. Diversity of polysaccharide structure arises from the infinite permutations of monosaccharides from which carbohydrates are constructed, and the variety of bonds between each monosaccharide. Polysaccharides can be linear or branched, allowing for further diversity in structure (Berg *et al.*, 2002).

In biological systems, glycans are typically used for energy storage and structural roles. Glycogen is a glucose polymer used for energy storage in mammalian and bacterial cells while starch performs a similar role in plant cells. The plant cell wall contains many classes of carbohydrate, each working in concert to provide the cell with strength, and resistance to environmental insults and pathogens.

Glycans are present on mammalian cell surfaces and fulfil roles in signalling and cell recognition. The bacterial cell capsule, expressed by some species of bacteria is made up of glycans to protect the cells in harsh environments.

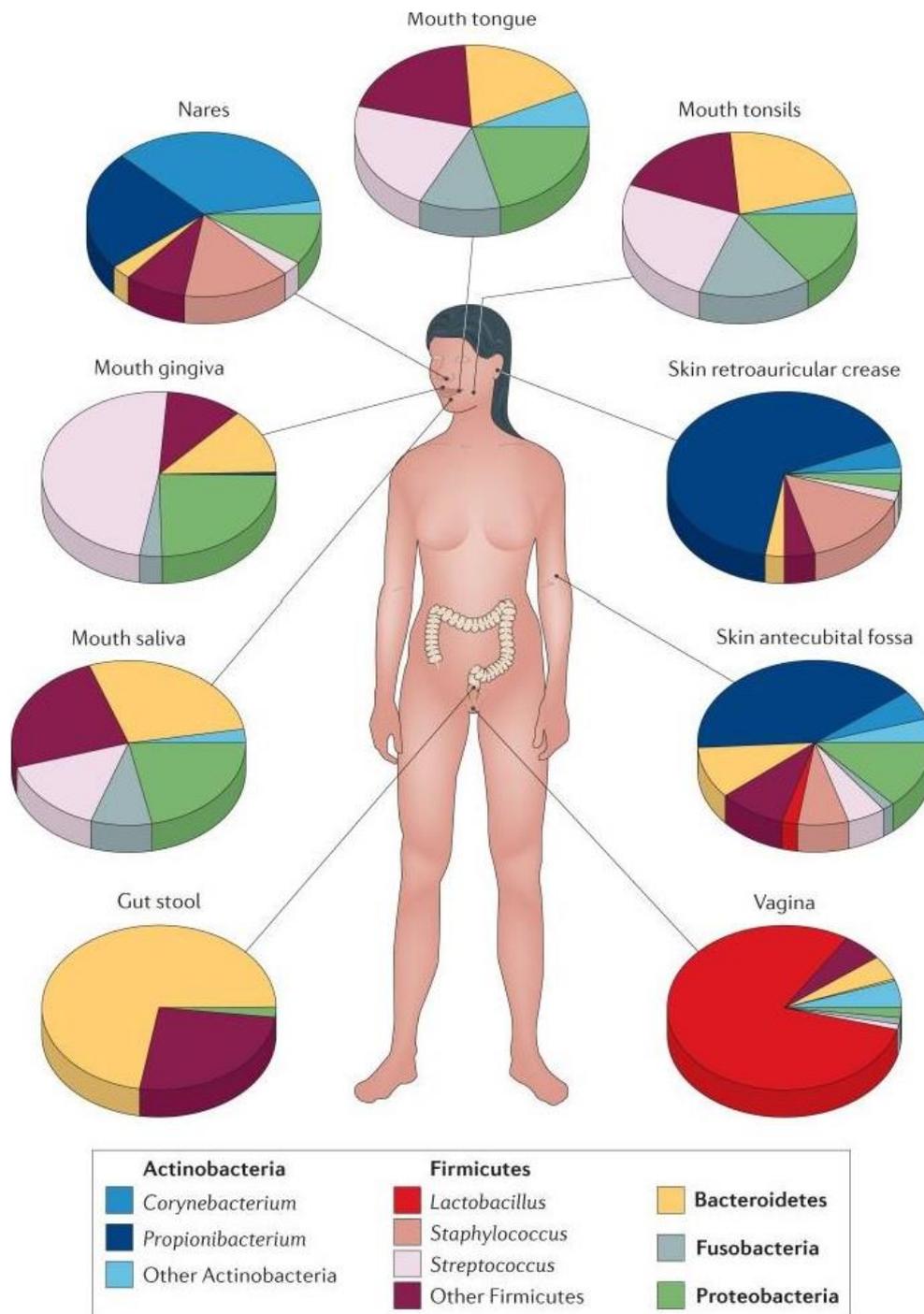


Figure 1.1 Overview of the composition of the human microbiota. Composition of the microbial populations around the body. Whereas other sites around the body show great diversity in phyla the gut bacteria show dominance of Bacteroidetes and Firmicutes. Taken from Lasken and Mclean (2014).

Carbohydrates can be fermented to produce energy for cellular functions. Although, due to the diversity of glycan structures present specific enzymes are required for breakdown into mono- or disaccharides. Bacteria in the human gut, the human gut microbiota (HGM), have developed a symbiotic relationship with their host. The human gut provides the HGM with carbohydrates in the form of indigestible (to the host) dietary glycans and a relatively safe niche in which to grow, and the HGM provide the host with energy rich short chain fatty acids (SCFAs) and various other health benefits (Backhed *et al.*, 2005). Other sites around the body carry microbes with health benefits (Figure 1.1), but are not as densely populated as the gut. The microbes at these other sites are diverse at the phylum level despite the lower microbe density (Figure 1.1). The focus of this thesis is the study of the breakdown of several classes of glycan by two *Bacteroides* species typically present in the HGM and any potential cross-feeding which occurs during the glycan utilisation process.

1.2 Glycans available in the Human Gut

The glycans available to the HGM are primarily derived from the host and diet, the latter being the most significant source of carbohydrate. Host glycans are those expressed on the surface of epithelial cells or secreted into the protective mucus layer of the gut. Dietary glycans are carbohydrates consumed by the host, and are usually storage or structural (component of the cell wall) plant polysaccharides.

1.2.1 Host Derived Glycans

Within the thick, protective mucous layer of the large intestine are high molecular weight (>1000 kDa) glycoproteins, mucins, that contribute to the lubricative and viscoelastic properties of mucus. Mucins can be secreted, form an extracellular gel, or are appended to the cell membrane of epithelial cells, each contributing to the defensive barrier of the gut. Mucin expression is tissue specific with over 21 human mucin genes identified to date (Dharmani *et al.*, 2009). Typically, mucins consist of a peptide core decorated with sugar side chains; a simplified structure is shown in Figure

1.2 (Bansil and Turner, 2006). Thus, the mucins produced by any one human are diverse with different structures found at epithelial sites around the body. The protein core of mucins undergo extensive O- with some N- glycosylation, with attached sugar representing 90 % of the mass of the glycoprotein (Thornton *et al.*, 2008). The glycosylated core of the peptide is typically flanked by cysteine rich domains (C domains) and von Willebrands factor domains (D domains, Figure 1.2). Terminal C domains are essential for polymerisation of mucins through disulphide bonds. The structure of the glycan decorations of mucins is complex often involving charged and uncharged sugars with many unique linkages (Bansil and Turner, 2006; Thornton *et al.*, 2008).

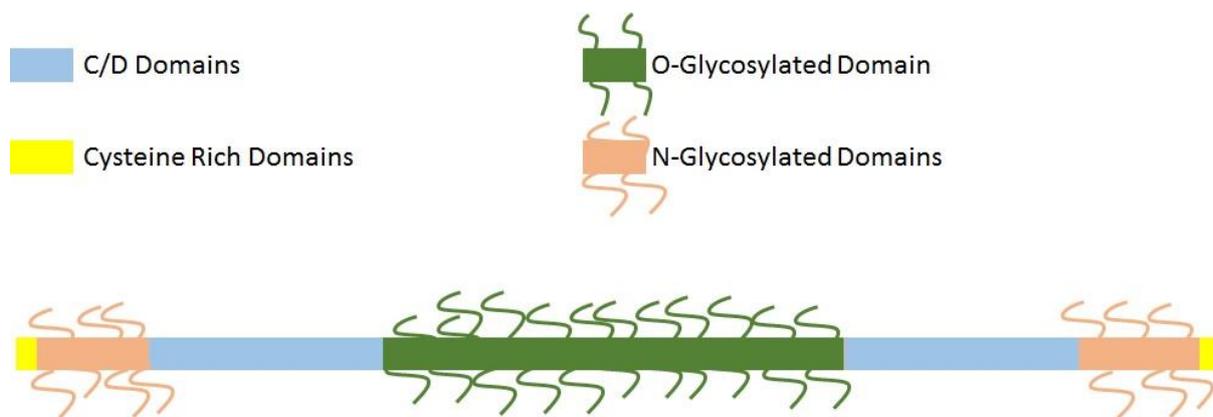


Figure 1.2 Diagram of mucin domain structure. Diagram of the common domain structure of gel forming mucins. Central mucin domains include Serine, threonine and proline residues providing many O-glycosylation sites, which are interrupted with cysteine rich domains. N- and C- terminal domains are cysteine rich allowing intra- and inter- molecular disulphide bonds. Based on figure from Ndeh (2013).

1.2.2 Dietary Glycans

As omnivorous mammals, humans consume nutrients from a wide range of sources including substantial amounts of plant material (Backhed *et al.*, 2005). Plant cells possess a thick polysaccharide rich cell wall which provides the cell with structural integrity, intracellular communication, microbial defence and aids water transport. The plant cell wall is made up of many different classes of polysaccharide which perform different roles at different stages of growth, matching the requirements of the cell at each stage (Figure 1.3) (Cosgrove, 2005).

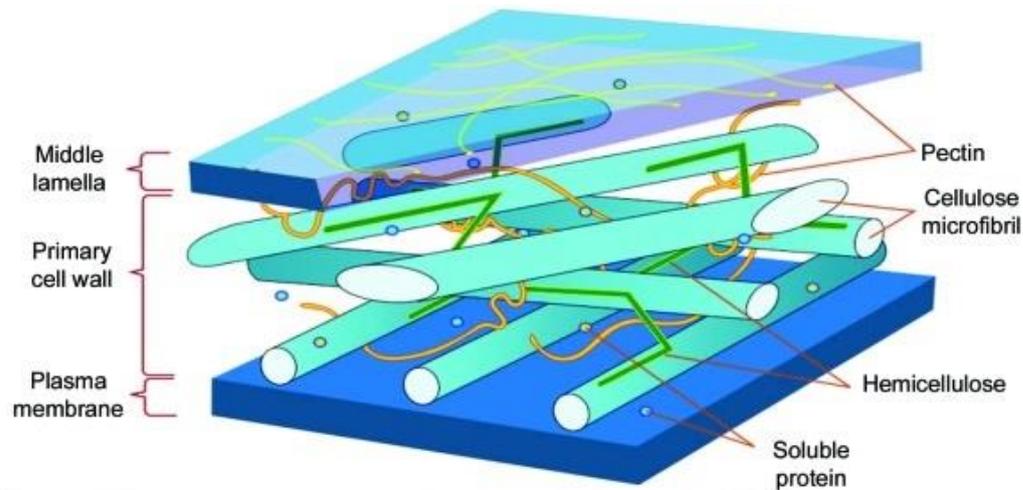


Figure 1.3 Schematic diagram of plant cell wall structure. Shown are 3 main regions of the plant cell wall, the middle lamella, primary cell wall and secondary cell wall, which extends out from the plasma membrane. Shown in the primary cell wall are pectin (yellow), hemicellulose (dark green) and cellulose microfibrils (light green). Adapted from Agirre *et al.* (2016).

The plant cell wall consists of multiple distinct layers. The inner most layer, deposited after cell division is the middle lamella. The next layer, the primary cell wall, is then deposited under the middle lamella which is followed by synthesis of a secondary cell wall, positioned between the plasma membrane and the primary cell wall. The secondary cell wall affords the greatest structural strength to the cell wall, due to ordered cellulose microfibril structures present in this layer.

Cellulose is also present in the primary cell wall but is thought to be in a more random arrangement. Primary cell wall, along with the middle lamella, contain pectin, a complex group of negatively charged polysaccharides (due to the presence of galacturonic acid) that form a gel allowing movement of cellulose microfibrils during growth (Figure 1.3) (Cosgrove, 2005).

1.2.2.1 Cellulose

Cellulose is the most abundant organic molecule, not only in the plant cell wall, but in the biosphere. Cellulose is an unbranched polymer of β -D-1,4 glucose where the repeating unit is a disaccharide (Figure 1.4a), cellobiose (Carpita *et al.*, 1997). Synthesised at the plasma membrane, cellulose

performs a purely structural role providing the plant cell wall with considerable strength (Figure 1.4c). Between 30 to 100 cellulose chains associate through extensive hydrogen bonds forming microfibrils (Beguin and Aubert, 1994). The dominant form, cellulose I, shows a paracrystalline/amorphous structure which is interspersed with highly ordered crystalline regions (Figure 1.4b). The ratio of these regions varies depending on plant species, Cellulose in cotton is 70 % crystalline while the alga *Valonia macrophysa* cellulose is 100 % crystalline (Wood, 1988; Lehtio *et al.*, 2003).

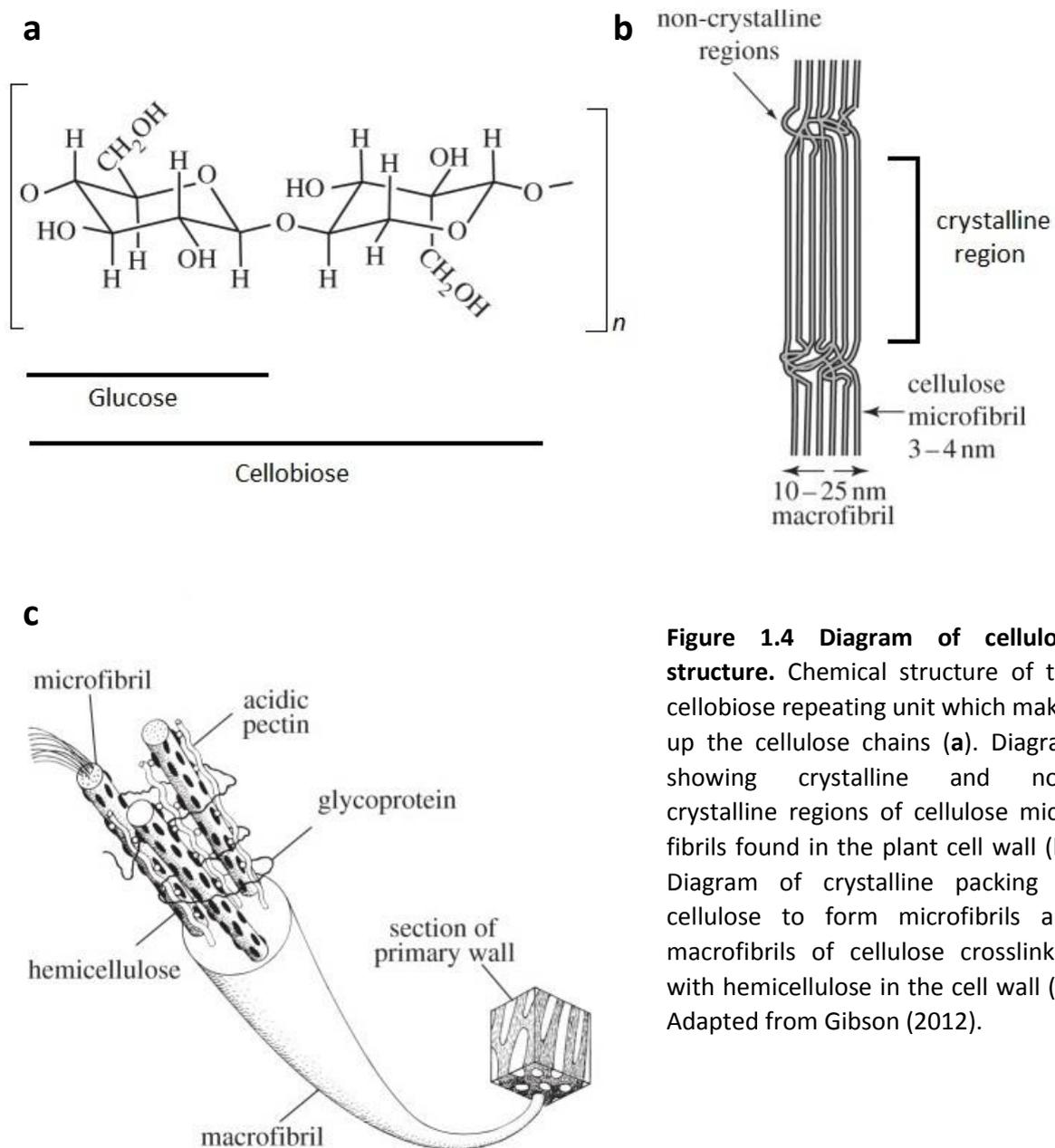


Figure 1.4 Diagram of cellulose structure. Chemical structure of the cellobiose repeating unit which makes up the cellulose chains (a). Diagram showing crystalline and non-crystalline regions of cellulose microfibrils found in the plant cell wall (b). Diagram of crystalline packing of cellulose to form microfibrils and macrofibrils of cellulose crosslinked with hemicellulose in the cell wall (c). Adapted from Gibson (2012).

1.2.2.2 Hemicellulose

Hemicellulose polysaccharides are a group of heterogeneous polysaccharides with a β -1,4-linked backbone of D-sugars including xylans, xyloglucans, mannans and β -glucans. They crosslink cellulose fibrils within the plant cell wall. The Golgi apparatus is the site of hemicellulose synthesis, where it is packaged into vesicles for transport to the plasma membrane (Scheller and Ulvskov, 2010).

1.2.2.2.1 Xylan

Xylan is the predominant hemicellulose in dicot secondary plant cell walls and the majority component of commelinid monocot primary cell walls. Simple forms of the polysaccharide comprise a linear polymer of β 1,4-linked xylose residues with a three-fold screw axis. The backbone xylose units are decorated with acetyl, 2-O-glucuronic acid (GlcA) and 4-O-methylglucuronic acid (Me-GlcA) substitutions in glucuronoxylans (GX, Figure 1.5). There is plant- and tissue-specific variation in the structure of xylans. Arabinoxylan (AX, Figure 1.5) have double or single arabinosyl substitutions to the xylan backbone at the O-2 and/or O-3 position. Different patterns of substitutions are observed in different taxonomic groups. Grass plant cell walls show mostly AX with O-3 arabinosyl substitutions, while double substituted AX is usually found in grass endosperm, and O-2 arabinosyl substitutions are favoured in dicot AX (Scheller and Ulvskov, 2010).

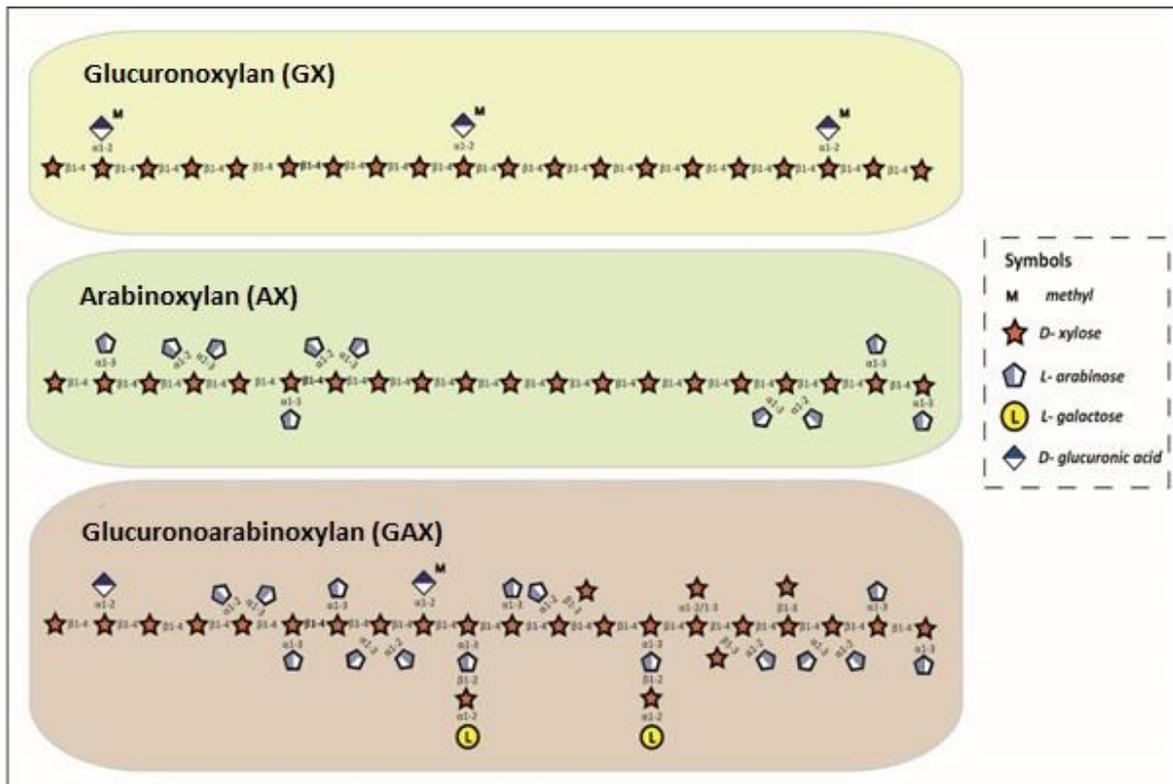


Figure 1.5 Diagram of xylans Structure. Schematic diagram of generalised glucuronoxylan (GX), arabinoxylan (AX) and glucuronoarabinoxylan (GAX) structures. Adapted from Rogowski et al. (2015).

The most complex xylan structure known is that of corn arabinoxylan (CX), a type of glucuronoarabinoxylan (GAX, Figure 1.5). Each of the substitutions of AX and GX are represented in CX along with 1,2/1,3-linked α - or β - xylopyranose and α -L- and α -D-galactose units. A feature common to dicot and conifer xylans is the tetrasaccharide 4- β -D-Xylp-(1 \rightarrow 4)- β -D-Xylp-(1 \rightarrow 3)- α -L-Rhap-(1 \rightarrow 2)- α -D-GalpA-(1 \rightarrow 4)-D-Xylp at the reducing end of the chain. This oligosaccharide is thought to be either an initiator or terminator of xylan backbone synthesis, although there is still much debate of the purpose of this structural feature (Scheller and Ulvskov, 2010). Interestingly, this oligosaccharide has not been found in grass xylans, despite conservation of genes responsible for its synthesis (Scheller and Ulvskov, 2010).

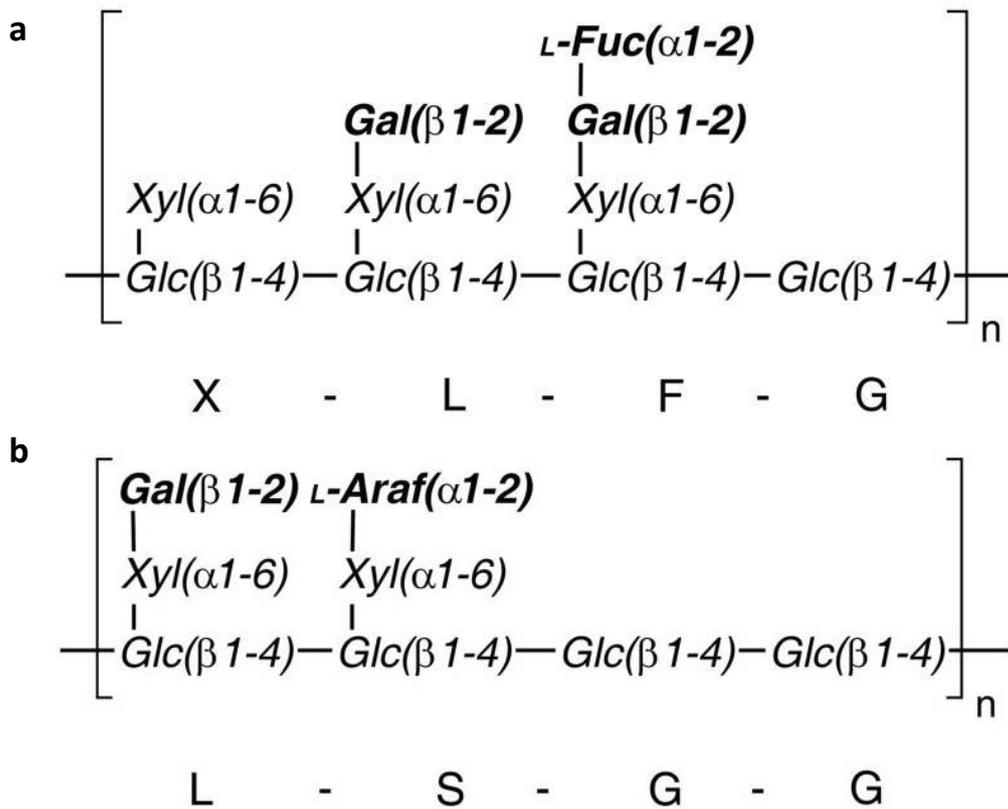


Figure 1.6 Diagram of Xyloglucan repeating unit structure. Representative structures of XXXG-type (a) and XXGG-type (b) xyloglucans with optional residues shown in bold. Taken from Larsbrink *et al.*, (2014).

1.2.2.2.2 Xyloglucan

Xyloglucan (XyG), found in all terrestrial plant species examined to date, has a β 1,4-linked glucose backbone with regular α 1,6-linked xylose substitutions. These side chains are then further substituted with either α -L-arabinofuranose or β -galactose at the O-2 of xylose. The galactose side chains can then be capped with an α -L-fucose, giving a highly variable structure to the polysaccharide. Despite this, repeating oligosaccharides in XyG have been observed in different groups of plant species. Vascular plants show two repeating core oligosaccharides XXGG (Figure 1.6a) and XXXG (where X and G denote glucose decorated with xylose and undecorated, respectively xylose and glucose, respectively, Figure 1.6b), while XXGG is prominent in commelinoid monocots and XXXG dominates in solanaceous plants, respectively (York *et al.*, 1996).

XyG are common in the human diet, found in quantities of up to 25% of the dry mass of vegetables such as lettuce, tomatoes and onions. Seed derived XyG are used as food thickening agents and in drug delivery systems in the intestine (Scheller and Ulvskov, 2010).

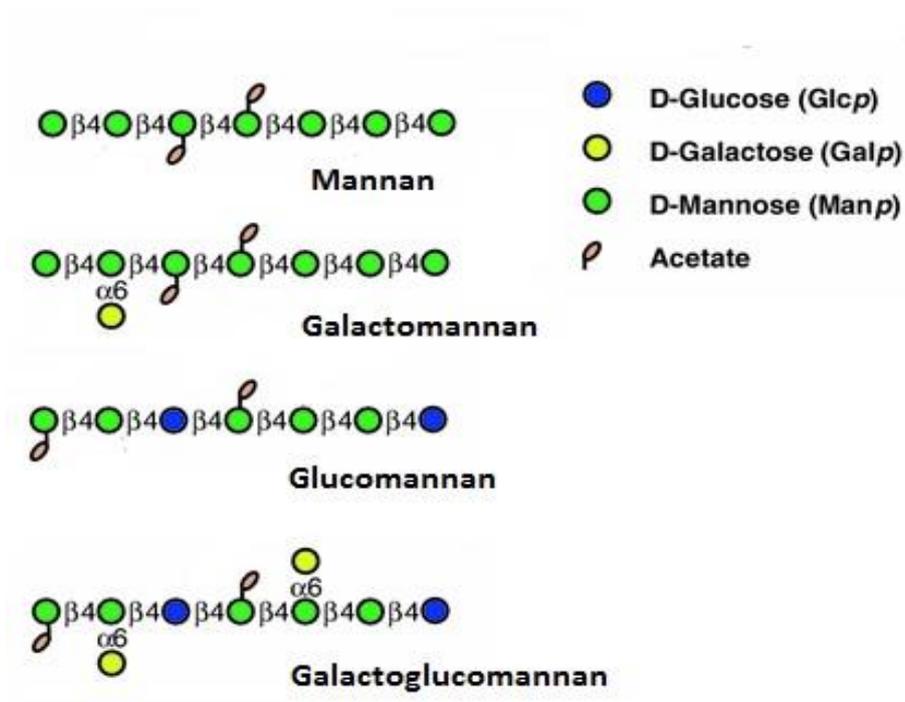


Figure 1.7 Schematic diagram of hemicellulose mannan. Mannan structures found in the hemicellulosic fraction of plant cell wall include mannan, galactomannan, glucomannan and galactoglucomannan. Each mannan structure can be acetylated to greater or lesser degrees. Adapted from Pauly *et al.*, (2013).

1.2.2.2.3 Mannan

The term mannan refers to several polysaccharides each incorporating mannose in their structure.

Mannan polysaccharides include, mannan, galactomannan, glucomannan, galactoglucomannan and glucuronomannan. Undecorated mannan is a β 1,4-linked mannose homopolymer (Figure 1.7), which can form cellulose-like crystalline microfibrils (Moreira and Filho, 2008). Galactomannan comprises a mannan backbone with some α 1,6-linked galactose side chains (Figure 1.7), which varies depending on plant species. Galactomannan contributes to secondary cell wall thickening and uptake of water in seeds (Brett, 1990). The glucomannan backbone comprises β 1,4-linked backbone of mannose and

glucose residues (Figure 1.7). The addition of α 1,6-galactose substitutions to mannose in the backbone creates galactoglucomannan (Figure 1.7). Each of these polysaccharides, with the exception of mannan, can be acetylated to varying degrees at O-2 and O-3 of the mannose units. (Moreira and Filho, 2008).

1.2.2.2.4 β 1,3, β 1,4 mixed linked glucans

Primarily composed of β 1,4-linked glucose interspersed with β 1,3-linked glucose, mixed linked glucan abundance in primary cell wall is dependent on stage of growth of the plant cell. The interspersed β 1,3-linked glucose in the backbone stops mixed glucans from forming cellulose-like crystalline structures, but instead form gel-like structures important during phases of cell growth (Scheller and Ulvskov, 2010).

1.2.2.3 Pectin

Pectic polysaccharides contribute up to 30% of the dry weight of dicot primary cell walls, although they are absent or found in very low abundance in secondary cell walls (Ridley *et al.*, 2001). Pectin contains high quantities of galacturonic acid in the linear backbone from which polysaccharide side chains extend. The major pectic polysaccharides include, homogalacturonan (HG), rhamnogalacturonan-I (RG-I) and rhamnogalacturonan-II (RG-II) (Ridley *et al.*, 2001; Mohnen, 2008).

1.2.2.3.1 Homogalacturonan (HG)

HG is an unbranched structure comprising α 1,4-linked D-galacturonic acid residues, also known as 'smooth regions' of pectins (Figure 1.9). As the most abundant pectic glycan HG accounts for up to 60% of potato (Bush *et al.*, 2001), 35% of tomato (Seymour *et al.*, 1990) and 52% of mango pectins (Muda *et al.*, 1995). The HG backbone generally forms a two-fold helical structure (Jarvis and Apperley, 1995). As a negatively charged polymer, HG is able to form complexes with Ca^{2+} through ionic interactions crosslinking chains of the polysaccharide resulting in stronger, more ridged cell wall structures (Wolf *et al.*, 2009). Methyl esterification reduces the negative charge of HG, which in

turn reduces its capacity to form Ca²⁺-complexes. Methyl esterification of HG is, therefore, associated with increased cell wall fluidity observed during cell separation (Mohnen, 2008; Wolf *et al.*, 2009).

1.2.2.3.2 Rhamnogalacturonan II (RGII)

RGII is a highly conserved complex polysaccharide consisting of 20 different glycosidic linkages and 13 distinct sugars (Figure 1.8), including rare sugars such as, apiose, aceric acid, 3-keto-3-deoxy-manno-octulosonic acid (Kdo) and 3-deoxy-lyxo-2-heptosaric acid (Dha) (Vidal *et al.*, 2000; Mohnen, 2008). These diverse monosaccharides are components of four side chains, two disaccharides and two highly complex oligosaccharides, covalently linked to the α 1,4-galacturonic acid backbone (Mohnen, 2008). The structure of RG-II is highly conserved and pectin is ubiquitous in plant cell walls indicating a vital role which is not fully understood at present, although the glycan may contribute to cell wall strength by forming dimers through inter-chain boron-diester bonds (Ishii and Matsunaga, 1996).

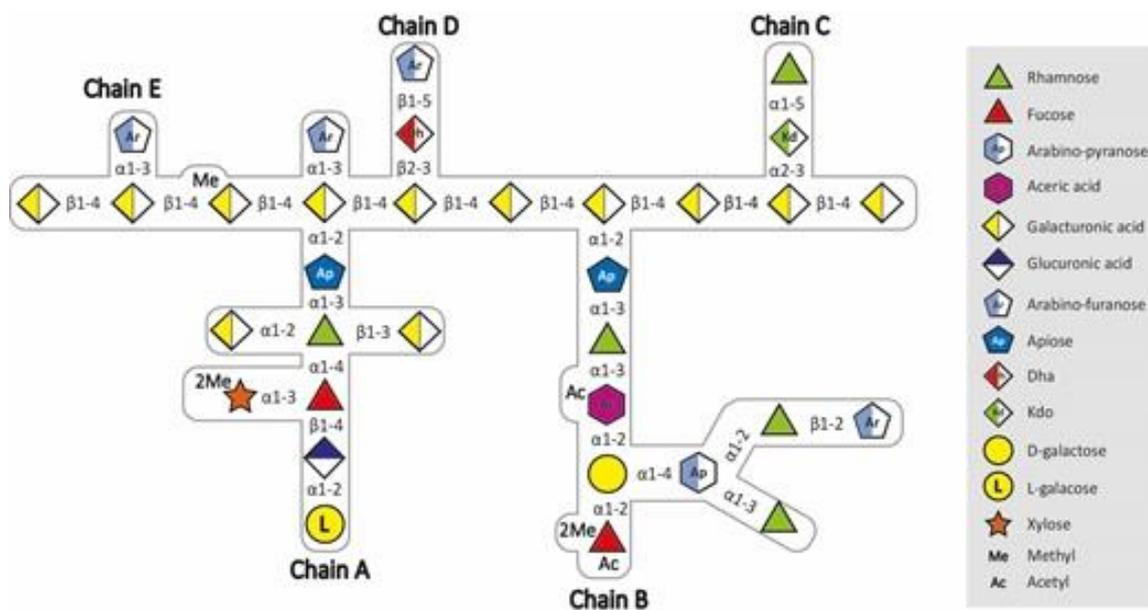


Figure 1.8 Diagram of RGII pectin. RGII pectin is the most complex polysaccharide in the plant cell wall with many unique linkages. The β 1,4-galacturonic acid backbone is decorated with four unique side chains, Chain A-D, with some forms of the pectic polysaccharide containing an arabinose decoration termed chain E. Used with permission of Dr Rogowski.

1.2.2.3.3 Rhamnogalacturonan I (RGI)

The RGI backbone comprises $[-\alpha\text{-D-GalpA-1,2-}\alpha\text{-L-Rhap-1,4-}]_n$ repeating units with approximately half of all Rhap residues substituted at C-4 by galactan, arabinan and arabinogalactan side chains (Figure 1.9) (Lau *et al.*, 1985; Mohnen, 2008). Unlike RGII, RGI side chains can extend to a considerable degree of polymerisation (DP). For example, soy bean galactan is shown to be between 43 and 47 residues in length (Nakamura *et al.*, 2002). RGI and associated polysaccharides can account for around 36% of potato tuber cell wall dry weight (Øbro *et al.*, 2004). The RGI backbone is predicted to form a threefold helix in an extended conformation (Engelsen *et al.*, 1996). The galactan side chain consists of β 1,4-linked galactopyranose (Galp) residues (Figure 1.9), which form a right-handed helix with 6-9 residues per turn (Mohnen, 2008; Cid *et al.*, 2010). Pectic arabinan is found in two configurations, linear arabinan, an α 1,5-linked arabinofuranose (Araf) backbone and branched arabinan (Figure 1.9), in which the linear Araf glycan is decorated at O-2 and/or O-3 with monomeric Araf residues (Oomen *et al.*, 2002). A short galactooligosaccharide connects pectic arabinan to the RGI backbone (Mohnen, 2008). An increased arabinan and galactan content is associated with the developmental switch from division to elongation, suggesting galactan biosynthesis is dependent on certain developmental cues (Ridley *et al.*, 2001). A decrease in arabinan and galactan is usually accompanied by an increase in HG-Ca²⁺ complexes, associated with cell wall strength and rigidity, leading to the hypothesis that proposes arabinan and galactan contribute to cell wall flexibility (Mohnen, 2008). Another suggested role for these long side chains is to crosslink RG-I with other cell wall components, xylans, xyloglucans, cell wall proteins and lignins (McNeil *et al.*, 1980; Ishii and Matsunaga, 1996; Fleischer *et al.*, 1999).

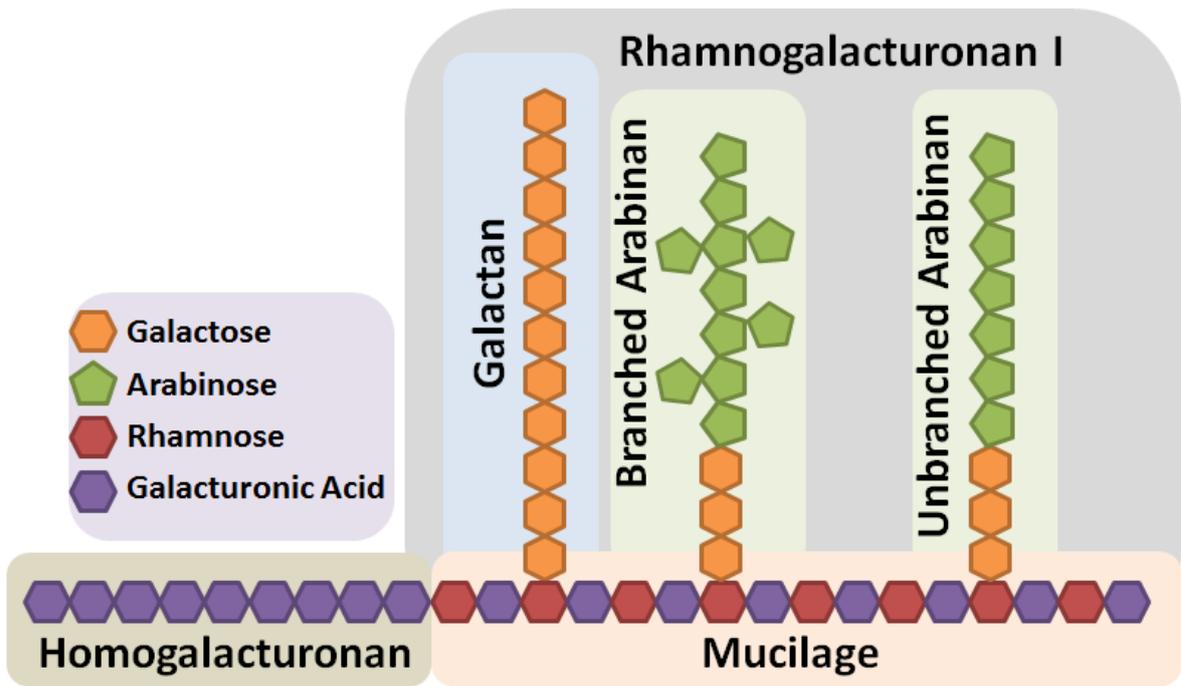


Figure 1.9 Diagram of RGI and HG pectin. RGI and HG pectin is a combination of homogalacturonic acid backbone and 'hairy regions' of RGI. The hairy regions of RGI are heavily decorated with long arabinan and galactan side chains. The hairy regions of RGI are formed from a repeating unit of Rhamnose-galacturonic acid with side chains attached to the rhamnose residue. The galactan side chains are homo-galactan polymers while the arabinan side chains, either branched or unbranched are attached to rhamnose via a short galactooligosaccharide. Aradopsis mucilage is comprised exclusively of the RGI backbone.

1.2.2.4 Fructans

Fructan polysaccharides are primarily composed of fructose units and synthesised by both plant and microbial cells. Unlike the polysaccharides discussed above, fructans are predominantly used as storage polymers, and are the second most abundant non-structural polysaccharides (Hendry, 1993).

1.2.2.4.1 Inulin

Inulin is a β 2,1-linked fructose polymer with an α 2,1-linked reducing end glucose cap (Figure 1.10c), which fulfils a storage role in plant cells. The fructose chain can range between 20-100 fructose residues in length (Roberfroid, 2005). Inulin is synthesised from sucrose (Figure 1.10b), a fructose-glucose disaccharide, which is elongated by accepting fructose through the action of a fructosyltransferase. Inulin oligosaccharides are referred to as kestooligosaccharides, while a

mixture of inulin oligosaccharides are referred to as fructooligosaccharides (FOS). Branching of inulin polysaccharides occurs rarely in microbes and not at all in plants. Inulin chains DP of 9 or more form a six fold helical structure, favouring right-handed helix, although left handed helices have also been observed (French, 1988).

Inulin is mainly produced in dicotyledons, such as chicory roots and is used, along with FOS, to supplement certain foods to boost *Bifidobacteria* and other gut bacteria within the large intestine (French, 1988). Inulin and FOS have been shown to increase mineral absorption of calcium and iron and increased production of SCFAs in the large intestine (Van de Wiele *et al.*, 2004). As well as these positive health effects inulin shows similar properties to fat when added to food with a neutral taste, making it an excellent alternative in low calorie dairy products (Akalın and Erişir, 2008; Buriti and Saad, 2014).

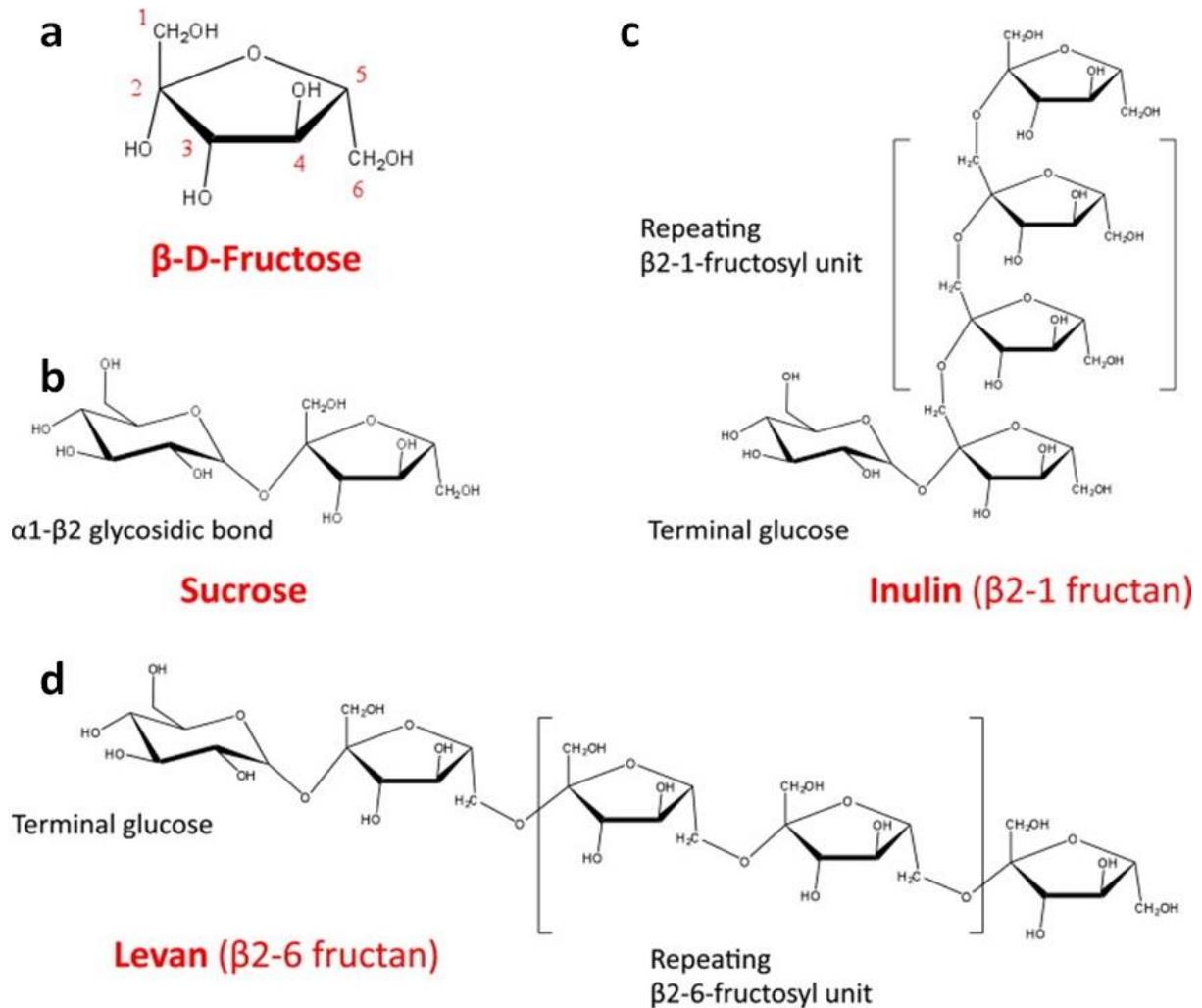


Figure 1.10 Fructan structures. Fructose (a) forms a disaccharide with glucose to become sucrose (b) which is used in plant and bacterial cells to create polysaccharides of β 2,1- or β 2,6-linked fructose, termed inulin (c) and levan (d) respectively. Each polysaccharide terminates in a glucose. Adapted from Sonnenburg *et al.*, (2010).

The medical industry has utilised inulin gel solution to protect drugs that have to transit the human digestive system to deliver pharmaceuticals to the colon. Inulin remains intact until it reaches the large intestine where it is broken down by the HGM and the drug is released (Fuchs, 1987). Inulin is a highly flexible polysaccharide, due to the bond type between each monosaccharide, which forces the furanose ring outwards, away from the plane of the backbone (Illustrated in Figure 1.10c and 1.10d). In addition to storage, inulin may play a regulatory role in cellular compartment sucrose levels, preventing feedback inhibition of photosynthesis in the plant cell (Vijn and Smeekens, 1999).

1.2.2.4.2 Levan

Predominantly found in microorganisms, although low molecular weight forms are produced by some monocotyledon plants, levan is a β 2,6-fructose polymer (Figure 1.10d). High molecular weight levan found in microbes are highly branched with β 2,6-linked chains of fructose attached to the backbone via a β 2,1-linkage. The degree and extent of branching in levans is dependent on species. Levan is less flexible than inulin, adopting a more stable left-handed helix due to more extensive hydrogen bonding than seen in inulin (Han, 1990).

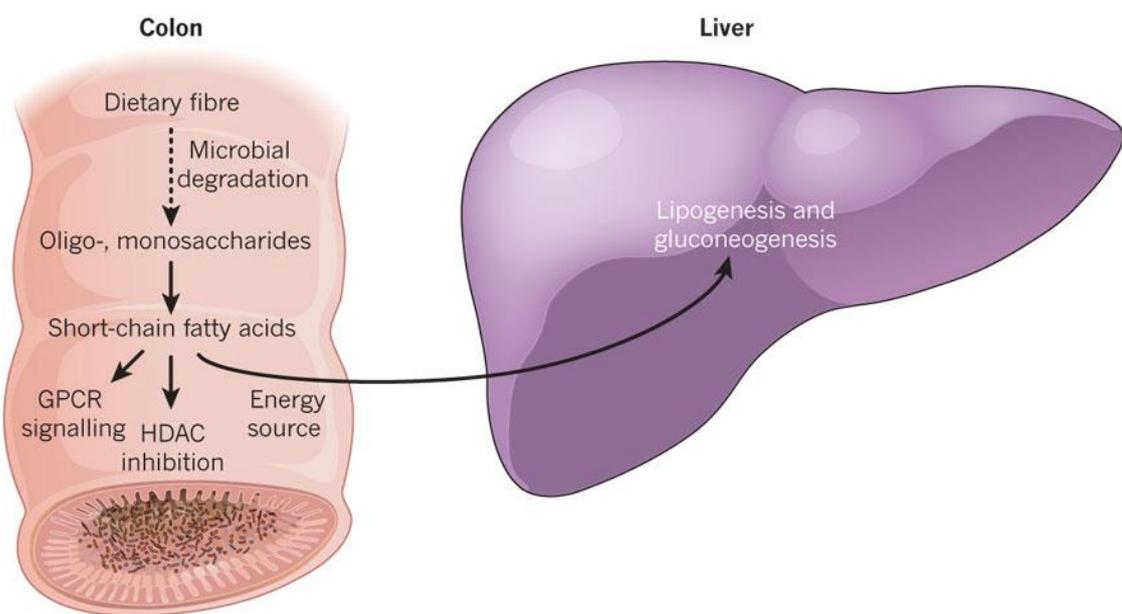


Figure 1.11 An overview of gut microbe metabolism from dietary fibre to effects on the host. Potential effects on the host of production of SCFA metabolites including signalling in host cells via G-protein coupled receptors (GPCR) and inhibition of Histone Deacetylases (HDACs) causing transcriptional changes in the host. SCFAs can also be metabolised in the liver and muscles. From Tremaroli and Backhed (2012).

Lower molecular weight levan in plant cells is used as a storage polymer, primarily found in root tissues, while microbial levan is secreted to form the capsule and biofilm layer. Levan producing bacteria include *Bacillus spp* and *Streptococcus salivarius*, which are found in soil and as oral symbionts, respectively. Levan of oral symbionts is often found in dental plaques (Higuchi *et al.*, 1970). *Erwinia herbicola*, *Zymomonas mobilis* and *B. subtilis* produce levan which assists in biofilm

adhesion, likely due to their adhesive and elastic aqueous properties (Blake *et al.*, 1982; Benigar *et al.*, 2014).

1.3 Human Gut Microbiota

The adult human intestine hosts a microbial population of trillions of organisms dominated by bacterial species. This microbial community is generally known as the human gut microbiota (HGM), and forms a symbiotic relationship with the host (Figure 1.11) (Backhed *et al.*, 2005). From birth the HGM develops in concert with the host. Shaped by the human diet at each stage of growth, the HGM varies throughout life.

1.3.1 Microbiota Composition and Diversity

Development of high throughput sequencing techniques have allowed probing of the incredibly complex HGM to explore the species composition, where previously it was only possible to identify species that could be cultured (Metzker, 2010). These advances gave way to the Human Microbiome Project (HMP) in 2008 with the aim of characterising the human microbiome, including the gut microbiome (genomes or metagenomes of the HGM). Originally those working on the project believed a core microbiome could be established for healthy individuals, but healthy individuals were found to possess incredibly divergent microbiomes. Three broad enterotypes were established during the MetaHIT study that were defined by variations in the prominence of the three dominant genera, *Bacteroides*, *Prevotella* and *Ruminococcus* (Arumugam *et al.*, 2011), although this can be argued to be an oversimplification of an incredibly complex microbial community (Knights *et al.*, 2014). The MetaHit and HMP studies generally use stool samples to probe the HGM, as this technique is much less invasive than endoscopy. Stool samples, however, may only give an idea of which species are present at the distal end of the colon. Animal studies using humanised microbiota have shown phylogenetic diversity between the early and late gastro-intestinal (GI) tract compared with mid GI tract, where diversity was found to drop (Gu *et al.*, 2013). Differences are believed to be due to variations in oxygen along the GI tract, where obligate anaerobes were found in locations

more likely to provide anaerobic conditions (Gu *et al.*, 2013). Similar data have shown microbe diversity at different locations in the GI tract of pigs, where the early GI tract was found to be Firmicute rich while the late GI tract showed dominance of Bacteroidetes (Kim and Isaacson, 2015). Sequencing of 16S ribosomal DNA was deployed to probe this vast population, results of which showed incredible diversity at the species level with 1000s of species present, while at the phylum level Bacteroidetes and Firmicutes represent 90% of the HGM (Figure 1.12) (Qin *et al.*, 2010). Despite greater diversity at the species level 25% of gut microbial species belong to the *Bacteroides* genera, a significant proportion of this microbial community (Martens *et al.*, 2009). The genus *Bacteroides* belongs to the phyla *Bacteroidetes* and includes over 20 distinct species. Although most species of *Bacteroides* are considered commensal under some conditions *Bacteroides fragilis* can become pathogenic. *Bacteroides spp* are anaerobic rod-shaped, bile-resistant, non-sporulating, gram-negative bacteria typically found in the gut (Wexler, 2007). Sequenced in 2003 (Xu *et al.*, 2003) and 2005 (Cerdeno-Tarraga *et al.*, 2005) respectively, *B. thetaiotaomicron* and *B. fragilis* both have a relatively low gene content for size of the genome, indicating a large number of high molecular weight proteins are expressed (Wexler, 2007). *Bacteroides spp* begin to populate the human gut approximately 10 days after birth, although *Bacteroides spp* are more prominent post-weaning or in infants which were not breast fed (Simon and Gorbach, 1986; Mackie *et al.*, 1999). Products of *Bacteroides* metabolism, primarily SCFAs, provide a significant contribution to the daily energy requirements of the host (Hooper *et al.*, 2002).

Another genus associated with positive health is *Bifidobacterium*. Belonging to the diverse phyla *Actinobacteria*, *Bifidobacterium spp* are gram-positive, non-motile, non-filamentous, Y-shaped bacteria without the ability to form a capsule (Barka *et al.*, 2016). Interestingly, the Y-shaped cells are only maintained in clinical isolates from the gut, when cultured *in vitro* however, the cell revert to a rod shape (Barka *et al.*, 2016). *Bifidobacterium spp* have been shown to exert antimicrobial activity by competitive exclusion while also adhering to the intestinal wall or mucus layer of the gut (Ouweland *et al.*, 2002).

Although each individual host has a unique bacterial population, trends have been identified that are associated with particular host diets or phenotypes (Qin *et al.*, 2010; Tremaroli and Bäckhed, 2012). Experiments investigating differences in gut microbe composition between genetically obese mice (*ob/ob*) and their lean littermates have revealed an association with increased Firmicute to Bacteroidete ratio in the obese phenotype (Ley *et al.*, 2005), an observation which is mirrored in human studies (Ley *et al.*, 2006). Further, *in murine* investigations have revealed genetically obese mice possess a greater intestinal SCFA concentration and reduced energy content of faecal matter than lean mice on the same diet, implying more efficient gut microbe composition in the obese phenotype. Metagenomic studies of the gut microbiota of obese mice demonstrated a greater capacity for glycan degradation and utilisation, which again is mirrored in studies on human gut microbes. Transplantation of the obese gut microbes into lean mice showed a two-fold increase in weight gain than lean mice given microbes from lean donors (Turnbaugh *et al.*, 2006). Interestingly, germ-free mice, without any gut microbes, require up to 30% more nutrients than littermates with normal gut microbiota to grow at the same rate (Gilmore and Ferretti, 2003).

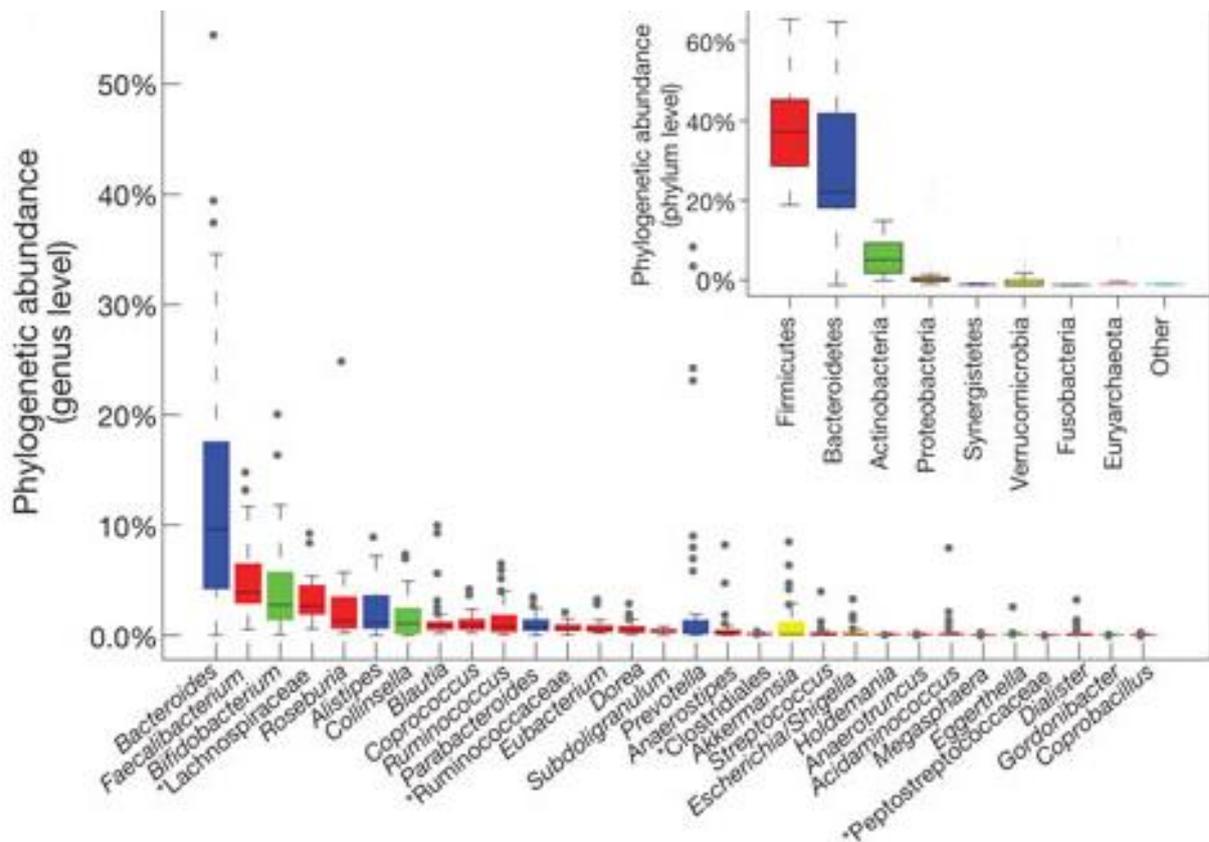


Figure 1.12 Diversity of phyla and genera within the HGM of healthy individuals. The box plot shows the thirty most abundant genera, while the inset shows abundance at the phylum level. Genus and phylum level abundances were calculated using reference genome based mapping using 85% and 65% cutoffs, respectively. Taken from Arumugan *et al.*, (2011).

Host diet plays a major role in shaping the gut microbiota composition (Tremaroli and Bäckhed, 2012). A recent study contrasted the microbial composition from faecal samples of children from Africa and Italy. The African diet included high amounts of plant polysaccharides correlating with an increased Bacteroidetes to Firmicutes ratio in faecal samples, with *Prevotella* species being particularly enriched. The Italian children gut microbiota showed higher levels of *Enterobacteriaceae*. The African gut microbiota had adapted to maximise energy yield from the polysaccharide rich diet, selecting for species with greater glycan utilisation capacity, in this case *Prevotella* and other Bacteroidetes (Turnbaugh *et al.*, 2009).

Studies in which participant diets are supplemented with resistant or non-fermentable starch have shown significant microbiota change in which *Eubacterium rectale*, *Oscillobacter spp* and

Ruminococcus bromii were enriched (Walker *et al.*, 2011). Each of these have been shown to utilise resistant starch. *Bi. adolescentis* also dramatically increased (Martinez *et al.*, 2010), consistent with its ability utilise starch (Duranti *et al.*, 2014). Drastic changes in HGM composition only seems possible with long-term dietary changes, as a study that monitored microbial composition as a result of dietary intervention showed no significant change during the 10 day experiment (Wu *et al.*, 2011). The investigators did observe a selective increase for *Bifidobacterium* spp when daily doses of inulin, a prebiotic, were administered (Wu *et al.*, 2011). In genetically obese mice enrichment for *Bifidobacterium* spp correlated with reduced adiposity and a reduction in lipopolysaccharide, a known microbial-derived inflammatory molecule, when compared to *ob/ob* mice on the same diet without prebiotics (Cani *et al.*, 2007). Interestingly, only a small number of species in the HGM appear to be effected by inulin dietary supplementation, and these organisms are in the genera *Bifidobacterium* and *Atopobium* (Costabile *et al.*, 2010). A wider range of species were enriched when participants were supplemented with FOS rather than inulin. In these studies, *Bifidobacterium* and *Bacteroides* spp are particularly abundant post intervention, while *Faecalibacterium prausnitzii* and *Roseburia intestinalis* were reduced in abundance (Benus *et al.*, 2010).

Unlike dietary fibre, fat appears to have an indirect effect on HGM by modulating bile acid secretion and composition. Interestingly, the *Bacteroides* enterotype (defined in the MetaHIT study) positively correlates with intake of saturated fat (Wu *et al.*, 2011). Polyunsaturated fats were found to affect adherence of gut microbes to the intestinal wall (Kankaanpaa *et al.*, 2001), leading to a reduced presence of *Bacteroides*, *E. rectale/Clostridium coccoides* group and *Bifidobacterium* (Cani *et al.*, 2007). Dietary proteins which reach the intestine are metabolised by microbial proteolysis or gut fermentation, generating gasses and SCFAs (Russell *et al.*, 2011). Over a six week dietary intervention, participants with high-protein low-carbohydrate diet were found to have reduced *Bifidobacterium* spp specifically and total bacterial abundance in the gut (Duncan *et al.*, 2007; Brinkworth *et al.*, 2009), potentially increasing the risk of infection with pathogenic bacteria, due to reduced competition from the HGM.

1.3.2 Short chain fatty acid production by the HGM

The microbes of the HGM provide the host with beneficial nutrients and energy in the form of SCFAs generated by metabolism of dietary components that are inaccessible to the human host. SCFAs are the main metabolite products of microbial glycan fermentation in the human gut and provide many of the microbiota associated benefits to the host (Figure 1.13). SCFAs are utilised by colonocytes for up to 70% of their energy requirements (Roy *et al.*, 2006). The three core SCFAs, acetate, butyrate and propionate are found in the gut in a combined concentration of 50-150 mM (Louis *et al.*, 2014). Generally, propionate is produced by Bacteroidetes, while butyrate is generated by Firmicutes such as *Roseburia* species, also by *Bifidobacterium* from the Actinobacteria phylum (Reichardt *et al.*, 2014). When dietary fibre is in short supply microbes can utilise less energetically favourable growth sources, such as amino acids from dietary or endogeneous proteins (Cummings and Macfarlane, 1991). These growth substrates result in reduced fermentative activity. SCFAs and some branched-chain fatty acids are generated from metabolism of branch chain amino acids (Cummings and Macfarlane, 1991). These metabolites are implicated in insulin resistance in host cells (Newgard *et al.*, 2009). SCFA levels are not constant throughout the length of the gut, with higher concentrations being found in the cecum and proximal colon, while levels steadily decline towards the distal end of the colon (Cummings *et al.*, 1987), most likely due to increased absorption through monocarboxylate transporters in the colon wall (Koh *et al.*, 2016).

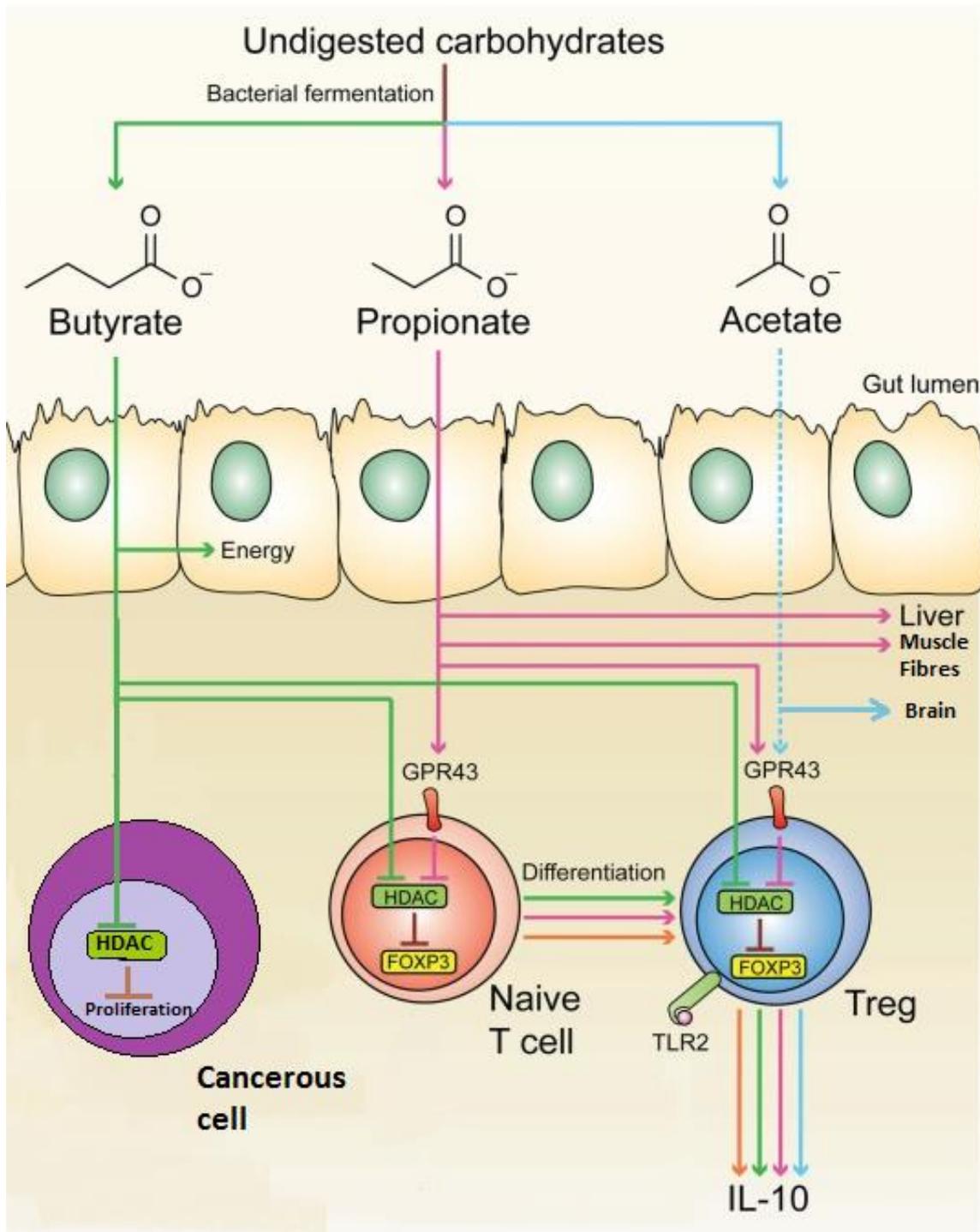


Figure 1.13 Effects of microbe produced SCFAs on host cells. Fermentation of carbohydrates in the anaerobic environment of the gut generates SCFAs, the most abundant are butyrate, acetate and propionate. All three SCFAs act as histone deacetylase inhibitors although in different cells to bring about different outcomes. Butyrate has been shown to inhibit proliferation of cancerous cells. Butyrate and propionate can force differentiation of naïve T-cells to Treg cells which go on to produce IL-10. Acetate is metabolised by the liver and muscle fibres but is also able to cross the blood brain barrier where it induces satiety signals. Butyrate can be utilised by colonocytes for energy. Adapted from Hoeppli *et al.*, (2015).

Propionate, acts as an anti-obesity factor delaying gastric emptying, promoting satiety in the host (Arora *et al.*, 2011). The most abundant SCFA in the human gut is acetate, which is produced either by acetogenic bacteria or as an additional fermentation product from Bacteroidetes (Miller and Wolin, 1996; Louis *et al.*, 2014). Acetate, is absorbed in the colon and metabolised by the liver and the muscles providing 1.5-2 kcal/g of energy for the host (Kien, 1996; Topping and Clifton, 2001), and through stimulation of intestinal epithelium plays a role in prevention of enteropathogenic infection (Fukuda *et al.*, 2011). Once in circulation acetate is capable of crossing the blood-brain barrier where it has been found to reduce appetite via a central homeostatic mechanism (Frost *et al.*, 2014).

Butyrate has been implicated in colorectal cancer prevention by acting as a histone deacetylase inhibitor in cancerous cells to alter expression of genes involved in cell proliferation, apoptosis and differentiation. Despite this effect on cancerous cells, butyrate acts as an energy source for non-cancerous cells, stimulating cell growth This discord is known as the butyrate paradox (Lupton, 2004). Cancerous cells show preference for glucose over butyrate for an energy source, causing higher cellular butyrate concentrations overcoming the threshold required to inhibit histone deacetylases and alter gene regulation in the cancerous cell (Donohoe *et al.*, 2012; Kaiko *et al.*, 2016). Acetate has also been shown to inhibit the action of histone deacetylases, although only in activated T-cells of the human immune system (Park *et al.*, 2015).

1.3.3 Interactions between members of the HGM

Within the HGM there is a complex network of interactions not just between this microbial ecosystem and the host, but also between bacteria within this microcosm.

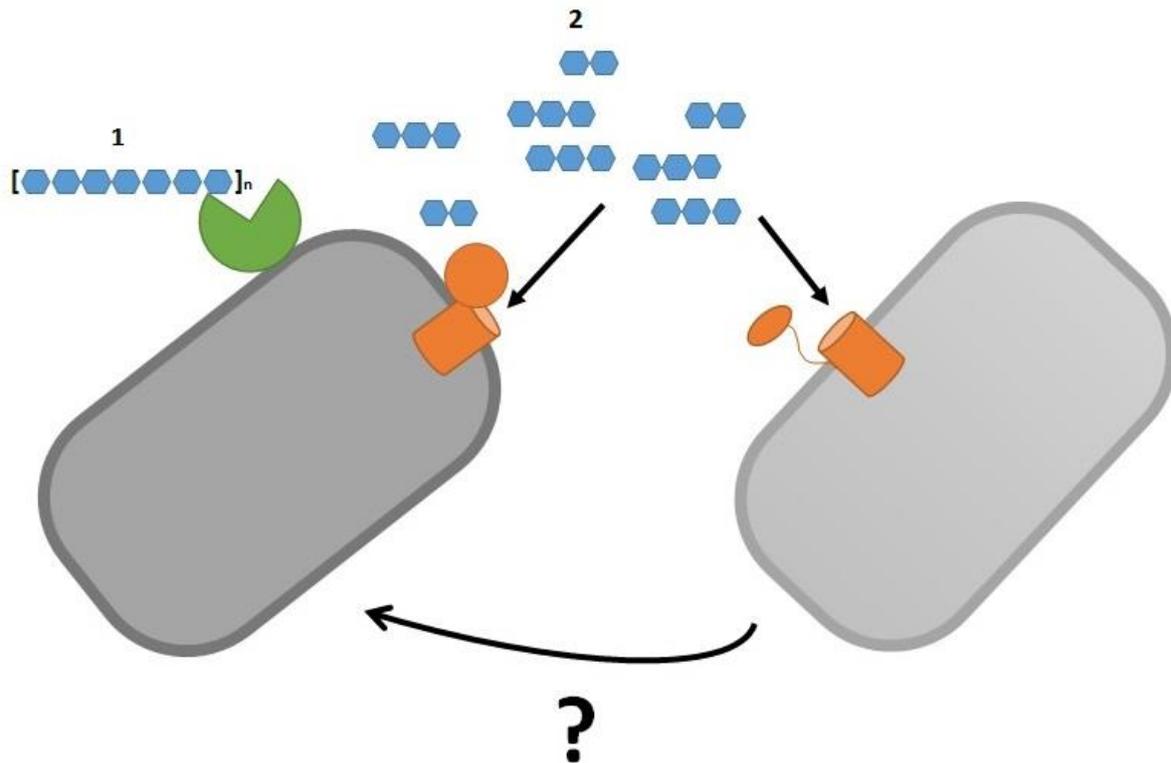


Figure 1.14 Diagram of glycan cross-feeding interactions between members of the HGM. Bacteria possessing the extracellular glycan degrading enzymes produce accessible substrate from inaccessible polysaccharide which is then utilised by other members of the gut microbiota. Polysaccharide (1) is degraded by surface or secreted CAZymes (green) to oligosaccharides (2) and transported by glycan transport proteins (orange). It is hypothesised that there is some benefit to the glycan degrading bacterium beyond generating oligosaccharides it can utilise, however this has yet to be elucidated.

Recent research has shown widespread cross-feeding within the gut microbiota, where one bacterium, through release of oligosaccharides, allows another species to grow on a previously inaccessible substrate (Figure 1.14). Growth of a *Bacteroides* spp unable to utilise the specific target glycan has been shown to grow in the presence of a species able to utilise the target polysaccharide (Rakoff-Nahoum *et al.*, 2014). Evidence of glycanase enzymes were found in outer membrane vesicles, which were secreted into the growth medium (Elhenawy *et al.*, 2014), where oligosaccharides were generated away from the *Bacteroides* cell surface and available for another species to utilise (Rakoff-Nahoum *et al.*, 2016). At first this appears counter intuitive as the HGM is densely populated and highly competitive. Another study showed this is not limited to *Bacteroides-Bacteroides* interactions, clearly demonstrating cross-feeding is an inter-genera interaction (Van der

Meulen *et al.*, 2006). Recently it has been shown that a surface endo-inulinase, expressed by *B. ovatus* in response to inulin, is not required by the bacterium to utilise the polysaccharide. The enzyme, however, generates FOS at the cell surface. FOS is known to be enrich for *Bifidobacterium spp* in the gut, hence termed bifidogenic, and has been connected to various health benefits for the host. This implies *Bacteroides* intentionally modulates the HGM by releasing publicly available oligosaccharides, although the extent to which this occurs, and the benefit of this action for other *Bacteroides*, remains unclear (Rakoff-Nahoum *et al.*, 2016). Although it should be noted that the *B. thetaiotaomicron* mannan utilisation system adopts a selfish utilisation system in which few if any oligosaccharides are released, leading to no cross-feeding during mannan utilisation (Cuskin *et al.*, 2015).

Competition for CO₂ in the gut may lead to competition and modulation of metabolites. *Bacteroides* and *E. rectale* both utilise CO₂, indicating the potential for competition for this resource. When grown alongside one another in the gut of a mouse, propionate is produced in much lower quantities than when *E. rectale* is not present, indicating CO₂ is depleted much quicker in the presence of *E. rectale* forcing *Bacteroides* to produce acetate over propionate (Rey *et al.*, 2010).

1.3.4 Interactions between the HGM and Pathogens

A significant benefit of colonisation of the human gut with commensal bacteria is resistance to colonisation of pathogenic bacteria (Sassone-Corsi and Raffatellu, 2015). Germ-free mice, those treated with antibiotics and bred in a sterile environment, are highly susceptible to pathogenic bacteria such as *Shigella flexneri* (Sprinz *et al.*, 1961), *Listeria monocytogenes* (Zachar and Savage, 1979) and *Salmonella Typhimurium* (Ferreira *et al.*, 2011). In a study of Swedish adults, susceptibility to infection with *Campylobacter jejuni* was shown to be dependent on microbiota composition. Those with increased diversity appeared to be more resistant to infection compared to individuals with lower microbial variation (Kampmann *et al.*, 2016).

The most prevalent theory of how the HGM protect against pathogenic bacterial infection in the gut is based on commensal bacteria out competing pathogens. Studies in which mice colonised with commensal bacteria and *Citrobacter rodentium* show diet plays an important role on the likelihood of the commensals out competing the pathogenic bacterium in the gut. When fed with a diet of mixed monosaccharides and polysaccharides *C. rodentium* is out competed by *Escherichia coli* but not *B. thetaiotaomicron*. The *Bacteroides* organism only outcompetes *C. rodentium* when fed a monosaccharide only diet in which both species compete for the same resource, whereas the monosaccharide-polysaccharide mixture allows *B. thetaiotaomicron* to utilise polysaccharides leaving the monosaccharides available for *C. rodentium* (Kamada et al., 2012).

Clostridium difficile is considered the leading cause of infectious diarrhoea and antibiotic associated pseudomembranous colitis (Rodriguez et al., 2015). Antibiotic treatment causes indiscriminate bacterial killing within the gut, leading to reduced diversity, and hence increases the likelihood of *C. difficile* infection and the resulting colitis. Reduction in both bacterial diversity and abundance in the gut can lead to an increase in metabolite and carbohydrate availability, which are then utilised for growth by antibiotic-resistant pathogens, increasing the severity and rate of infection (Rodriguez et al., 2015).

1.3.5 Effect of prebiotics on the HGM

Prebiotics are described as non-digestible food supplements that beneficially affect the host by specially enriching a limited number of beneficial bacteria (to the host) in the colon (Gibson and Roberfroid, 1995). This was later updated to include the requirement that prebiotic material must resist gastric acidity and not be absorbed by the host (Gibson et al., 2004). Currently FOS and galactooligosaccharides (GOS) are being used as prebiotics, although new evidence suggest xylooligosaccharides and isomaltooligosaccharides may give similar positive health benefits by stimulating growth of targeted species within the HGM (Chapla et al., 2012; Patel et al., 2012;

Christensen *et al.*, 2014). Table 1.1 shows a list of prebiotic glycans shown to have health benefits in human studies.

Glycan	Prebiotic effects	References
Wheat dextran	Increased <i>Bacteroides</i> and decreased <i>Clostridium perfringens</i>	Lefranc-Millot <i>et al.</i> 2012
Inulin/FOS	Bifidogenic	Costabile <i>et al.</i> 2010, Ramnani <i>et al.</i> 2010
GOS	Bifidogenic	Eli <i>et al.</i> 2008
Acadia gum	Bifidogenic	Howarth <i>et al.</i> 2001
Psyllium	Prebiotic potential	Lanza <i>et al.</i> 2007
Polydextrose	Bifidogenic	Howarth <i>et al.</i> 2001, Ley <i>et al.</i> 2006
Gum Arabic	Prebiotic potential	Calame <i>et al.</i> 2008

Table 1.1 Reported probiotic effects of various glycans in human studies. Effects observed are not limited to the references given.

The importance of prebiotics has increased in recent years as the amount of dietary glycans has reduced in the western diet, while refined sugar has increased dramatically (Jew *et al.*, 2009). Introduction of prebiotics into foods may help alleviate the negative effects of the western diet, increasing species diversity and stability of the HGM. Prebiotic FOS and GOS supplements in infant formula has worked to decrease the differences in microbial profiles of infants breast fed and those that are formula fed (Bakker-Zierikzee *et al.*, 2005; Knol *et al.*, 2005). There is evidence that supplementation of inulin in foods can help to re-establish a healthy microbiota during and immediately after antibiotic treatment (Johnson *et al.*, 2015). *Bifidobacteria* are able to ferment colonic glycans to produce butyrate and have been implicated in modulating the immune system of the host, explaining why they have been the target of many of the prebiotic treatments devised (Gibson and Roberfroid, 1995; Gibson *et al.*, 2004; Kanauchi *et al.*, 2013).

1.4 Carbohydrate Active Enzymes

Utilisation of dietary and host glycans requires a number of carbohydrate active enzymes (defined henceforth as CAZymes) and non-catalytic carbohydrate binding modues (CBMs) or proteins. The human genome encodes relatively few enzymes involved in the deconstruction of dietary glycans, when the omnivorous human diet is taken into account. Energy from plant polysaccharides would pass through the digestive tract unutilised if it were not for the action of gut microbial CAZymes, which breakdown polysaccharides into mono- or di-saccharides that are subsequently fermented. The gut microbiome represents an incredible coding capacity for CAZymes (Figure 1.15a), the majority of which are directed against plant glycans (Figure 1.15b), reflecting the importance of these carbohydrate polymers in HGM metabolism (El Kaoutari *et al.*, 2013).

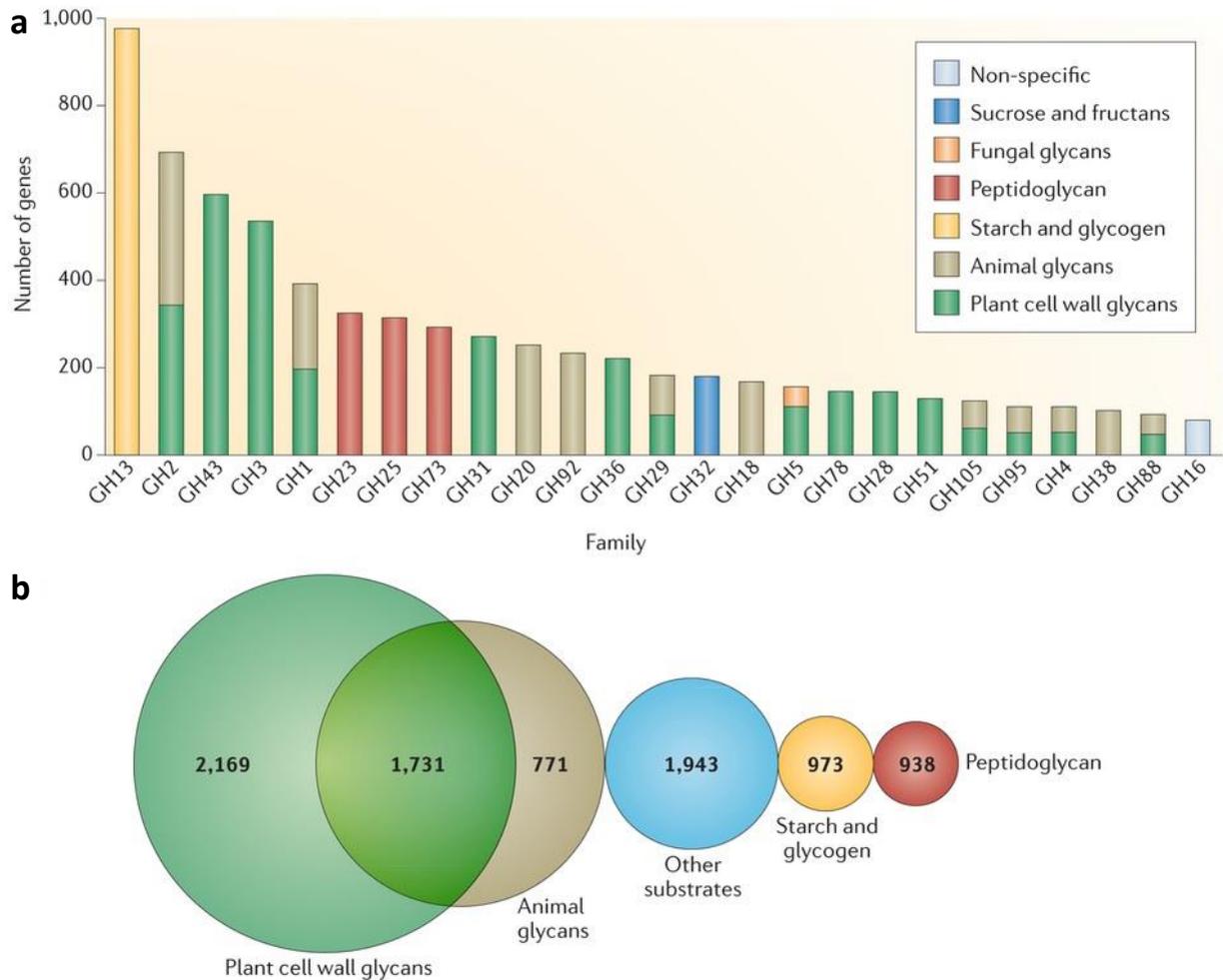


Figure 1.15 Expansion of glycoside hydrolases (GHs) in gut microbiota. Diversity of identified GH genes in a representative gut microbiota, predicted or confirmed substrate specificity is denoted by colour (a). Venn diagram of substrate specificity of the glycoside hydrolases showing dominance of GHs directed against plant glycans (b). Taken from El Kaoutari *et al.* (2015).

The data presented in this thesis are focused on glycoside hydrolases, although it is worth noting there are a considerable number of enzymes, other than those that hydrolyse glycosidic bonds, which contribute to the deconstruction of glycans in the gut.

1.4.1 Classification of Carbohydrate Active Enzymes (CAZymes)

Classification of CAZymes into sequence based families on the CAZy database (www.cazy.org) has been an invaluable tool in the studying these enzymes. The CAZy database aims to catalogue all

known CAZymes (Cantarel *et al.*, 2009; Lombard *et al.*, 2014). Enzyme families are annotated according to their functional classes, Glycoside Hydrolase Families (GHs), Polysaccharide Lyases (PLs), Carbohydrate Esterases (CEs) and Auxiliary Activities (AA). CAZy classification is primarily based on amino acid sequence similarity to members of known families, although in borderline cases classification is made on a module by module basis. CAZymes are typically modular proteins, often including catalytic and non-catalytic modules, and are thus located in multiple families. Unlike DNases and RNases, CAZymes with similar substrate specificities usually show significant sequence similarity and hence, classification into a CAZy family can provide insight into substrate specificity (Cantarel *et al.*, 2009; Lombard *et al.*, 2014). It should be emphasised, however, that some of the larger families, such as GH5, GH2 and GH43, contain enzymes with a range of specificities. The delineation of these large families into subfamilies will help in defining the activities of different members of these families (Aspeborg *et al.*, 2012; Mewis *et al.*, 2016).

1.4.2 Catalytic Mechanisms

Glycoside hydrolases employ two main acid-base assisted mechanisms of action, retaining (double-displacement) mechanism, which generates a glycosyl enzyme intermediate, or an inverting mechanism, which involves a single displacement without the need to generate a covalent intermediate.

1.4.2.1 Retaining Mechanism

The retaining mechanism involves a two-step reaction, shown in Figure 1.16, in which the enzyme becomes glycosylated and then deglycosylated through an activated water molecule. Here, retaining refers to retention of the configuration of the anomeric carbon, for example, when a β -bond is cleaved and a β -bond is formed. α and β refer to the configuration of the anomeric carbon, C1, with respect to the furthest chiral centre typically C5. If both have the same configuration this is termed β -configuration while sugars with differing configuration at C1 and C5 are termed α -configuration (Koshland *et al.*, 1954; Sinnott, 1990).

Retaining mechanism for a β -glycosidase:

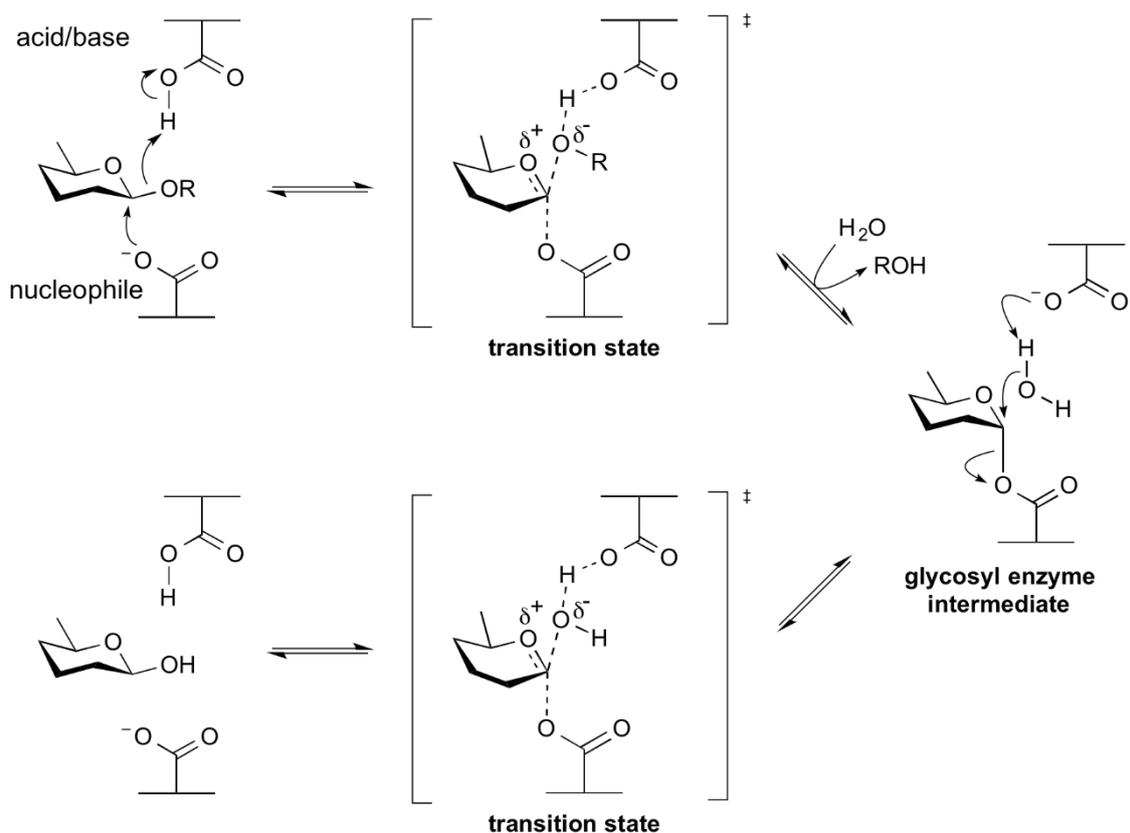


Figure 1.16 Retaining enzymes mechanism. Schematic diagram of the glycoside hydrolase retaining mechanism. Taken from www.cazypedia.org

The retaining mechanism utilises a general acid/base and a nucleophile as the catalytic machinery required for efficient catalysis. During the reaction the general acid/base requires protonation at the optimum pH of the enzyme, usually pH 7.0, despite acidic amino acids possessing a pKa of approx. 4.0. To compensate for this disparity the general acid/base is usually located in an apolar region of the enzyme around 5.5 Å away from the catalytic nucleophile, causing protonation of the general acid base to become more favourable at a physiological pH (Davies and Henrissat, 1995).

The first step of the retaining mechanism (for a D- β - or L- α -glycone involves glycosylation of the enzyme through an oxocarbenium ion-like transition state. The catalytic nucleophile attacks the anomeric carbon, C1, from the α -face of the sugar while the general acid/base, here acting as an acid, donates a proton to the leaving group, which is typically a sugar. Catalysis then proceeds

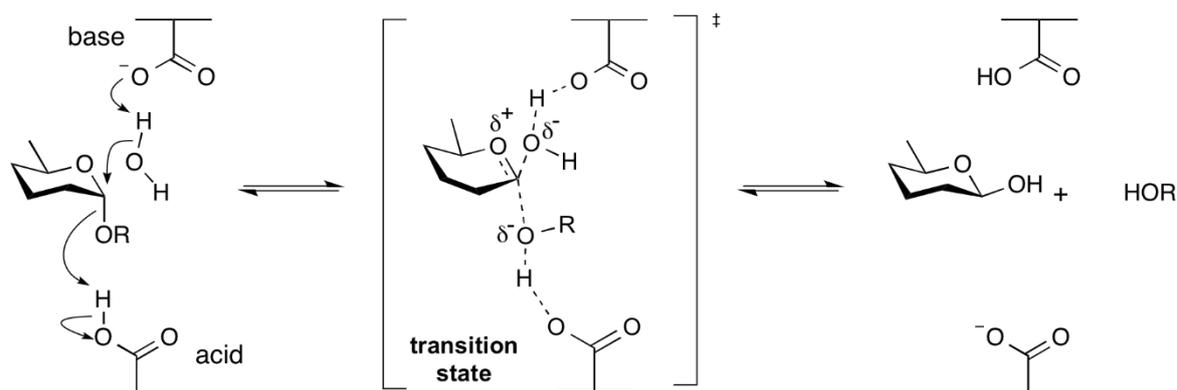
through an oxocarbenium ion-like transition state, promoting development of a positive charge at the anomeric carbon. The sugar is distorted from the relaxed 4C_1 chair conformation into one of four possible conformations (4H_3 , 3H_4 , ${}^{2,5}B$ or $B_{2,5}$), any of which place C1, C2, C5 and the endocyclic ring oxygen in the same plane, promoting sharing of the positive charge between C1 and the endocyclic ring oxygen, stabilising the transition state. The positive charge facilitates attack by the catalytic nucleophile leading to the formation of a glycosyl-enzyme covalent intermediate, breaking the target glycosidic bond in the process (Davies and Henrissat, 1995).

The second step involves deglycosylation of the glycosyl-enzyme intermediate, releasing product from the enzyme active site to make way for a new substrate. In this step the general acid/base acts as a base accepting a proton from an incoming water molecule, promoting the formation of a hydroxyl ion, which attacks C1 from the β -face of the sugar. This action cleaves the α -glycosyl-enzyme linkage, releasing the sugar molecule from the enzyme with a β -orientation hydroxyl group as the final reaction product. Cleavage of the α -bond occurs via an oxocarbenium ion-like intermediate just as in the glycosylation step (Davies and Henrissat, 1995). If another sugar molecule is present in place of the incoming water molecule a transglycosylation reaction occurs, generating a product with a degree of polymerisation the sum of the both substrate glycans, a phenomenon currently exploited to generate GOS (mimic oligosaccharides in breast milk), which are used as prebiotics (Rodriguez-Colinas *et al.*, 2014).

1.4.2.2 Inverting Mechanism

In contrast to the retaining mechanism, the inverting mechanism makes use of two amino acids as a separate general acid and base for proton donation and reception (Figure 1.17), respectively. These catalytic amino acids are typically 10 Å apart (Davies and Henrissat, 1995). Some enzymes which employ the inverting mechanism, including those in GH43, possess a third catalytic amino acid, a pKa modulator which along with the apolar environment surrounding the general acid ensure protonation (Nurizzo *et al.*, 2002).

Inverting mechanism for an α -glycosidase:



Inverting mechanism for a β -glycosidase:

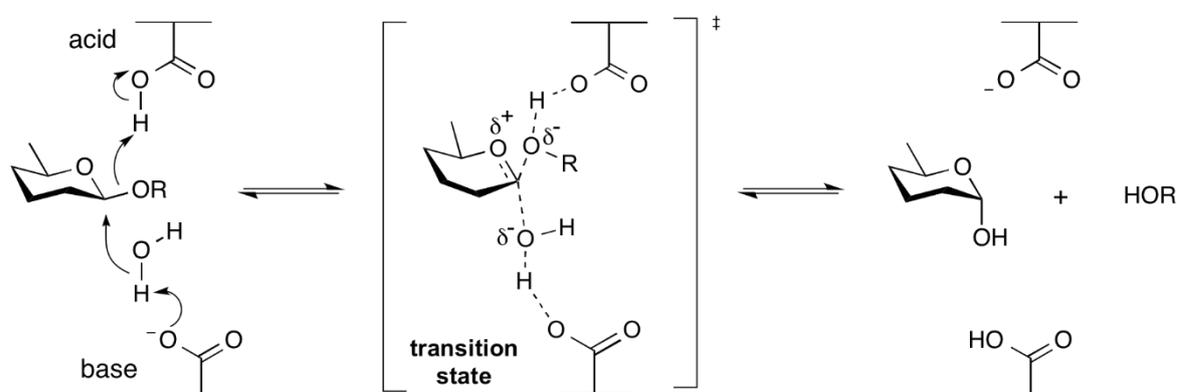


Figure 1.17 Inverting enzymes mechanism. Schematic diagram of the glycoside hydrolase inverting mechanism. Taken from www.cazypedia.org

The inverting mechanism involves just one step to mediate glycosidic bond cleavage. The general base accepts a proton from an incoming water molecule generating a hydroxyl ion in the process. The hydroxyl ion attacks the anomeric carbon, C1, from the β -face of the sugar as the general acid donates a proton to the glycosidic oxygen to assist departure of the leaving group. Similar to the retaining mechanism the reaction must progress through an oxocarbenium ion-like transition state distorting the 4C_1 chair conformation to ensure C1 and the endocyclic oxygen are in the same plane and thus can share the positive charge at C1. The development of the positive charge promotes nucleophilic attack by the nucleophilic water (Davies and Henrissat, 1995).

1.4.3 Subsite Nomenclature

A single subsite of a glycoside hydrolase constitutes all amino acids which interact with a single sugar molecule of the substrate. Subsites either side of the scissile glycosidic bond (the bond between subsites -1 and +1) are given positive or negative values. The subsites towards the non-reducing end of the scissile glycosidic bond are given negative values while those on the reducing end have positive values (Figure 1.18). The -1 subsite houses the catalytic machinery of the enzyme, and the sugar occupying this subsite becomes distorted during the oxocarbenium ion-like transition state (Davies and Henrissat, 1995).

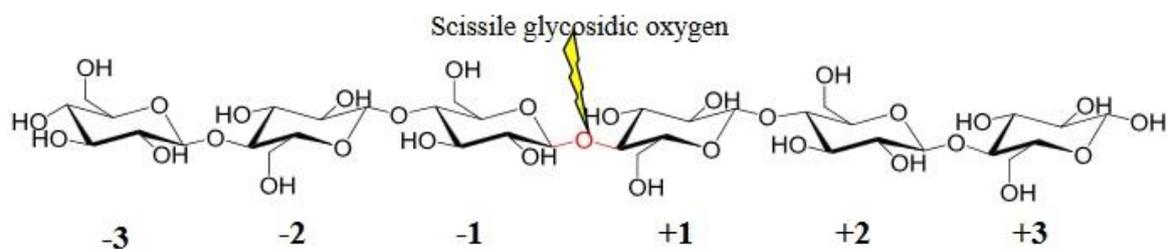


Figure 1.18 Schematic diagram showing subsite nomenclature of glycoside hydrolases. Scissile glycosidic oxygen shown in red. Subsite numbers are given under each sugar. Taken from Cartmell (2010).

1.5 Glycan Utilisation Systems

Glycans are the primary growth substrate for gut bacteria, either from the human diet or the host epithelium. The glycan systems currently identified share common features, CAZymes, to deconstruct polysaccharides, glycan binding proteins to capture polysaccharides and oligosaccharides, transport systems, to import oligosaccharides across cell membrane(s), and a regulator that binds an intermediate breakdown product to control expression of the utilisation system (Cockburn and Koropatkin, 2016).

1.5.1 *Bacteroides* Glycan Utilisation

Bacteroides species, which are well represented in the HGM, devote considerable genomic coding capacity to the degradation and utilisation of glycans. The first identified glycan utilisation system of *B. thetaiotaomicron* targeted starch, the *sus* locus (D'Elia and Salyers, 1996; Shipman *et al.*, 2000; Cho *et al.*, 2001). This archetypal system was used to identify further glycan utilisation systems sharing similar features, namely adjacent genes *susC* and *susD* homologues encoding the outer membrane transport system alongside CAZymes. *B. thetaiotaomicron* was found to encode 88 polysaccharide utilisation loci (PULs), with each PUL averaging 10 genes (Martens *et al.*, 2008; Martens *et al.*, 2011). A major challenge of glycan utilisation is coupling degradation with transport; being a Gram negative bacterium *Bacteroides* must transport glycans over two membranes and a periplasmic compartment (Martens *et al.*, 2009). A simplified diagram of a *Bacteroides* glycan utilisation system based on the *sus*-paradigm is shown in Figure 1.19.

1.5.1.1 Extracellular Glycan Degradation

With the exception of inulin [at least low molecular weight inulin from chicory (Rakoff-Nahoum *et al.*, 2016)], in the overwhelming majority of *Bacteroides* glycan utilisation systems studied to date, a degree of substrate degradation is required at the cell surface prior to transportation into the periplasm (Sonnenburg *et al.*, 2010; Larsbrink *et al.*, 2014; Cuskin *et al.*, 2015; Rogowski *et al.*, 2015). At first surface degradation of glycans appears inefficient with high risk of product loss to competitors, however this appears to be an intentional strategy employed to maintain species diversity in the HGM (Rakoff-Nahoum *et al.*, 2016). This is further consolidated by secretion of CAZymes in outer membrane vesicles (OMVs). Among the enzymes purified from OMVs was BT_1760 (Elhenawy *et al.*, 2014), an enzyme shown to degrade levan prior to import of the resulting oligosaccharides (Sonnenburg *et al.*, 2010). Contrasting with this, *B. thetaiotaomicron* employs a selfish mechanism for yeast mannan utilisation. Surface endo- α 1,6-mannanases generate large oligosaccharides, minimising extracellular metabolism and maximising substrate available to the

periplasmic exo- α 1,6-mannosidases. This renders *B. thetaiotaomicron* unable to support mannose-/mannan-utilising strains of *Bacteroides* on intact *Saccharomyces cerevisiae* α -mannan (Cuskin *et al.*, 2015).

Initial prediction of surface localisation is performed by signal sequence analysis based on conserved features of *E. coli* proteins. This analysis is not always reliable. A more reliable technique involves using fluorescence microscopy to identify labelled antibodies specific for the suspected surface protein (Cuskin *et al.*, 2015; karunatilaka 2014 *et al.*, 2014; Larsbrink *et al.*, 2014; Rogowski *et al.*, 2015).

1.5.1.2 Surface Glycan Binding Proteins and the SusCD-homologue Complex

Surface glycan binding proteins (SGBPs) are not always encoded by *Bacteroides* PULs. As SGBPs show little to no sequence conservation they can be difficult to identify, although the gene encoding the SGBP is typically found adjacent to the *susD*-homologue. Functionally, SGBPs bind the target glycan at the cell surface increasing local substrate concentration to modulate the activity of surface CAZymes. The prototypic *sus*-system encodes two SGBPs, SusE and SusF, each expressed at the cell surface. Both SusE and SusF contribute to surface binding of starch, although they have been found to be non-essential in the utilisation of soluble starch (Cho *et al.*, 2001). Although, when grown on starch alongside wild type *B. thetaiotaomicron*, a mutant lacking SusE is at a competitive disadvantage (Koropatkin *et al.*, 2008).

SGBPs operate alongside a SusCD-homologue complex to bind and transport glycans. The SusD-like protein is thought to be primarily a binding protein with the capability to channel captured ligand into the SusC-like protein. Despite showing lower affinity for starch than SusE or SusF, SusD has been shown to be vital to starch utilisation through genetic knockouts (Reeves *et al.*, 1997; Cho *et al.*, 2001).

Currently there is little published data regarding the SusC protein and its homologues, despite being essential for glycan utilisation in *Bacteroides*. The closest characterised homologue of SusC is FepA, a TonB-dependent iron transporter expressed by *E. coli*. The crystal structure of FepA, showed a 22-stranded β -barrel structure creating a channel that is blocked at the periplasmic end by a plug domain, which extends into the channel. The periplasmic end of the protein interacts with TonB, a protein that provides the transporter with energy to open the channel (Buchanan *et al.*, 1999; Jordan *et al.*, 2013).

1.5.1.3 Periplasmic glycan degradation

Once in the periplasm, oligosaccharides are sequestered away from other microbes which occupy the same ecological niche. Here degradation can be completed by periplasmic CAZymes, typically exo-acting, releasing monosaccharides. These monosaccharides are then transported into the cytoplasm entering fermentation pathways resulting in energy for the cell and SCFAs, which are secreted into the gut lumen (D'Elia and Salyers, 1996; Martens *et al.*, 2009) and can be metabolised by the host or other members of the HGM.

1.5.1.4 PUL regulation

Typically, *Bacteroides* glycan utilisation systems include one of two regulatory systems, hybrid two-component systems (HTCS) or extracytoplasmic function (ECF) sigma factor/anti-sigma factor systems. These systems sense a signal, usually in the form of short oligosaccharides, to upregulate the associated enzymes, binding and transport proteins, from the basal expression level employed by the bacterium to survey the environment for glycans (Martens *et al.*, 2009).

Many PULs are regulated by an HTCS. The *B. thetaiotaomicron* genome encodes 32 HTCSs, 23 of which are directly upstream of PULs (Xu *et al.*, 2003). The prototypic two-component system (TCS) consist of two core-modules, a membrane bound sensor histidine kinase (HK) and a cytoplasmic response regulator (RR), which operate together to initiate signal transduction upon binding of a

chemically defined ligand. The HK domain recognises a defined ligand resulting in autophosphorylation of conserved histidine residues. The phosphoryl groups are then transferred to conserved aspartate residues in the N-terminal of the corresponding RR domain, resulting in activation of the associated output module, which mediates the appropriate cellular response (Stock *et al.*, 2000). HTCSs comprise of a single polypeptide which incorporates all the features of a typical bacterial TCS signal transducer, possessing sensor histidine kinase, phosphoreceptor and DNA binding domains (Stock *et al.*, 2000). The N-terminal sensor kinase domain, which extends into the periplasm where it binds activation ligands usually an oligosaccharide, is immediately followed by five predicted transmembrane domains. The C-terminal domain protrudes into the cytoplasm, and possessing a histidine kinase and phosphoreceptor domains, along with an AraC helix-turn-helix domain, responsible for DNA binding (Lowe *et al.*, 2012). The DNA binding domain may be released upon activation allowing DNA binding (Miyazaki *et al.*, 2003). The sensor domain of the HTCS of the heparin PUL, Bt4663, displays a β -propeller fold followed by a Y_Y_Y domain of unknown function. The sensor domain forms a dimer at the periplasmic face of the inner membrane. When the sensor domain is bound to its activating ligand the dimerization face is altered bringing the two C-termini closer together, reminiscent of a scissor blade closing. The closing motion is most likely transmitted via the transmembrane helices to trigger autophosphorylation in the cytoplasmic histidine kinase domains of the dimers (Lowe *et al.*, 2012). This mechanism contrasts with that of canonical sensor histidine kinases such as the LuxPQ chemoreceptor. Ligand binding in these LuxPQ-like TCS trigger a piston-like mechanism for signal transduction (Falke and Erbse, 2009).

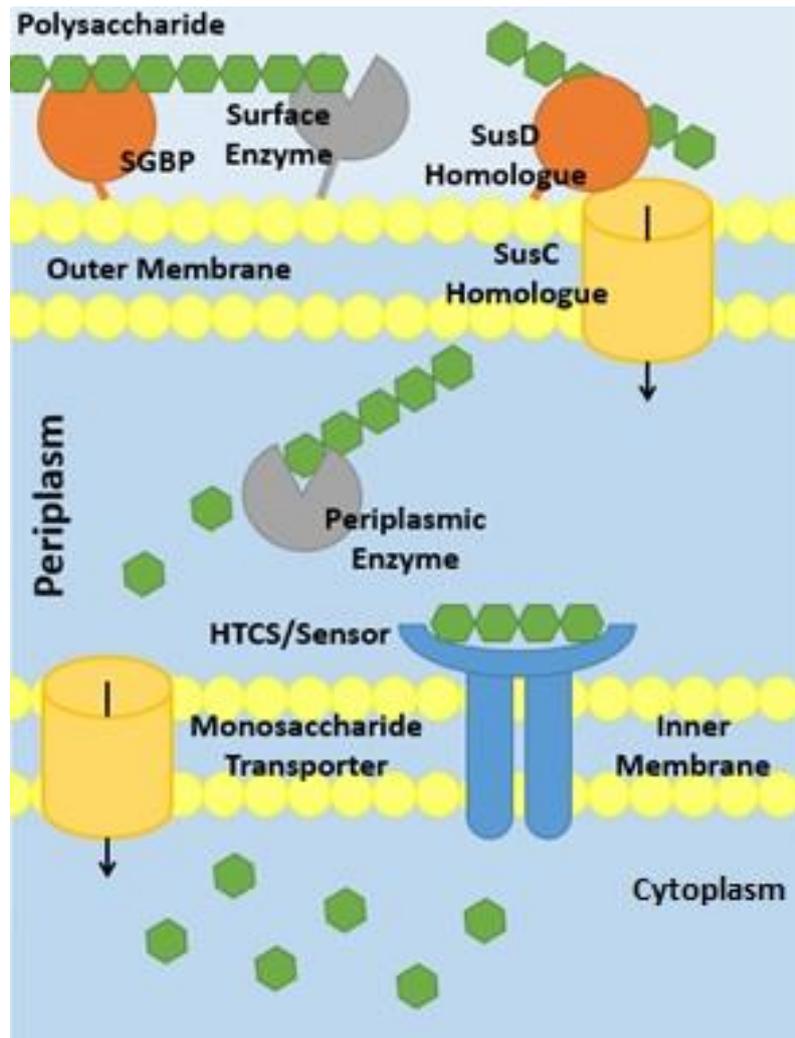


Figure 1.19 Simplified *Bacteroides* utilisation system. A generalised *Bacteroides* utilisation system, numbers of each component may vary depending on specific system. The polysaccharide is bound to the surface by the action of a surface glycan binding protein (SGBP) where it is degraded to oligosaccharides by lipoprotein glycoside hydrolases. The oligosaccharides are bound by a SusCD homologue complex and imported through the SusC homologue into the periplasm. In the periplasm oligosaccharides are degraded to monosaccharides by periplasmic glycoside hydrolases and imported into the cytoplasm to enter fermentation pathways.

The second regulatory strategy employed by *Bacteroides* PULs, the ECF sigma/anti-sigma factor system, is usually found in loci directed against host O-glycans rather than dietary polysaccharides (Martens *et al.* 2008). The inner membrane anti-sigma factor sequesters the ECF sigma factor until a signal is received indicating its target ligand is available, at which point the sigma factor is released to upregulate the expression of the associated PUL (Martens *et al.*, 2008).

Not all PULs are given the same priority by *Bacteroides*. Some are considered more important, for example, in the presence of glucose transcription of almost all PULs are repressed (Rogers *et al.*, 2013). Among polysaccharides, galactan and homogalacturonan appear to be higher in utilisation priority than other more complex fractions of pectin (RGI backbone for example). *Bacteroides* can rapidly respond to multiple glycans in a mixture, altering expression profile in a matter of minutes due to changes in the glycans available in the environment (Rogers *et al.*, 2013).

Recent studies have shown new regulation systems based on antisense small RNA (sRNA) responsible for repression of PUL expression in *B. fragilis* (Cao *et al.*, 2014). A 125 nt sRNA sequence has been found in an intergenic region upstream of the *susC* homologue of the *B. fragilis* N-glycan utilisation locus. This sequence was found in other *B. fragilis* PULs, the majority of which include sigma/anti-sigma regulation systems (Cao *et al.*, 2014). Mutation of the sRNA regions lead to a slightly increased expression of the PUL, while over expression of the sRNA lead to a 400-fold repression of the locus in the presence of the target glycan (Cao *et al.*, 2014). Similar regions have been identified in *B. theattotaomicron*, indicating the use of sRNA repressions systems in PULs is not just limited to one species (Cao *et al.*, 2014).

1.5.2 *Bifidobacterium* Glycan Utilisation

Bifidobacterium utilisation profiles show a preference for oligosaccharides over polysaccharides (Mei *et al.* 2011). Unlike *Bacteroides*, *Bifidobacteria* possess a single membrane with a peptidoglycan cell wall. *Bi. longum* was shown to possess 10 ABC transporter systems responsible for the uptake of various carbohydrates, along with glycoside hydrolases (Lorca *et al.*, 2007). A simplified *Bifidobacterium* glycan utilisation system is shown in Figure 1.20

1.5.2.1 Extracellular Solute Binding Proteins and ABC-Transporter

The vast majority of transport proteins thought to be involved in glycan utilisation in *Bifidobacterium* fall into the ATP-Binding Cassette (ABC)-Transporter classification. ABC-transporters are ubiquitous membrane protein complexes, using ATP hydrolysis to drive import of a wide range of solutes across biological membranes (Saier, 2000). Most ABC transporter complexes demonstrate a modular design, with two transmembrane domains and two cytosolic nucleotide-binding domains. Substrate specificity of these transporters lies in the transmembrane domains which show no sequence homology between transporters. These transmembrane domains can vary in the number of helices, between ABC-transporters. The hydrolysis of ATP to ADP, which drives solute translocation occurs in the cytoplasm by the action of the nucleotide binding domains. These domains contain a conserved LSGG(N)QQ signature motif which is characteristic of ATPases associated with ABC transporters (Saier, 2000). *Bifidobacteria* have been shown to encode a number of extracellular solute binding proteins (ESBP) along with the permeases which make up the ABC-transporter complex (Ejby *et al.*, 2013; Ejby *et al.*, 2016). The glycan-specific ESBPs characterised to date are high affinity oligosaccharide binding proteins, which bind considerably tighter to their ligands than any of the characterised SGBPs of *Bacteroides*. A recent study showed in mixed cultures *Bi. animalis* was capable of outcompeting *B. ovatus* on raffinose, and speculate that the presence of the ESBP-ABC-transporter system was responsible (Ejby *et al.*, 2016). Previous to this, the ESBP responsible for binding xylooligosaccharides (XOS) for delivery to the associated ABC-transporter in *Bi. animalis* was able to bind arabinoxylooligosaccharides (AXOS) with similar affinity to XOS, indicating a high tolerance of backbone substitutions in the protein (Ejby *et al.*, 2013). This affinity for AXOS/XOS was not shared for xylan polysaccharides (Ejby *et al.*, 2013), demonstrating a strong preference of these glycan utilisation systems for oligosaccharides. This could be interpreted as a method of co-existence with the polysaccharide utilising gut bacteria, which target longer substrates while *Bifidobacterium* has formed a niche of its own by targeting only oligosaccharides, partially avoiding competition (Ejby *et al.*, 2013, Ejby *et al.*, 2016).

1.5.2.2 *Bifidobacterium* Glycoside Hydrolases

Encoded alongside ABC-transporter components and ESBP glycan capture proteins are CAZymes, typically exo-acting glycoside hydrolase enzymes which target the non-reducing end of oligosaccharides releasing monosaccharide products. Although exo-acting CAZymes are most prevalent in *Bifidobacteria* some endo-acting enzymes are also present (van den Broek *et al.*, 2008).

Bi. longum possesses a membrane anchored modular putative glycoside hydrolase with the domain organisation, SignP-CBM22-GH43-CBM6-TmD, where SignP is a secretion signal, TmD indicates a trans-membrane domain, while CBM22, CBM6 and GH43 are protein modules located in these three CAZy families. This is an unusual example of a surface *Bifidobacterium* protein. Although this enzyme has not been characterised, it may possess exo-activity is common among GH43 enzymes (van den Broek *et al.*, 2008). This enzyme may act as an arabinofuranosidase removing arabinose side chains from AXOS or branched arabinooligosaccharides prior to transport via the associated ABC-transporter. *Bi. adolescentis* possesses a locus likely targeting XOS. The bacterium grows well on XOS but is unable to utilise xylans. During growth on XOS, xylose accumulation was observed, indicating that surface xylosidase(s) plays a role in the utilisation of these oligosaccharides in *Bi. adolescentis*, despite all identified potential xylanases/xylosidases expressed by *Bi. adolescentis* lacking any known secretion signal peptide (Amaretti *et al.*, 2013). This may indicate *Bifidobacteria* use secretion signals that are not recognised by current secretion prediction programmes. These XOS-utilising *Bifidobacteria* may secrete a number of XOS/AXOS targeting enzymes although current data suggest these are deployed to remove side chains, possibly to allow better recognition by the associated ESBP and downstream degradation apparatus.

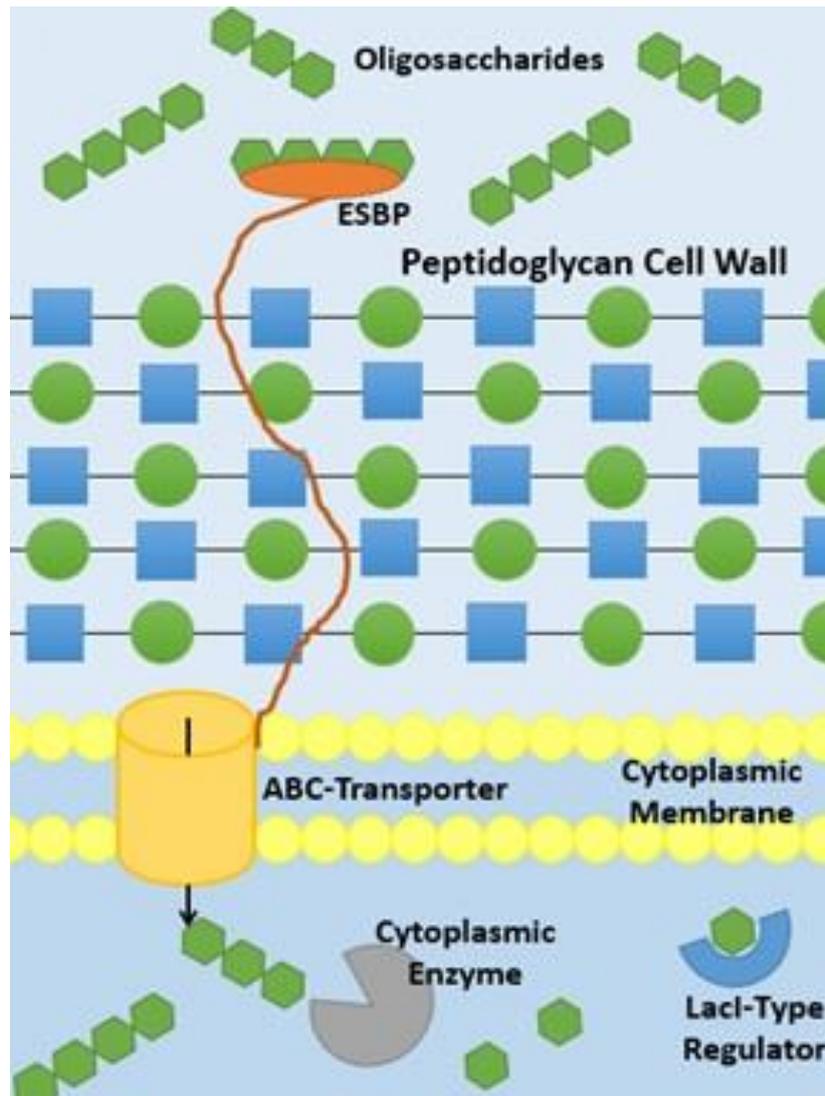


Figure 1.20 Simplified *Bifidobacterium* glycan utilisation system. Generally *Bifidobacterium* are oligosaccharide utilisers, however, some have been shown to possess endo-acting surface attached glycoside hydrolases. Exogenous or surface produced oligosaccharides are bound by extracellular solute binding proteins (ESBP) and delivered to the associated ABC-transporter complex and imported into the cytoplasm.

Bi. breve is able to utilise the polysaccharide galactan by use of an extracellular GH53 β -galactanase and a cytoplasmic β -galactosidase. Galactan encountered at the cell surface is degraded in an endo-fashion, generating galactotriose. Once imported into the cytoplasm, galactotriose is degraded to galactose by the exo-action of the cytoplasmic β -galactosidase (O'Connell Motherway *et al.*, 2011). Another operon found in *Bi. breve* was identified and shown to target fructose-containing glycans. This operon only included one CAzyme, a GH32 β -fructofuranosidase believed to be located

intracellularly. The exo-acting enzyme released glucose from sucrose and cleaved the β 2,1-bond between fructose of oligosaccharides. The activity of this enzyme along with being the only CAzyme in the locus shows the system targets FOS or the shorter fractions of inulin specifically (Ryan *et al.*, 2005). The ability of *Bifidobacterium* to utilise polysaccharides appears to correlate with lower complexity of the substrate. Short Inulin and galactan each have one unique linkage in the polysaccharide backbone, β 2,1- and β 1,4- bonds, respectively, requiring a minimum of two enzymes to degrade these polymers to monosaccharides. *Bifidobacterium* appears to leave degradation of more complex substrates to other members of the HGM but has developed systems for scavenging oligosaccharides released during utilisation of these polysaccharides.

1.5.2.3 Regulation

Similar to the *Bacteroides* utilisation systems, the expression of the corresponding *Bifidobacterium* apparatus is substrate dependent, mediated by LacI-type regulators (Rodionov *et al.*, 2001; Rodionov, 2007). *Bi. breve* utilisation of ribose was shown to be under the control of RbsR, a LacI-type regulator, which causes repression of the ribose utilisation system in the absence of the target substrate (Pokusaeva *et al.*, 2010). Another LacI-type regulator, GalR, was shown to regulate the galactan utilisation locus of *Bi. breve*. The repressor GalR binds to two sites in the promoter of *galA*, a GH53 galactanase, and *galC*, a component of the ABC-transporter. These binding regions of DNA were found to be 9 bp inverted repeats overlapping the -10 and -35 promoter recognition sequences. Interestingly, GalR remained bound to the promoter regions of *galA* and *galC* in the presence of lactose and galactose. GalR only disassociated when galactobiose was present, indicating this as the activating molecule (O'Connell Motherway *et al.*, 2011).

Bi. longum possess a pair of ABC-transporters shown to import xylose; both were found to be downregulated in the presence of glucose, indicating a hierarchy of glycan utilisation is present in *Bifidobacteria* with glucose being a more valuable substrate than xylose (Lee and O'Sullivan, 2010;

Liu *et al.*, 2011; Pokusaeva *et al.*, 2011). Thus classical catabolite repression appears to operate at least in *Bi. longum*.

1.5.3 Glycan Utilisation by other Members of the HGM

Although *Bifidobacteria* and *Bacteroides* are substantial glycan utilisers in the HGM, other bacteria employ glycan utilisation systems to forage for energy rich polysaccharides. As mentioned above *R. bromii* shows exceptional activity against resistant starch and 15 of its 21 glycoside hydrolases belong to GH13, a family typified by activity on starch. Four of these proteins are predicted to include dockerin modules, which along with cohesin modules mediates formation of a multi-enzyme complex, the amylosome (Figure 1.21). The amylosome is a large protein complex consisting of scaffold proteins (Sca2, Sca3, Sca4) each bearing cohesin modules and catalytic proteins (Amy4, Amy9, Amy10 and Amy12) with dockerin modules. Amy4 contains both cohesin and dockerin modules, allowing it to act as both a catalytic and scaffold protein. Similar to cellulosomes of cellulolytic *Ruminococcus spp* (Ben David *et al.*, 2015), the amylosome forms a large extracellular complex targeting a specific glycan, which may or may not be membrane anchored. In contrast to the substrate inducible glycan utilisation systems discussed above, the amylosome was found to be constitutively expressed even when starch is absent (Ze *et al.*, 2012; Ze *et al.*, 2015). Interestingly, when in mixed culture with other bacteria that can use soluble starch, *R. bromii* enabled these organisms to utilize the resistant form of this α -glucan. *R. bromii* increases accessibility of resistant starch for the large number of starch utilising bacteria in the gut (Ze *et al.*, 2012; Ze *et al.*, 2015).

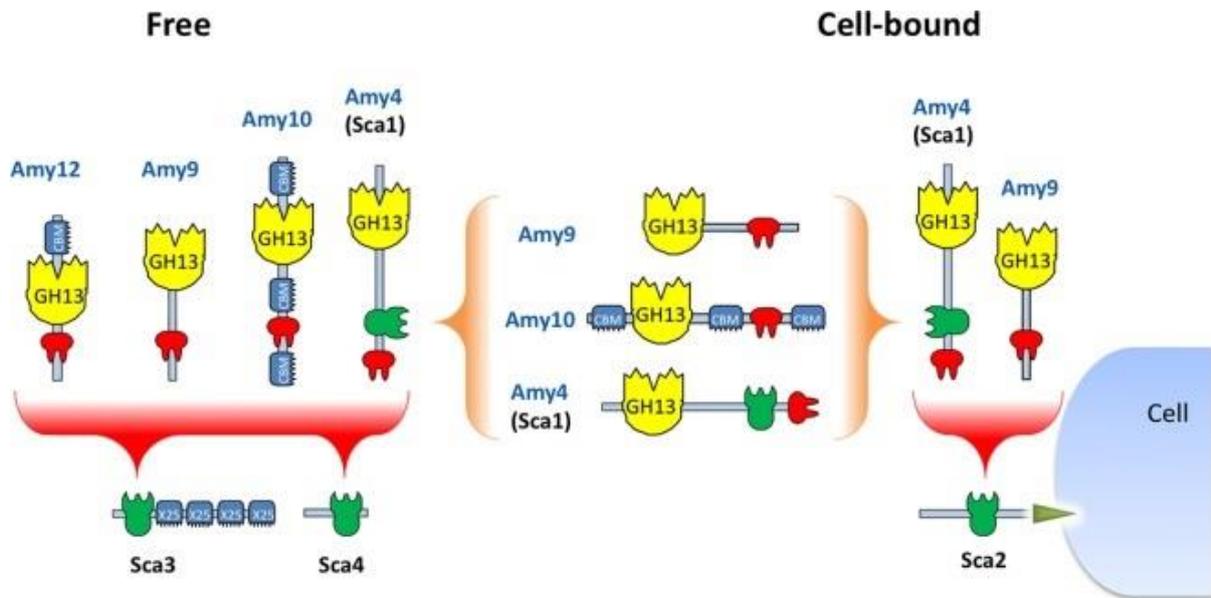


Figure 1.21 Model of the *R. bromii* multi-enzyme amylosome complex. Potential interactions of the amylosome between scaffold protein and catalytic proteins based on observed recombinant cohesin-dockerin interactions. The scaffold proteins Sca3 and Sca4 are likely to be secreted while Sca2 is most likely membrane anchored. Amy4, Amy9, Amy10 and Amy12 each possess GH13 catalytic domains and dockerin domains. Amy4 also contains a cohesin domain that, by binding to the dockerin domains of other proteins, is integral to the assembly of the multienzyme complex. Sca3, Amy10 and Amy12 each have non-catalytic carbohydrate binding modules (CBM or X25) that bind glycans increasing local concentration of substrate. Taken from Ze *et al.* (2016).

Another resistant starch utilising bacterium which employs a different strategy of glycan utilisation is *E. rectale*. This bacterium expresses two GH13 enzymes and three ABC glycan binding proteins in response to starch exposure. The first of the GH13 enzymes is a large modular lipoprotein attached to the cell surface with several CBMs (Figure 1.22). These CBM domains allow *E. rectale* to anchor itself onto starch granules in the gut and begin degrading the polysaccharide into maltooligosaccharides to be captured and imported into the cytoplasm by the associated ABC transporter (Figure 1.22). The second GH13 enzyme is membrane attached like the first, however it is unclear as to which face of the membrane it localises (Figure 1.22). Other amylases expressed by *E. rectale* are predicted to contribute to maltooligosaccharide breakdown in the cytoplasm (Cockburn *et al.*, 2015).

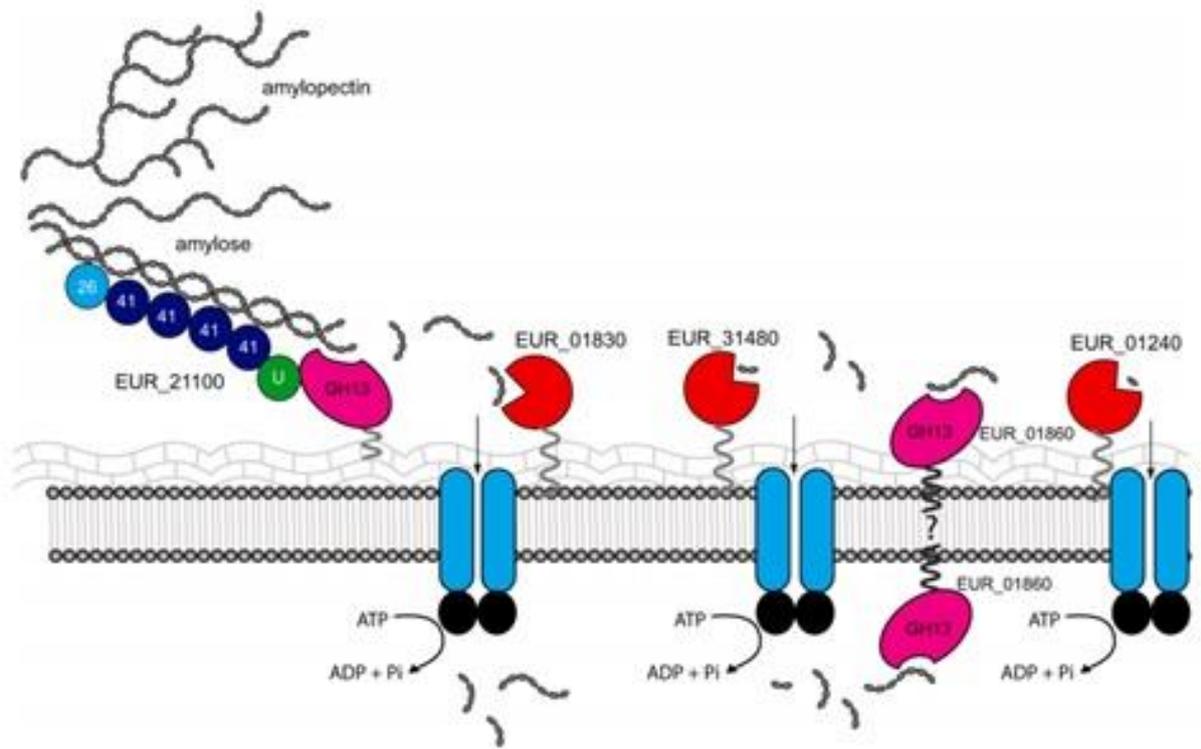


Figure 1.22 Model of *E. rectale* starch utilisation. The multi domain GH13 enzyme (pink) binds starch at the cell surface with five CBMs (blue). The polysaccharide is degraded into maltooligosaccharides by the action of the GH13 catalytic domain of the same protein. The second GH13 enzyme may be present on the surface or inner leaflet of the cell membrane. Three solute binding proteins bind and channel resulting maltooligosaccharides to their respective ABC transporters for import into the cytoplasm. Taken from Cockburn *et al.* (2015).

1.6 Objectives of this Study

The primary focus of the research in this thesis is to characterise the pectin degradation and utilisation systems of *Bacteroides thetaiotaomicron* (Chapter 4) and the xylan utilisation system of *Bacteroides ovatus* (Chapter 3). This involved biochemical characterisation of various glycoside hydrolases and glycan binding proteins associated with the utilisation systems. Microbiological methods were also used to introduce mutations in components of the glycan utilisation systems to assess the importance of individual proteins. During investigation of surface enzymes the potential for glycan cross-feeding was discovered and explored through the co-culture of *Bacteroides* with

another bacterium unable to utilise the glycan. Cross-feeding was explored during *Bacteroides* utilisation of xylans (Chapter 3), pectins (Chapter 4) and fructans (Chapter 5).

Previous work has shown *B. thetaiotaomicron* and *B. ovatus* to be capable of utilising a wide range of glycans. Transcriptomic data identified loci involved in utilisation of glycans. To date only a few *B. thetaiotaomicron* glycan utilisation systems have been characterised, each system revealing some novel enzyme activity, binding protein or strategy for glycan degradation or transport. By characterising more of these systems valuable insight can be gained into glycan degradation and may reveal new strategies that can be used to selectively enrich for *Bacteroides* in the HGM.

Glycan cross-feeding is a recently identified bacterial interaction within the HGM, where a glycan utilising bacterium increases substrate accessibility for a second bacterium which cannot use the glycan in its original form. This interaction may allow identification of novel glycan-based prebiotics. Specific objectives are listed at the beginning of each results chapter.

Chapter 2: Materials and Methods

2.1 Bacterial strains and mutants

The bacterial strains and mutants used are listed in Table 2.1, with mutants generated during the course of this study in bold.

Name	Genotype features	Description
BL21 (DE3)	<i>F</i> – <i>ompT gal dcm lon hsdSB</i> (rB-mB-) λ (DE3 [<i>lacI lacUV5-T7 gene 1 ind1 sam7 nin5</i>])	<i>E. coli</i> strain optimised for protein expression using a T7 promoter. Routinely used to over-express recombinant proteins. (Studier & Moffatt, 1986)
Tuner (DE3)	<i>F</i> – <i>ompT hsdS_B</i> (r _B ⁻ m _B ⁻) <i>gal dcm lacY1</i> (DE3)	As BL21 with the addition of a lac permease mutation to allow uniform diffusion of IPTG across cells, establishing a linear relationship between IPTG concentration and expression levels. Tuner cells also carry a <i>lacZ</i> mutation. These cells were used to express enzymes suspected to have β -galactosidase activity to avoid false positive results from contamination with genomic <i>lacZ</i> (Novagen)
Top10	<i>F</i> - <i>mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) ϕ 80 <i>lacZ</i> Δ M15 <i>ΔlacX74nupGreca1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-</i>	Routinely used for plasmid propagation and cloning. (Invitrogen)
CC118 λ -pir	<i>Δ(ara-leu) araD ΔlacX74 galE galK phoA20 thi-1 rpsE rpoB argE (Am) recA1 λ pir</i>	Used for plasmid propagation and cloning (pExchange <i>tdk</i> only) (Herrero <i>et al.</i> , 1990)
S17.1 λ -pir	<i>hsdR recA pro</i> RP4-2 (Tc:: <i>Mu</i> ;KmTn7)	Conjugation of pExchange <i>tdk</i> plasmids from this strain to <i>Bacteroides ovatus</i> (Skorupski & Taylor, 1996)
<i>B. ovatus</i>	Wild-Type ATCC 8483	DSM-1896, Type strain
<i>B. ovatus</i> Δ <i>tdk</i>	Δ <i>tdk</i>	<i>B. ovatus</i> ATCC 8483 lacking thimadine kinase. Used to generate genomic mutants or gene deletions in <i>Bacteroides ovatus</i> through FuDR selection (Chapter 2.9.4). Provided by Martens Lab, Center for

		Microbial Systems, University of Michigan.
ΔGH10	<i>Δbacova_04390</i>	<i>B. ovatus Δtdk</i> harbouring an inactive copy of <i>bacova_04390</i>
ΔGH98	<i>Δbacova_03433</i>	<i>B. ovatus Δtdk</i> harbouring an inactive copy of <i>bacova_03433</i>
<i>B. ovatus Tag 1</i>	<i>Tag 1</i> insertion into ATT site 1	<i>B. ovatus Δtdk</i> harbouring a unique sequence termed <i>tag1</i>
<i>B. ovatus Tag11</i>	<i>Tag 11</i> insertion into ATT site 1	<i>B. ovatus Δtdk</i> harbouring a unique sequence termed <i>tag11</i>
ΔGH98tag11	<i>Δbacova_03433</i> <i>Tag 11</i> insertion into ATT site 1	<i>Δbacova_03433</i> harbouring a unique sequence termed <i>tag11</i>
ΔGH91	<i>Δbacova_04502</i>	<i>B. ovatus</i> with inactive <i>bacova_04502</i>. Provided by Sarah Shapiro (Shapiro 2015).
ΔGH91 tag11	<i>Δbacova_04502</i>, <i>Tag11</i> insertion into att site 2	<i>B. ovatus</i> with inactive <i>bacova_04502</i>. Provided by Sarah Shapiro (Shapiro 2015). With Tag11 insetion in ATT site 2.
<i>B. thetaiotaomicron 8764</i>	Wild-Type	Provided by Martens Lab, Center for Microbial Systems, University of Michigan.
<i>B. thetaiotaomicron Δtdk</i>	<i>Δtdk</i>	Used to generate genomic mutants or gene deletions in <i>Bacteroides ovatus</i> through FuDR selection (Chapter 2.9.4). Provided by Martens Lab, Center for Microbial Systems, University of Michigan.
<i>B. thetaiotaomicron tag1</i>	<i>tag1</i> insertion into ATT site 2	<i>B. thetaiotaomicron</i> Tag1 insertion into ATT site 1
<i>B. thetaiotaomicron tag11</i>	<i>tag11</i> insertion into ATT site 2	<i>B. thetaiotaomicron</i> Tag11 insertion into ATT site 1
ΔBt_4667	<i>Δbt_4667</i>	<i>B. thetaiotaomicron</i> with inactivation of <i>bt_4667</i>
ΔBt_4668	<i>Δbt_4668</i>	<i>B. thetaiotaomicron</i> with inactivation of <i>bt_4668</i>
ΔBt_4668tag11	<i>Δbt_4668</i> <i>tag11</i> insertion into ATT site 2	<i>B. thetaiotaomicron</i> with inactivation of <i>bt_4668</i>. With <i>tag11</i> insertion into ATT site 1
ΔBt_4669	<i>Δbt_4669</i>	<i>B. thetaiotaomicron</i> with deletion of <i>bt_4669</i>
ΔBt_4670	<i>Δbt_4670</i>	<i>B. thetaiotaomicron</i> with deletion of <i>bt_4670</i>
ΔBt_4667-73	<i>Δbt_4667 Δbt_4668 Δbt_4669</i> <i>Δbt_4670 Δbt_4671 Δbt_4672</i> <i>Δbt_4673</i>	<i>B. thetaiotaomicron</i> deletion of the entire galactan PUL. Provided by Abbott lab, Lethbridge Research Centre,

		Agriculture and Agri-Food Canada.
$\Delta Bt_{4669-70}$	$\Delta bt_{4669} \Delta bt_{4670}$	<i>B. thetaiotaomicron</i> with deletion of <i>bt_{4669}</i> and <i>bt_{4670}</i>
ΔBt_{0360}	Δbt_{0360}	<i>B. thetaiotaomicron</i> with inactivation of <i>bt_{0360}</i>
ΔBt_{0362}	Δbt_{0362}	<i>B. thetaiotaomicron</i> with deletion of <i>bt_{0362}</i>
ΔBt_{0364}	Δbt_{0364}	<i>B. thetaiotaomicron</i> with deletion of <i>bt_{0364}</i>
ΔBt_{0365}	Δbt_{0365}	<i>B. thetaiotaomicron</i> with deletion of <i>bt_{0365}</i>
ΔBt_{0367}	Δbt_{0367}	<i>B. thetaiotaomicron</i> with inactivation of <i>bt_{0367}</i>
$\Delta Bt_{0360} \Delta Bt_{0367}$	$\Delta bt_{0360} \Delta bt_{0367}$	<i>B. thetaiotaomicron</i> with inactivation of <i>bt_{0360}</i> and <i>bt_{0367}</i>
$\Delta Bt_{0360} \Delta Bt_{0367} tag11$	$\Delta bt_{0360} \Delta bt_{0367} tag11$ insertion into ATT site 2	<i>B. thetaiotaomicron</i> with inactivation of <i>bt_{0360}</i> and <i>bt_{0367}</i> . With <i>tag11</i> insertion into ATT site 1
ΔBt_{4156}	Δbt_{4156}	<i>B. thetaiotaomicron</i> with inactivation of <i>bt_{4156}</i>
ΔBt_{4170}	Δbt_{4170}	<i>B. thetaiotaomicron</i> with deletion of <i>bt_{4170}</i> . Provided by Didier Ndeh for this project (Ndeh & Gilbert, ICaMB, Newcastle University).
ΔBt_{4175}	Δbt_{4175}	<i>B. thetaiotaomicron</i> with deletion of <i>bt_{4175}</i> . Provided by Didier Ndeh for this project (Ndeh & Gilbert, ICaMB, Newcastle University).
$\Delta Bt_{4667-73} \Delta Bt_{4156}$	As $\Delta Bt_{4667-73}$ with Δbt_{4156}	<i>B. thetaiotaomicron</i> deletion of the entire galactan PUL. Provided by Abbott lab, Lethbridge Research Centre, Agriculture and Agri-Food Canada. And inactivation of <i>bt_{4156}</i>
$\Delta Bt_{4667-73} \Delta Bt_{4170}$	As $\Delta Bt_{4667-73}$ with Δbt_{4170}	<i>B. thetaiotaomicron</i> deletion of the entire galactan PUL. Provided by Abbott lab, Lethbridge Research Centre, Agriculture and Agri-Food Canada. And inactivation of <i>bt_{4170}</i>

ΔBt_4667-73ΔBt_4170 tag11	As Δ Bt_4667-73 with Δbt_4170 tag11 insertion into ATT site 2	<i>B. thetaiotaomicron</i> deletion of the entire galactan PUL. Provided by Abbott lab, Lethbridge Research Centre, Agriculture and Agri-Food Canada. And inactivation of bt_4170 . With tag11 insertion into ATT site 1
ΔBt_4667-73ΔBt_4175	As Δ Bt_4667-73 with Δbt_4175	<i>B. thetaiotaomicron</i> deletion of the entire galactan PUL. Provided by Abbott lab, Lethbridge Research Centre, Agriculture and Agri-Food Canada. And inactivation of bt_4175
Δ Bt_1023	Δ bt_1023	<i>B. thetaiotaomicron</i> with deletion of bt_1023 . Provided by Didier Ndeh (Ndeh & Gilbert, ICaMB, Newcastle University).
ΔBt_1023 tag11	Δbt_1023 tag11 insertion into ATT site 1	<i>B. thetaiotaomicron</i> with deletion of bt_1023 . Provided by Didier Ndeh (Ndeh & Gilbert, ICaMB, Newcastle University). With tag11 insertion into ATT site 1
Δ RGII PUL	Δ RGII PUL	<i>B. thetaiotaomicron</i> mutation of the regulator controlling expression of the RGII PUL. Provided by Abbott lab, Lethbridge Research Centre, Agriculture and Agri-Food Canada.
ΔRGII PUL tag11	ΔRGII PUL tag11 insertion into ATT site 1	<i>B. thetaiotaomicron</i> mutation of the regulator controlling expression of the RGII PUL. Provided by Abbott lab, Lethbridge Research Centre, Agriculture and Agri-Food Canada. With tag11 insertion into ATT site 1
<i>Bacteroides vulgatus</i>	Type strain	
<i>Bacteroides vulgatus tag11</i>	As wild-type, with tag11 insertion into ATT site 2	As wild-type, with tag11 insertion into ATT site 2
<i>Bifidobacterium adolescentis</i>	Wild-Type	DSM-20083, Type strain
<i>Bifidobacterium longum</i>	Wild-Type ATCC 15707	DSM-20219, Type strain

Table 2.1 Bacterial strains and mutants used in this study. Mutants shown in bold were made during the course of this study.

2.2 Vectors

Several vectors were used in this study; they are listed in Table 2.2.

Plasmid	Supplier/Reference	Features
pET28a-b	Novagen	Kan ^r , T7 promotor, <i>lac</i> , <i>laciq</i> , integrated His tag
pET28-<i>bt4673</i>	Generated during this study	As above, with ligand binding domain of <i>bt_4673</i>.
pET22b	Novagen	Amp ^r , T7 promotor, <i>lac</i> , <i>laciq</i> , includes option to integrate <i>PelB</i> leader sequence and His tag.
pET21a-b	Novagen	Amp ^r , T7 promotor, <i>lac</i> , <i>laciq</i> , includes option to integrate His tag.
pExchange- <i>tdk</i>	Provided by Nicole Koropatkin (Koropatkin <i>et al</i> , 2008)	Amp ^r , erm ^r , <i>tdk</i> modified suicide vector
pNBU2- <i>tag11</i>	Provided by Nicole Koropatkin (Koropatkin <i>et al</i> , 2008)	Amp ^r , tet ^r , NBU2 transposon, <i>tag11</i> sequence, modified suicide vector
pNBU2- <i>tag1</i>	Provided by Nicole Koropatkin (Koropatkin <i>et al</i> , 2008)	Amp ^r , tet ^r , NBU2-transposon, <i>tag1</i> sequence modified suicide vector
pExchange-<i>tdk-GH98KO</i>	Generated during this study	As pExchange-<i>tdk</i>, with inactive <i>bacova_03433</i>
pExchange-<i>tdk-04390KO</i>	Generated during this study	As pExchange-<i>tdk</i>, with inactive <i>bacova_04390</i>
pExchange-<i>tdk-bt4668KO</i>	Generated during this study	As pExchange-<i>tdk</i>, with inactive <i>bt_4668</i>
pExchange-<i>tdk-bt4669KO</i>	Generated during this study	As pExchange-<i>tdk</i>, with <i>bt_4669</i> flanks
pExchange-<i>tdk-bt4670KO</i>	Generated during this study	As pExchange-<i>tdk</i>, with <i>bt_4670</i> flanks
pExchange-<i>tdk-bt0360KO</i>	Generated during this study	As pExchange-<i>tdk</i>, with inactive <i>bt_0360</i>
pExchange-<i>tdk-bt0362KO</i>	Generated during this study	As pExchange-<i>tdk</i>, with <i>bt_0362</i> flanks
pExchange-<i>tdk-bt0364KO</i>	Generated during this study	As pExchange-<i>tdk</i>, with <i>bt_0364</i> flanks
pExchange-<i>tdk-bt0365KO</i>	Generated during this study	As pExchange-<i>tdk</i>, with <i>bt_0365</i> flanks
pExchange-<i>tdk-0367KO</i>	Generated during this study	As pExchange-<i>tdk</i>, with inactive <i>bt_0367</i>

pExchange-tdk-bt4156KO	Generated during this study	As pExchange-tdk, with inactive bt_4156
pExchange-tdk-bt4170	Provided by Didier Ndeh	As pExchange-tdk, with bt_4170 flanks
pExchange-tdk-bt4175	Provided by Didier Ndeh	As pExchange-tdk, with bt_4175 flanks

Table 2.2 Vectors used in this study. Those in bold were generated during the course of this study.

2.3 Bacterial Growth conditions

2.3.1 Growth Media composition

The composition of growth media used in this study are listed in Table 2.3

Medium	Composition	Quantity (per litre)	Description/Method
Luria-Bertani (LB)	LBB granules, as supplied (sigma-Aldrich)	25 g	Dissolved in MiliQ water and autoclaved before use
Tryptone-Yeast Extract-Glucose (TYG)	Tryptone Yeast Extract Glucose Cysteine, free base 1 M KPO ₄ pH 7.2 0.4 mg/ml FeSO ₄ 1 mg/ml Vitamin K 0.8 % CaCl ₂ 0.25 mg/ml Resazurin TYG Salt Solution (MgSO ₄ 0.5 g/l, NaHCO ₃ 10 g/l, NaCl 2 g/l)	10 g 5 g 2 g 0.5 g 100 ml 1 ml 1 ml 1 ml 4 ml 40 ml	Dissolved in MiliQ water and autoclaved before use
Minimal Media Bacteroides (MM+0.5 % target Glycan)	NH ₄ SO ₄ Na ₂ CO ₃ Cysteine, free base 1 M KPO ₄ pH 7.2 0.4 mg/ml FeSO ₄ 1 mg/ml Vitamin K 0.01 mg/ml Vitamin B ₁₂ 0.25 mg/ml Resazurin MM Salt Solution (NaCl 18 g/l, CaCl ₂ 0.53 g/l, MgCl ₂ 0.4 g/l, MnCl ₂ 0.2 g/l, CoCl ₂ 0.2 g/l)	1 g 1 g 0.5 g 100 ml 10 ml 1 ml 0.5 ml 4 ml 50 ml	Dissolved in MiliQ water and autoclaved before use
Brain-Heart Infusion (BHI)	Used as recommended by manufacturer		Dissolved in MiliQ water and autoclaved before use
Bifidobacterial Minimal Medium (BiMM+0.5 % target Glycan)	Peptone Tryptone KCL NaCO ₃ NaCl MgSO ₄ CaCl ₂ MnSO ₄ FeSO ₄	6.5 g 2.5 g 2 g 0.2 g 4.5 g 0.5 g 0.45 g 0.2 g 0.005 g	Dissolved in MiliQ water and autoclaved before use. Modified from Van der Meulen <i>et al.</i> , 2006.

	ZnSO ₄ Cysteine, free base 1 mg/ml Vitamin K 1 mg/ml Vitamin B12	0.005 g 0.5 g 0.005 ml 0.005 ml	
Clostridial Media (CM)	Used as recommended by manufacturer		Dissolved in MiliQ water and autoclaved before use
His-Heme	Hematin 0.42 g/l Histidine-HCl pH 8.0	1.2 g 1 l	Used to supplement TYG, bacteroides MM, CM and Bifidobacterial BM. Added in a 1 in 1000 dilution prior to inoculation to enhance <i>Bacteroides</i> and <i>bifidobacterial</i> growth.

Table 2.3 Composition of bacterial growth media used in this study

2.3.2 *E. coli* Growth Conditions

The *E. coli* strains described in Table 2.1 were grown at 37°C in Luria-Bertani broth (LB) medium (1 % (w/v) Bacto®Tryptone, 1 % (w/v) NaCl and 0.5 % (w/v) yeast extract, pH 7-7.4. Aeration during growth was achieved by rotary shaking at 150-180 rpm. Agar plates for growth on solid medium were made by addition of 2 % (w/v) Bacteriological agar N°1 (Oxoid) to LB dissolved in water prior to sterilisation by autoclave. While the suspension was still molten approximately 25 ml of LB-Agar was poured into plastic Petri-dishes (Thermo). Prior to pouring relevant antibiotics (Table 2.2) were added for selection of *E. coli* with desired plasmids.

2.3.3 *Bacteroides* Growth Conditions

Specific growth requirements for *Bacteroides spp.* were met by Tryptone-Yeast extract-Glucose (TYG) media (Table 2.3). *Bacteroides spp.* were grown at 37°C in anaerobic conditions, achieved using an anaerobic cabinet.

2.3.4 *Bifidobacterium* Growth Conditions

Specific growth requirements for *Bifidobacterium spp.* were met by Clostridial media, CM (Table 2.3). *Bacteroides spp.* were grown at 37°C in anaerobic conditions, achieved using an anaerobic cabinet.

2.3.5 Selective Media

Antibiotic	Stock Concentration (mg/ml)	Working Concentration ($\mu\text{g/ml}$)	Storage
Ampicillin	50	50	-20°C for 1 month
Kanamycin	10	10	-20°C for 1 month
Tetracycline	1	1	-20°C for 1 month
Gentamycin	0.2	0.2	Made prior to use
Erythromycin	0.025	0.025	Made prior to use

Table 2.4 Antibiotic stocks used in this study

For selection of positive transformants or distinct bacterial species antibiotics were added to media using stock solutions listed in Table 2.4 and diluted appropriately to achieve an effective dose to eliminate unwanted bacteria.

Isopropylthio- β -D-galactoside (IPTG) was added to strains containing *lacIq* either on plasmids or in the genome for induction of transcription of recombinant genes controlled by *lacO*. IPTG stock concentration was made at 1 M in MiliQ H₂O for use with BL21 *E. coli* and 0.2 mM for use with TUNER *E. coli* cells.

2.4 Basic Lab Methods

2.4.1 Storage of DNA and Bacteria

Plasmid and linear DNA solutions were stored at -20 °C while genomic DNA was stored at 4 °C in elution buffer (EB, 10 mM Tris/HCl buffer, pH 8.5). Bacterial colonies on agar plates are stored for a maximum of 2 weeks at 4 °C. For long term storage bacterial culture was mixed 1:1 with 50 % glycerol and stored at -80 °C.

2.4.2 Sterilisation

All media and growth substrates were sterilised by autoclaving at 121 °C under 32 lb/inch² of pressure for 20 min using either an Astell Hearson 2000 Series Autoclave or a Prestige[®] Medical Series 2100 Autoclave. Smaller volumes were sterilised by filter sterilisation through a 0.22 µm sterile filter disc (Stupor[®] Acrodisc[®] 3.2 Gelman Sciences).

2.4.3 Centrifugation

Centrifugation of bacterial cultures ranging from 100-1000 ml was performed at 5000 x g, in 500 ml Nalgene bottles using a Beckman J2-21 centrifuge and JA-10 rotor. Cultures of 1-5 ml were harvested by centrifugation in a MSE Mistral 3000i bench top centrifuge with a fixed angle rotor. A Heraeus Instruments Biofuge pico bench top centrifuge was used to centrifuge Eppendorf tubes up to 13000 x g.

2.4.4 Plating Bacteria

In close proximity to a Bunsen burner flame, 100-200 µl of bacterial cell suspension was dispensed on to an agar plate. A metal spreader was sterilised by immersion into 100 % ethanol which is removed by passing through the Bunsen flame. The spreader was allowed to cool and used to spread the bacterial cell suspension across the surface of the plate.

2.4.5 Chemically Competent *E. coli*

E. coli strains listed in Table 2.1 were made chemically competent following a variation on a protocol outlined in Cohen *et al.* (1972) to allow the uptake of plasmid DNA using calcium chloride. A 1 ml aliquot of a 5 ml overnight culture of the required *E. coli* strain was used to inoculate 100 ml LB (no antibiotic) in a sterile 1 L non-baffle flask. The flask was incubated at 37°C whilst shaking (180 rpm) until log phase was reached ($A_{600nm} = 0.4$). The cells were incubated on ice for 10 min, and then harvested by centrifugation at 5000 x g at 4°C for 5 min. The supernatant medium was removed and cells re-suspended gently in 8 ml of ice cold 100 mM CaCl₂. This was repeated again using 4 ml ice cold 100 mM CaCl₂ to re-suspend the cells. After 2 h on ice, the cell cells were chemically competent

for transformation with purified plasmid DNA. Aliquots of 100 µl of the competent cell suspension were stored at -80 °C with 25 % (v/v) glycerol in Eppendorf tubes.

2.4.6 Transformation of Chemically Competent *E. coli*

An aliquot of chemically competent *E. coli* (Table 2.1) was thawed on ice to retain competence. Still on ice 1 – 5 µl of purified plasmid DNA or ligation mixture was added to 100 µl competent *E. coli* cells. The mixture was incubated on ice for 20 – 30 min, after which the cells were heat shocked at 42 °C for 1 – 2 min and returned to the ice box for 5 min. The cells were then plated onto LB agar plates containing an appropriate antibiotic for selection of positive transformants. The plates were incubated upside-down in a 37 °C incubator for 16 h.

Transformation of ligation mixtures or products of site directed mutagenesis require incubation for 1 h at 37 °C with shaking in 500 µl LB prior to plating out. After incubation the cells are harvested by centrifugation at 13000 rpm, 500 µl of supernatant was removed and the remaining 100 µl was used to re-suspend cells. The 100 µl of cells was then plated onto agar plates containing appropriate antibiotic.

2.4.7 Small scale, rapid purification of plasmid DNA from *E. coli*

The desired plasmid is used to transform Top10 competent *E. coli* cells and grown from single colonies for 16 h in 5 ml LB with an appropriate antibiotic. The cultures were centrifuged at 5000 x g for 10 min. The supernatant was removed. Plasmid purification was then achieved using a QIAGEN® QIAspin Prep kit as per the manufacturer's instructions.

2.4.8 Restriction digest of DNA

The required DNA, where possible 1000 µg per digest, was dissolved in EB buffer (10 mM Tris-HCl, pH 8.5) and mixed with 10 x concentrated reaction buffer in a 1.5 ml Eppendorf tube. The buffer selected was optimal for the specific enzymes used in the reaction. The volume was made up to 40 µl with sigma ultra-pure water or EB buffer and 0.5 – 1 µl enzyme was added. The digest mixture was incubated at 37 °C for 1 – 2 h. the enzymes were inactivated by incubation at 60 °C for 10 min.

2.4.9 Measuring DNA concentration

DNA concentration was determined by spectroscopy using the NanoDrop 2000 benchtop spectrophotometer. The NanoDrop was blanked with either sigma ultrapure water or EB buffer (depending on which the DNA is in solution with) and a reading was taken to confirm the blank as approximately Abs = 0. A 2 µl sample of the DNA solution was measured at A_{260} .

2.4.10 Agarose gel electrophoresis

Size of linear DNA molecules was determined by electrophoresis through submerged horizontal gels. The gels used in this work for rapid analysis of DNA samples were mini gels. Gels were prepared by dissolving 400 mg of agarose (sigma) in 50 ml of 1 x TBE buffer (89 mM Tris Base, 89 mM Boric acid and 2 mM EDTA), giving a 0.8 % (w/v) solution. The suspension was gently mixed and boiled until the agarose had completely dissolved. The gel was then cooled, still molten, and 0.5 µg/ml of ethidium bromide was added, which allows for visualisation of DNA under UV light. The gel was then poured into a mini gel system mould (Applied Biosystems) complete with a comb to create wells for the DNA. Once set the gel was submerged in 1 x TBE buffer and ran at 70 volts for 45 – 60 min (LKB Bromma 2197 Power Supply). Samples were mixed with 6 x loading dye in a sample: dye ratio of 5: 1 and loaded into wells in the gel. Hyperladder I (Bioline) of known DNA sizes was run alongside the samples to allow size determination by comparison with bands in the Hyperladder I lane.

All buffer reagent stocks are shown below in Table 2.5

Buffer	Ingredient	Amount
Electrophoresis Buffer-TBE (pH 8.3 checked)	Tris base	108 g l ⁻¹
	Boric acid	55 g l ⁻¹
	0.5 M EDTA pH 8.0	40 ml
DNA Sample Buffer	Bromophenol Blue	0.25 %
	Glycerol	50 %
	TBE	10 x

Table 2.5 Buffers for agarose gel electrophoresis (x10 stocks)

2.4.11 Visualisation and photography of agarose gels

After electrophoresis a BioRad Gel Doc 1000 using Molecular Analyst™/PC Windows software was used to visualise DNA. Photographs were produced using a Mitsubishi Video Copy Processor (Model P68B) with Mitsubishi thermal paper.

2.4.12 Determination of DNA fragment size

The size of double stranded DNA fragments can be determined by comparison to known standards run alongside the sample in the same agarose electrophoresis gel. Migration rate through a gel matrix is inversely proportional to the log₁₀ of the size of the oligonucleotide fragment.

2.5 Purification of DNA fragments and cloning into vectors

2.5.1 Purification of Vector DNA (Gel extraction)

Vector DNA linearized by digestion with endonucleases was purified by gel electrophoresis using 0.8 % high purity Seachem Gold™ Agarose (w/v) dissolved in 1 x TBE buffer. The DNA was excised from the gel with a scalpel and purification from the gel carried out using a QIAquick Gel extraction kit (Qiagen) as per the manufacturer's instructions.

2.5.2 Purification of inserts and PCR products

Insert DNA digested with endonucleases and PCR products were purified using QIAquick PCR purification Kit (Qiagen) as per manufacturer's instructions.

2.5.3 Ligation of insert and vector DNA

Ligation was carried out by mixing insert and vector digested with compatible restriction endonucleases with 5 x ligation buffer (250mM Tris-HCl (pH7.6), 50 mM MgCl₂, 5 mM ATP, 5 mM DTT, 25% (w/v) polyethylene glycol-8000). Insert to vector ratios used were 3:1 or 6:1. The amount of vector DNA used in the reactions was 20 ng, the amount of insert DNA used was calculated using the below equation:

$$\text{Ratio of insert:vector} \times \left(\frac{\text{Size of vector (bp)}}{\text{Size of insert (bp)}} \right) \times \text{Amount of vector (ng)}$$

Where:

Amount of vector = 20 ng

Ratio of insert to vector = 3 or 6

Ligation reactions were made up in 50 µl PCR tubes.

Amount	Component
Calculated using the above equation	Insert DNA
20 ng	Vector DNA
2 µl	5 x ligation buffer
Up to 7 µl	Sterile water
1 µl	Invitrogen T4 DNA Ligase (4 u/µl)
20 µl	Total volume

Table 2.6 Ligation mixture

The ligation reaction was incubated at room temperature for 30 - 60 min before being used to chemically transform competent One Shot™ Top10 *E. coli* cells.

2.5.4 Ligation independent cloning

An alternative to using ligation it is possible to use T4 DNA polymerase to integrate the insert into the vector. This technique was adapted from (Jeong *et al.*, 2012). Primers were designed with 15 bp extensions which are homologous to the region of vector to allow for homologous recombination into this site. The desired insert with the 15 bp extensions at both ends was generated using PCR. The vector was digested with endo nucleases to linearize the vector. The vector and insert were

mixed in a 1:1, 1:2 or 1:3 ratio (amounts of DNA were calculated using the same equation as in ligation reactions). Buffer (x5), T4 DNA polymerase and 1mg/ml BSA was added to the mixture as shown in Table 2.7.

Component	Stock concentration	Volume added	Final Concentration
Linearized vector	100 ng/μl	1 μl	10 ng/μl
Insert	40 ng/μl	1 μl	4 ng/μl
BSA	x10	1 μl	x1
NEB Buffer 2	x10	1 μl	x1
T4 DNA polymerase	3 U/μl	0.2 μl	0.06 U/μl
H ₂ O		5.8 μl	

Table 2.7 Ligation independent cloning mixture

Once mixed, the reaction solution was incubated at room temperature for 2 min 30 s before being immediately placed on ice for 10 min. Without removing the mixture from the ice 1-5 μl of the the reaction mixture was used to transform chemically competent *E. coli*.

2.5.5 Polymerase Chain Reaction

The polymerase chain reaction was (PCR) developed by Mullis and Faloona (1987) was used to amplify target DNA throughout this study. Regions of DNA were targeted by use of primers 10-20 bases in length with a G/C content of approximately 40 % and a melting temperature (T_m) of > 45 °C that anneal to the boundary of the gene/DNA region of interest. Where possible primers were designed to contain G or C bases at both ends to aid annealing increasing amplification efficiency. Where required restriction sites were included in the 5' –ends of primers capped with a 6 base sequence to allow restriction enzymes to cut the fragment. Primers were manufactured by MWG-Biotech and lyophilised. Primers were in sterile water to a working concentration of 100 pmol/μl.

Primer melting temperature was calculated using the formula:

$$T_m = 64.9 + 41 (yG + zC - 16.4)/(wA + xT + yG + zC)$$

Where w, x, y, z are the number of bases A, T, G, C in the sequence, respectively.

A thermostable DNA polymerase catalyses the synthesis of the complementary DNA strand in the presence of dNTPs. Control reactions lacking DNA template were always carried out. PCR reaction mixture was made according to the amounts in Table 2.8 in sterile 0.2/0.5 ml Eppendorf tubes.

Reagent	Amount
MgCl ₂	25 mM
dNTPs	2 mM each
Oligonucleotide primer 1	50 μM
Oligonucleotide primer 2	50 μM
Template DNA	1-5 ng
Novagen KOD DNA Polymerase	2.5 U/μl
PCR grade water	Volume up to 50 μl

Table 2.8 PCR reaction mixture

The standard thermocycler program was as follows:

Step	Temperature (°C)	Time (min)
1	95	1
2	95	1
3	5 lower than T _m of primer pair	1
4	68	1 per kb of target DNA region
5	68	10
6	4	Until removed from machine

Table 2.9 PCR program

To improve DNA yield steps 2-4 in Table 2.9 were repeated 30-40 times. After each PCR an aliquot of reaction mixture was analysed by electrophoresis.

2.5.6 Site-directed Mutagenesis

Mutagenesis of single amino acids was carried using the site directed mutagenesis method. The site-directed mutagenesis method utilizes an appropriate double-stranded recombinant plasmid DNA and two synthetic oligonucleotide primers (MWG-Biotech AG, Germany) containing the desired mutation flanked by 10-15 nucleotides that fully complemented the DNA template. The

oligonucleotide primers are extended during temperature cycling by using a thermostable KOD polymerase. The same reaction mixture as standard PCR is used (Table 2.8) and thermocycler program shown in Table 2.10.

Step	Temperature (°C)	Time (min)
1	95	1
2	95	1
3	55	1
4	68	1 per kb of plasmid length
5	68	10
6	4	Until removed from machine

Table 2.10 Site directed mutagenesis thermocycler program

After thermocycling the reaction mixture was cooled on ice before the addition of 1µl of *DpnI* (Fermentas 10 U) into each reaction mixture before being vortexed and centrifuged briefly. The digestion was then incubated for 1 h at 37 °C. *DpnI* digests methylated template DNA leaving the unmethylated PCR product intact. The DNA was then used to transform *E. coli* TOP10 competent cells.

2.5.7 PCR Overlap Extension

A plasmid carrying a region homologous to part of the *Bacteroides* genome, but lacking the target gene, was required for genomic disruption. The CC118 λ pir strain of *E.coli* was used for cloning of these fragments into the pExchange *tdk* plasmid. Restriction sites BamHI and XbaI were routinely used, and Sall or SpeI used if necessary.

To remove a gene, a fragment was required which possessed homology to a region of at least 1,000 bases upstream and 1,000 base pairs downstream of the gene, but lacking the gene ORF entirely. These regions are referred to as the upstream and downstream flanks.

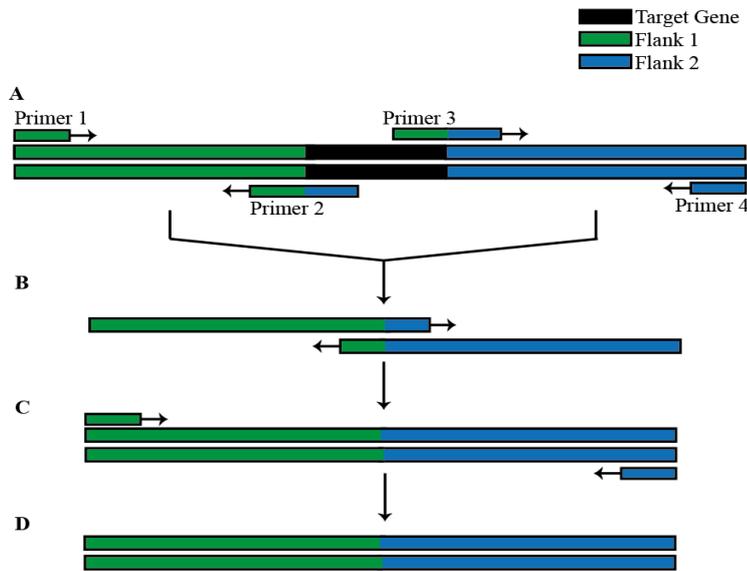


Figure 2.1 Sewing PCR was used to Generate Gene Knockout Fragments. Sewing PCR was used to remove a gene from a fragment of DNA. Two 1,000 bp regions flanking the gene of interest were amplified during two distinct reactions using an amplification primer (primer 1, primer 4) and a sewing primer containing homology to the terminus of the flank to be joined (primer 2, primer 3). This step creates two fragments with a region of complementarity (a). The first step of the two-step sewing PCR cycle contains only the two gene flanks, which act as both primer and template DNA to each other (b). The second step of the two-step sewing PCR cycle is undertaken with the addition of the two amplification primers (primer 1, primer 4, c). A hybrid fragment is created which lacks the target gene (d). Adapted from Shapiro (2015).

Four sets of primers were required for construction of this fragment, two amplification primers and two internal “sewing” primers which allowed the two flanks to be joined (Figure 2.1). The upstream primer (Primer 1) and the downstream primer (Primer 4) were designed as for routine amplification primers with appropriate restriction sites. Primer 3 and 4 were designed by selecting the 20 base pairs immediately upstream from the start codon and the 20 base pairs directly after the stop codon and adding these together to form a 40 base pair sequence. This sequence was used in antisense (Primer 3) and sense (Primer 4) to target each DNA strand.

The first step (Figure 2.1a) is to create the two flanks, with homologous region. This is done by using primer set 1 + 2 and primer set 3 + 4 in two distinct routine PCR amplifications.

The second step (Figure 2.1b) is to join these flanks together. The first cycle is performed with no primers or template DNA, but both flanks in equal concentration around 1 µg. As the flanks contain complementarity they will act as both primer and template to each other. After 10 cycles the amplification primers were added (Primer set 1 + 4) in order to efficiently amplify the newly constructed 2,000 base pair fragment (Figure 2.1c).

It was essential to begin with both flanks at equal concentration, otherwise the reaction would favour the amplification of one flank above the other and result in an undesirable 1,000 base pair product. Once obtained, the 2 kb fragment was then purified and ligated into pExchange *tdk* vector.

2.5.8 Automated DNA Sequencing

DNA sequencing was conducted using the Value Read service provided by MWG Viotech AG, Ebersberg, Munich, Germany using ABI 3700 sequencers and Big dye technology (applied Biosystems). Each clone was sequenced in both the forward and reverse direction. As required by the company, 7-15 µl of 50-100 ng plasmid DNA or linear DNA fragment was sent per sample. Where possible plasmids were sequenced from the T7 promoter and T7 terminator. When not possible specific primers were designed for the desired regions to be sequenced. T7 primer sequences are as follows, T7 promoter – TAATACGACTCACTATAGGG, T7 terminator – CTAGTTATTGCTCAGCGGT.

2.6 Protein expression and purification

2.6.1 Protein expression

Expression plasmids carrying the sequence encoding the protein of interest were used to transform BL21 or TUNER *E. coli* cells which were used to inoculate 5 ml LB containing the corresponding antibiotic to preserve the vector during bacterial growth. The inoculated media was then grown in a shaking incubator set at 37 °C for 16 h. The 5 ml culture was then used to inoculate 1 L LB (in 2 L flask) with the same antibiotic as the previous culture. The flask was then placed in a shaking incubator for approximately 4 h, or until an OD_{600nm} of 0.6 is achieved. The flask was removed and

cooled for 30 min before the addition of 1 ml 1 M IPTG and returned to the incubator, set this time to 16 °C, for 16 h.

The culture was removed and dispensed into 500 ml centrifuge tubes before being centrifuged at 5000 rpm using a JA10 rotor in a Beckman Avanti centrifuge. The supernatant was removed and cells resuspended in TALON buffer (20mM Tris, 300mM NaCl pH 8.0) and sonicated for 2 min using a B. Braun Labsonic U sonicator set at low intensity ~42 watts and 0.5 second cycling before being transferred to a 50 ml centrifuge tube (Nalgene) and pelleted by centrifugation at 15000 rpm for 30 min using a JA25.50 rotor. The supernatant, here after referred to as cell lysate, was collected to purify recombinant protein. The pellet was resuspended in 10 ml TALON. Protein was purified from the cell lysate by Immobilised Metal Affinity Chromatography (IMAC) with TALON Resin (Clontech).

2.6.2 Immobilised Metal Affinity Chromatography (IMAC)

His-tagged proteins can be purified by (IMAC) as the Histidine side-chain interacts with electropositive transition metal immobilised in a column; this interaction can be disrupted by imidazole. The His-tag protein can be eluted from the column matrix using an imidazole gradient. TALON™ (Clontech Laboratories Inc.) columns with 2 ml TALON™ resin containing cobalt ions were prepared by washing in 10 volumes water and then TALON™ buffer (20 mM Tris/HCl buffer, pH 8.0, containing 300 mM NaCl). Cell lysate was filtered (0.45µm) and poured through the column. The column was washed with 4 x 5 ml of TALON™ buffer. The protein was eluted with 2 x 5 ml of TALON™ buffer containing 10 mM imidazole followed by 2x 5 ml of TALON™ buffer containing 100 mM imidazole. Analysis by SDS-PAGE of each stage of the purification indicated which fractions contained the purified protein.

2.6.3 SDS-PAGE

Protein was visualised by SDS-PAGE as described by Laemmli (1970) to determine the size, relative purity and relative quantity of the protein. 12.5 % polyacrylamide gels (Acrylogel 3; BDH Electran®) were routinely used for protein visualisation in conjunction with the AE-6450 apparatus from ATTO

Corporation (Genetic Research Instruments) which utilises 12 cm x 10 cm glass plates sealed with a rubber gasket. The SDS-PAGE running buffer is given in Table 2.11. The resolving gel (Table 2.12) was poured into the plates, covered with water and allowed to polymerise. The water was then removed and the stacking gel poured on top of the resolving gel. A comb is added and polymerisation of this second layer is allowed to take place. Before use, the comb and rubber seal is removed and the gel affixed within the gel tank, which is filled with running buffer. Loading dye was added to samples at a ratio of 1:2 and samples were boiled for 2 minutes to denature the proteins. Samples were loaded, alongside standards for comparison into the gel wells and a current of 35 A (per gel) was applied. Gels were run for approximately 30-45 min or until the loading buffer band has reached the end of the gel.

Component	Reagent	Volume or Concentration
Running Buffer (For 1 l*)	32 mM Tris/190mM glycine, pH 8.3	350 ml
	SDS	0.1 % (w/v)
Loading Buffer (For 10 ml*)	SDS	10 % (w/v)
	0.25M Tris/HCl, pH 8.8	5 ml
	Glycerol	25 % (w/v)
	β-mercaptoethanol	2.5 ml
	Bromophenol blue dye	0.1 % (v/v)

Table 2.11 SDS-PAGE Gel Running buffer

Component	Reagent	Volume per gel
Resolving Gel	0.75 M Tris/HCl, pH 8.8 with 0.2 % SDS	2.35 ml
	40 % Acrylamide (BDH Electran acrylamide, 3 % (w/v) bisacrylamide)	1.45 ml
	d.d. H ₂ O	0.875 ml
	10 % (w/v) Ammonium persulphate	22.5 µl
	TEMED	7.5 µl
Stacking Gel	0.25 M Tris/HCl, pH 8.8 with 0.2 % SDS	0.938 ml
	40 % Acrylamide (BDH Electran acrylamide, 3 % (w/v) bisacrylamide)	0.188 ml
	d.d. H ₂ O	0.75 ml
	10 % (w/v) Ammonium persulphate	15 µl
	TEMED	5 µl

Table 2.12 SDS-PAGE Gel preparation

After electrophoresis, the gel was soaked in InstantBlue™ stain (Expedeon) for 15 minutes to reveal protein bands, after which they were washed in distilled water. Gels were then photographed and catalogued (Bio-Rad Gel Doc 1000, Molecular Analyst™/PC windows Software).

2.6.4 Determination of Protein concentration

Protein concentration was deduced by use of a NanoDrop 1000 and the protein concentration analysis feature of the NanoDrop 1000 software. Purified protein sample was taken and 2 µl loaded onto the pedestal for OD_{280/320nm} analysis. The optical density was calculated using the equation:

$$A = \epsilon CID$$

Where A = absorbance at 280 nm – absorbance at 320 nm, ϵ = molar extinction coefficient, I = length of light path (cm), D = dilution factor and C = molar concentration of sample.

The extinction coefficient for each protein was found by entering the protein amino acid sequence into the ProtParam tool (www.expasy.com).

2.6.5 Protein Concentration

Protein solutions were concentrated using 20 ml or 2 ml Viaspin™ centrifugal concentrators (VivaScience) with 10, 30 or 50 kDa molecular weight cut off filters (as appropriate). Centrifugation was performed at 3000-3500 xg using a MSE Mistral 3000i bench centrifuge with a swingout rotor at 10 °C.

2.7 Bioinformatics

2.7.1 Alignments

Amino acid sequence searches were carried out using the Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1997), using the NCBI (National Centre for Biotechnology Information) version hosted at the European Bioinformatics Institute (EBI) website (www.ebi.ac.uk). Amino acid sequences were aligned using ClustalO at the EBI website.

2.7.2 Prediction of prokaryotic signal peptides

The presence or absence of signal peptides, and the signal peptide cleavage position, was predicted using the SignalP 3.0 software hosted at www.cbs.dtu.dk/services/SignalP/ (Bendtsen et al., 2004).

2.7.3 Identification of loci in *Bacteroides* and *Bifidobacterium* spp

Currently identified and predicted *Bacteroides* polysaccharide utilisation loci are archived in the PULDB (www.cazy.org/PULDB/). *Bifidobacterium* potential loci were found by manual search of the Kyoto Encyclopedia of Genes and Genomes (KEGG, www.genome.jp/kegg/).

2.8 Biochemistry

2.8.1 Enzyme assays

Unless otherwise stated, all reactions were performed in pre-warmed solutions, prior to addition of enzyme, at 37 °C. Assays were repeated at least three times where possible. Graphs were plotted in GraphPad Prism 6.0 and used to calculate slopes, gradients and standard errors. Non-linear 'one-site' binding model was used to fit kinetic data to estimate K_M and k_{cat} .

2.8.1.1 3,5-Dinitrosalicylic acid (DNSA) Reducing Sugar Assay

The rate of hydrolysis was monitored by the increase in reducing sugar formed over time. The free anomeric carbon at the end of a polysaccharide can open from its more common cyclic conformation and act as a weak reducing agent. Each time a glycosidic bond is hydrolysed a new reducing end is formed, the concentration of which can be determined with the DNSA reagent using the Miller method (Miller, 1959). A 100 µl aliquot of an enzyme reaction was added to 100 µl DNSA reagent (1 % (w/v) DNSA, 0.2 % (v/v) phenol, 1 % (w/v) NaOH, 0.002 % glucose, 0.05 % (w/v) NaSO₃) to terminate the reaction. The tube was then boiled for 20 min, placed on ice for 10 min, equilibrated to room temperature. The samples were then transferred to a 96-well plate (CoStar) and inserted into a benchtop plate reader with the absorbance read at 575 nm. A standard curve of

0-1000 $\mu\text{g/ml}$ monosaccharide (plus polysaccharide substrate) was used to quantify the released reducing sugar.

2.8.1.2 PNP substrates

Relevant 4-Nitrophenyl substrates were used to show enzyme activity in initial activity tests. A 10 mM stock of the relevant substrate was made. A final reaction concentration of 1 mM substrate, 0.1 mg/ml BSA and 1 μM enzyme in 20 mM sodium phosphate 150 mM sodium chloride pH 7.5.

Reactions were performed in 1 ml plastic cuvettes with optical density readings at 405nm taken by a pharmacia Biotech Ultrospec 4000 spectrophotometer. The enzyme, if active, cleaves the glycosidic bond releasing *p*-Nitrophenol, a chromogenic substrate which is detected at 405 nm. Concentration of *p*-Nitrophenol was calculated using the extinction coefficient, 18,000 $\text{M}^{-1} \text{cm}^{-1}$.

2.8.1.3 Galactose/arabinose detection kit

The release of galactose/arabinose was monitored through the use of a linked assay utilising galactose dehydrogenase (Megazyme), respectively (galactose dehydrogenase is also able to oxidise arabinose). Galactose dehydrogenase catalyse the oxidation of galactose/arabinose and reduces NAD^+ to generate NADH in a 1:1 molar ratio. NADH absorbs at $A_{340\text{nm}}$ and thus its synthesis can be directly monitored. An extinction coefficient of 6230 $\text{M}^{-1}\text{cm}^{-1}$ was used to calculate the NADH concentration. All reactions were carried out at 37 °C in 20 mM sodium phosphate 150 mM NaCl buffer pH 7.5. Glass cuvettes were used and NADH release monitored using a Biotech Ultrospec 4000, UV/vis spectrophotometer. Table 2.13 lists the reaction composition for the linked assay.

Volume (μl)	Component
50	Appropriately diluted enzyme
50	0.02 M Sodium phosphate pH 7.5
50	10 mM NAD^+

50-345	Substrate
5	Galactose dehydrogenase 500 U/ml (1 U = oxidisation of 1 μmol galactose min^{-1})
50-345	Sterile water
500	Total volume

Table 2.13 Composition of typical galactose dehydrogenase linked assay

2.8.1.4 Acetic acid detection kit

Using a similar spec assay protocol as galactose/arabinose detection acetic acid accumulation although with a 3 step reaction:

- (1) Acetic acid + ATP + CoA \rightarrow acetyl-CoA + AMP + pyrophosphate
- (2) Acetyl-CoA + oxaloacetate + H₂O \rightarrow citrate + CoA
- (3) L-Malate + NAD⁺ \leftrightarrow oxaloacetate + NADH + H⁺

Where reactions are mediated by acetyl-CoA synthetase (1), citrate synthase (2) and L-malate dehydrogenase (3).

Accumulation of NADH is produced at a 1:1 molar ratio as acetic acid and measured at $A_{340\text{nm}}$ with a Biotec Ultrspec 4000, UV/vis spectrophotometer and reactions were performed in plastic disposable 1 ml cuvettes. All reactions were performed at 37 °C in 20 mM sodium phosphate 150 mM NaCl pH 7.5.

2.8.2 High Pressure Liquid Chromatography/associated assays

Enzyme assays were performed at 37 °C in 20 mM sodium phosphate buffer, pH 7.5, and included 500 μM rhamnose as an internal standard, and 0.1 mg/ml BSA to prevent enzyme absorption to surfaces. Each substrate was at a final concentration of 500 μM in the reaction, which was suspected to be $\ll K_M$ from previous work with similar GHs. Scoping experiments were performed to find an enzyme concentration that would give ~80% degradation for each substrate in 60 min. Aliquots of 50 μl were taken at time points which was then heated to inactivate the enzyme, centrifuged at 13000 x

g for 10 min, and diluted 10-fold in MiliQ water in HPLC Vials. Samples were analysed using an analytical CARBOPAC™ PA-100 anion exchange column (Dionex) equipped with a CARBOPAC™ PA-100 guard column, sugars were detected by Pulsed Amperometric Detection. Elution conditions were 0-10 min 66 mM NaOH, 10-25 min 66 mM NaOH with a 0-75 mM sodium acetate linear gradient, followed by a wash with 500 mM sodium acetate for 10 min then 500 mM NaOH for 10 min. Data were collected and manipulated using Chromeleon™ Chromatography Management System V .6.8 (Dionex) via a Chromeleon™ Server (Dionex). To ensure sensitivity remained constant over the entire run all peak areas were normalised to an internal rhamnose standard. Rates were determined by substrate depletion, using the following formula:

$$K = \ln (S_0/S_t)$$

$K = k_{cat}/K_M$ [enzyme], \ln = natural log, S_0 = substrate concentration at time 0, S_t = substrate concentration at time t.

Peaks corresponding to different oligosaccharides were identified by co-elution with standards of known oligosaccharides. The concentration of each oligosaccharide was calculated by comparing the peak area with standards of each oligosaccharide at known concentrations.

2.8.3 Isothermal Titration Calorimetry

Isothermal Titration Calorimetry (ITC) can determine the thermodynamic parameters driving macromolecular interactions, by titration of ligand into protein at 25 °C. Enthalpic and entropic changes are measured by the heat released or consumed in the reaction cell compared to a reference cell. In this project both protein and ligand were in 50 mM HEPES/NaOH pH 7.5. The protein in the cell, at 100 µM, was titrated with 20 x 10 µl injections of ligand, which, in the syringe, was at a concentration of either 5 mM for oligosaccharides and 10 mg/ml for polysaccharide. Integrated heat effects were analysed by non-linear regression using a single site-binding model (Microcal Origin v7) giving the association constant (K_a) and enthalpy of binding (ΔH),

Other parameters were calculated using:

$$-RT \ln K_a = \Delta G = \Delta H - T\Delta S$$

R = gas constant, ln = natural log, ΔG = change in Gibbs free energy, ΔH = change in enthalpy, T = absolute temperature (Kelvin), ΔS = change in entropy

2.8.4 Thin Layer Chromatography

Thin Layer Chromatography (TLC) allows visualisation of mixtures of oligosaccharides in solution, as migration patterns vary for different oligosaccharides with respect to degree of polymerisation (d.p.) and structure. TLC plates (Silica gel 60, Merck) were cut to size and 3-6 μl of samples were spotted on plates, dried and placed in a tank containing 1 cm running buffer (1-butanol/acetic acid/water, 2:1:1 (v/v)). Plates were then left to allow the running buffer to migrate to the top of the plate, dried and left to migrate again. Plates were dried and submerged in developer (sulphuric acid/ethanol/water, 3:70:20 v/v, 1% orcinol) for 5 seconds to allow visualization of the sugars. Plates were dried and developed by heating to 80-100 °C.

2.8.5 Acid hydrolysis/oligosaccharide purification

Galactooligosaccharides, Gal3-6, were produced by acid hydrolysis of potato galactan (Megazyme) using 100 mM HCl at 100 °C for 1 h. The reaction was neutralised by titration with NaOH until neutral pH was achieved, and samples of the reaction were run on TLC to assess the degree of hydrolysis (data not shown). The samples were freeze-dried, resuspended in 3 ml MiliQ water and applied to a P2 (Bio-Rad) matrix packed in 2 Glass Econo-Column™ for purification by size exclusion. The oligosaccharides were eluted in distilled water at a flow rate of 0.23 ml/min; 1.8 – 2 ml fractions were collected and assessed by TLC. Fractions containing pure oligosaccharides were pooled, freeze dried and resuspended in volumes to allow for stock concentrations between 6 – 15 mM.

2.9 Microbiology

2.9.1 Culture Preparation and Monitoring

Bacterial cultures were set up by inoculation of the desired medium using a rate of 10 µl active bacterial culture or 20-50 µl glycerol stock per 1ml of growth medium. All experimental cultures were inoculated from actively growing cultures (these were seeded using glycerol stocks).

Bacterial cultures could be monitored directly throughout growth using a 96 well or 24 well corning® costar® culture plate (Sigma-Aldrich) in conjunction with an Epoch microplate spectrometer (Biotek Instruments Ltd.) inside of an anaerobic chamber (Don Whately Scientific). Data were manipulated in Gen5 2.05 software and later plotted using Prism 6.0 (GraphPad). 96 well plates allowed for culture volumes of 200 µl, and 24 well plates of 2 ml. The plate reader measured and recorded the optical density (at 600 nm) of each well at 15 minute intervals. Each well was prepared in triplicate and the data averaged. Media without bacterial inoculum was always run as a control to ensure no contamination has occurred throughout the growth period and positive control of media with glucose was always used.

For larger monocultures and all co-cultures glass test tubes were used to hold 5 ml aliquots of media. These were plugged with cotton wool to prevent contamination prior to sterilisation. Tubes were inoculated and incubated in an anaerobic chamber (Don Whately Scientific). OD (at 600 nm) was measured using a CO 7500 spectrophotometer (Biochrom).

2.9.2 *Bacteroides* and *Bifidobacterium* Co-culture

Co-cultures were inoculated from overnight (16 h) incubations of approximately the same volume and density of overnight culture washed twice in PBS. Cultures were grown in 5 ml volume glass test tubes and monitored at 600nm using a CO 7500 spectrophotometer (Biochrom). Each culture was performed in triplicate.

Samples were taken at time points were serially diluted in PBS and micro-plated onto a clostridial media-agar plate divided into 9 equal sections. 10 µl of each dilution was plated onto each section and incubated under anaerobic conditions for 2 days. Colonies were counted in each dilution where possible and an average CFU/ml was calculated for each time point. In the case of *Bacteroides-Bifidobacterium* co-cultures colonies were differentiated based on colony morphology and confirmed by replica plating onto clostridial agar media supplemented with gentamycin. *Bifidobacterium* is unable to grow on gentamycin media while *Bacteroides* is able. In the case of *Bacteroides-Bacteroides* co-cultures 0.5 ml samples were taken and genomic DNA was purified from the culture and subjected to quantitative real-time PCR to differentiate between mutant and wild type *Bacteroides* by presence of unique sequences inserted in the genome.

2.9.3 Genomic DNA Extraction

DNA extraction from 5 ml cultures was undertaken using the GenElute™ Bacterial Genomic DNA Kit (Sigma Aldrich) according to the manufacturer's instructions.

2.9.4 *Bacteroides* mutagenesis and genomic insertions

The modified suicide plasmid, pExchange *tdk*, containing a knockout fragment were transformed into S17 λ *pir E.coli* cells, referred to as the "donor" strain. *B. ovatus tdk-* or *B. thetaiotaomicron tdk-* is the "recipient" strain. The donor and recipient strains were cultured (5 ml) to roughly equivalent cell densities in LB broth and TYG media respectively (Figure 2.2a). Cells were harvested by centrifugation and washed in TYG medium. Equal sized cell pellets were then re-suspended in 1 mL TYG medium and spread evenly on the surface of BHI plates with no antibiotic. These plates were incubated agar side down and grown for 16-24 hours until a thick lawn has formed; *E.coli* should grow first, creating an anaerobic environment underneath this growth in which *Bacteroides* can thrive, providing the necessary conditions for plasmid conjugation from the donor to the recipient strains (Figure 2.2b). This biomass was scraped from the plate and re-suspended in 5 ml TYG

medium. Then 100 µl of this solution, along with three serial dilutions (1:10, 1:100, 1:1000) were plated onto BHI + gentamycin (200 µg/ml) + erythromycin (25 µg/ml) plates. These antibiotics select for the recipient strain and the pExchange *tdk* plasmid, thus colonies represent single recombinant where the pExchange *tdk* has recombined with the genomic DNA via one of the flanks. These plates were incubated anaerobically for up to 2 days or until colonies formed, then 10 colonies were picked and re-streaked onto fresh BHI + gentamycin + erythromycin plates to minimise wild type contamination (Figure 2.2c). Then 10 colonies were cultured overnight in TYG medium, 1 ml of each culture was taken and a pooled stock created. A glycerol stock can be made at this stage for safekeeping (Figure 2.2d).

The pooled stock alongside three serial dilutions (1:10, 1:100, 1:1000) was plated upon BHI + FUdR (200 µg/ml) and allowed to grow anaerobically for 2 days or until colonies appeared. FUdR is toxic to strains able to synthesise thymidine. The recipient strain lacks the *tdk* gene, but this has been complemented within the pExchange *tdk* plasmid, in this manner FUdR selects for the second recombination event, whereby the second flank incorporates into the genome and the pExchange *tdk* sequence is eliminated. Following this growth, 10 FUdR resistant colonies were re-streaked onto fresh BHI + FUdR plates to minimise wild-type contamination (Figure 2.2e). 10 resistant colonies were picked and cultured in 5 ml of TYG so that genomic DNA could be extracted and glycerol stocks could be made (Figure 2.2f).

Isolated DNA was screened for successful knockout mutations using PCR. The downstream and upstream primers used to create the plasmid (primer 1 & primer 4, Figure 2.1) were used to amplify the clones, using wild-type *Bacteroides* as a control; the wild-type strain will produce a fragment which is the length of the target gene (500-2000 bp), plus the length of both flanks (1,000 bp each). Any successful knockouts will lack the target gene, yielding a fragment of 2,000 bp. Clones which appeared successful after screening were then sequenced to ensure the correct mutation had taken place.

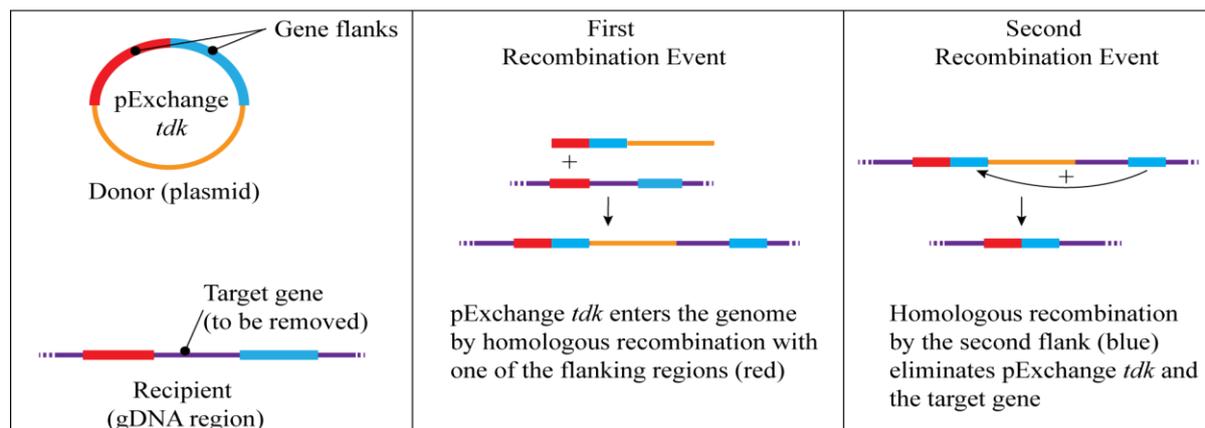
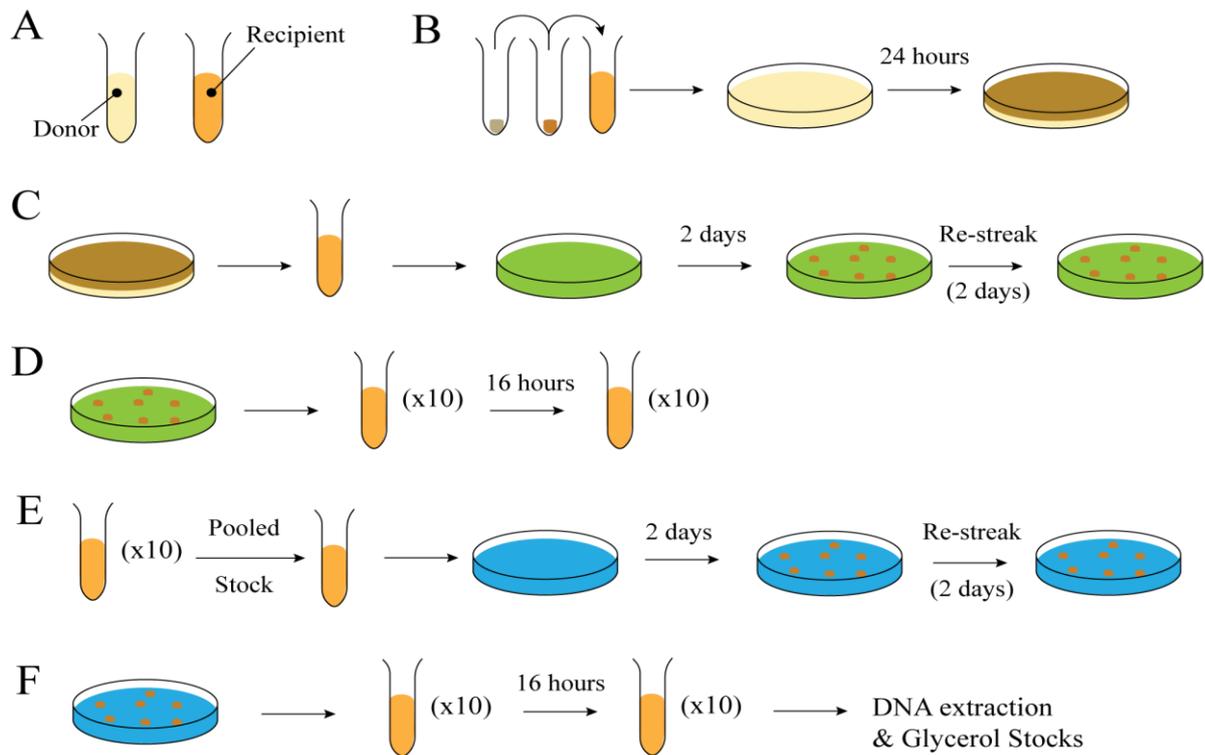


Figure 2.2 Generating Knockout Strains of *Bacteroides ovatus* and *Bacteroides thetaiotaomicron*. (a) The donor and recipient strains were cultured in 5 ml of LB and TYG media respectively. (b) Equal size cell pellets were harvested by centrifugation, washed in TYG, combined and re-suspended in 5 ml TYG and plated onto BHI plates containing no antibiotics (yellow). These plates were not inverted during growth. (c) The plates were scraped and the biomass re-suspended in 5 ml TYG. This was plated onto BHI plates containing gentamycin (200 $\mu\text{g}/\text{ml}$) and erythromycin (25 $\mu\text{g}/\text{ml}$) (green). Resistant colonies were re-streaked onto fresh plates to minimise wild-type contamination. (d) 10 colonies (these represent the first recombination event) were picked and cultured overnight in TYG. (e) The cultures were pooled into one stock, which was plated onto BHI containing FUdR (200 $\mu\text{g}/\text{ml}$) (blue) to select for the second recombination event, as before these are re-streaked. (f) 10 resistant colonies are cultured overnight in TYG. Glycerol stocks are prepared and DNA extracted for analysis. Inset: A visualisation of the donor and recipient DNA, and the first and second recombination events. Figure taken from Shapiro (2015).

In a similar method, unique sequences of DNA referred to here as Tags can be inserted into one of two ATT sites of *Bacteroides* to allow for differentiation of species or mutants within a co-culture. The tags are carried on modified suicide vectors, pNBU2-*tag11* and pNBU2-*tag1*, which were used to transform CC118 *E. coli* competent cells which are plated on LB-ampicillin plate overnight. Overnight cultures were then made from the resulting colonies and from glycerol stocks of the *Bacteroides* in LB and TYG, respectively. The overnight cultures were then used to inoculate sterile media for a 4 h outgrowth. The 4 h cultures were centrifuged (6000 x g, 15 min) and resuspended in 1 ml TYG together. This resulting *E. coli-Bacteroides* culture was plated onto BHI-agar without antibiotic. There plates were then incubated aerobically for 24-30 h and the resulting biomass was removed and resuspended in fresh TYG (5 ml). The resulting cell suspension was diluted 10^{-1} , 10^{-2} and 10^{-3} and plated onto BHI-agar with gentamycin (200 µg/ml) and tetracycline (2 µg/ml). The plates are incubated anaerobically for 2 days, then 10 colonies were selected and re streaked onto fresh BHI-agar with gentamycin ant tetracycline which were incubated anaerobically for a further 3 days. Single colonies were selected and grown in TYG overnight. These cultures were subjected to genomic DNA extraction and PCR was performed to check for tag insertion. A second PCR was performed to show which site has been destroyed by insertion of the tag. PCR controls are used with wild type genomic DNA.

2.9.5 Whole cell assay

Overnight 5 ml cell cultures were gently harvested by centrifugation (6,000 x g), the supernatant removed and the pellet washed with Phosphate Buffered Saline Buffer (PBS). This step was repeated twice to ensure thorough washing before the pellet was finally re-suspended using 1 ml PBS. In the presence of oxygen, the cells are metabolically inactive but retain structural integrity. Thus, proteins which do not require ATP (such as CAZymes) presented at the cell surface remain active and can be observed.

Whole cells preparations were used as the catalytic agent during assays to detect CAZyme activity. A 1 ml total reaction volume was used (500 µl whole cell assays, 500 µl 1 % glycan solution in PBS). Cells were boiled for 10 minutes and this matter used as a control reaction. To stop this reaction, samples were centrifuged (to remove cells) and the supernatant boiled. These samples were analysed by TLC.

2.9.6 Real-time Quantative PCR (qPCR)

Genomic DNA harvested from co-cultures could be quantified using qPCR (in this thesis, all qPCR was real time qPCR) by amplification of unique regions of DNA know as tags. These tags were inserted into the genome of wild type and mutant *Bacteroides* to allow differentiation in co-culture. qPCR relies on a dye which releases measurable light during amplification. This project used SYBR Green I (Roche), an intercalating dye which absorbs light at 497 nm and emits light at 520 nm when intercalated into double stranded DNA. During each amplification step the samples were illuminated at 497 nm and emitted light at 520 nm was measured. The more amplification which occurs during the PCR reaction, the greater the intensity of light released, as the dye is incorporated into new double strands. A CQ (quantification cycle) value is obtained at the cycle where fluorescence from the sample exceeds background florescence, this is the point at which it is clear that a fragment is being amplified. A low CQ value means fewer cycles were required to detect amplification, whilst a higher CQ value shows that more cycles were required. Every experiment included a control with the probe set but no template DNA to ensure that background florescence was not classed as amplification throughout the experiment. And a melt curve analysis was performed with each run to ensure a single PCR product was present in all reactions.

10 µl Reactions were set up using 5 µl of the SYBR Green I Master Mix (Roche), 1 µl of each probe (5 µM forward primer, 5 µM reverse primer), 2 µl of template gDNA or cDNA and 1 µl of PCR grade water. qPCR was carried out using a LightCycler® 480 (Roche) or using a Roche LightCycler® 96. Program used throughout this project is given in Table 2.14.

Step	Process	Temperature (°C)	Time (s)
1	Initial Denaturation	95	600
2	Denaturation	95	10
3	Annealing	57	10
4	Elongation	72	10
5	Measurement	72	-

Table 2.14 qPCR program and parameters used throughout this project. Steps 2-5 are repeated 45 times and the measurement was taken by exciting samples at 497 nm and measuring light emitted at 520 nm.

Chapter 3: Xylan Utilisation and Cross-feeding by *Bacteroides ovatus*

3.1 Introduction

The majority of data presented in this chapter were published previously (Rogowski *et al.*, 2015).

3.1.1 Background

Xylan is a major component of plant cell walls, and one of the most variable plant structural polysaccharides. It is present in particularly high concentrations in the endosperm and cell walls of cereals, an important component of the human diet. Xylan is a hemicellulose consisting of a highly conserved β 1,4-linked xylopyranose backbone decorated with various sugars, primarily glucuronic acid (GlcA) and arabinofuranose, and acetyl groups (Allerdings *et al.*, 2006; Agger *et al.*, 2010). Decorations to the linear backbone increase the complexity of the polysaccharide. Glucuronoxylans (GX) as found in birchwood xylan (BX), are at the simpler end of the spectrum of possible xylan structures, with α -GlcA substitutions at the O2 of the backbone xylose residues (Figure 3.1). Arabinoxylans (AX) structures, found in wheat xylan (WX), are slightly more complex with either single or double arabinofuranose substitutions at the O2 and/or O3 positions on xylopyranose backbone (Figure 3.1) (Allerdings *et al.*, 2006; Agger *et al.*, 2010). The most complex xylan structure used in this study was glucuronoarabinoxylan (GAX) found in corn xylans (CX) present in the grain wall. CX possesses the methylated GlcA of BX and the arabinofuranosyl substitutions of WX along with 1,2 or 1,3 linked α - and β -xylopyranose, and α -L- and β -D-galactose units (Figure 3.1).

Bacteroides species degrade and utilise glycans either part of the host diet or glycans expressed by host cells encountered in the human gut (Martens *et al.*, 2011). These systems are encoded by gene clusters known as polysaccharide utilisation loci (PULs). These PULs encode carbohydrate active enzymes (CAZymes), which are identified and assigned to sequence based families in the CAZY database (Lombard *et al.*, 2014; Terrapon *et al.*, 2015). PULs are identified due to the presence of adjacent *susC* and *susD*-homologues along with CAZymes and a regulator that controls transcription

of the locus. The CAZymes expressed in PULs are localised to the cell surface or periplasm. Surface enzymes, typically endo-acting, degrade polysaccharides into oligosaccharides which are transported by the SusCD-like proteins into the periplasm where periplasmic CAZymes complete degradation to mono- or disaccharides for fermentation (Martens *et al.*, 2009).

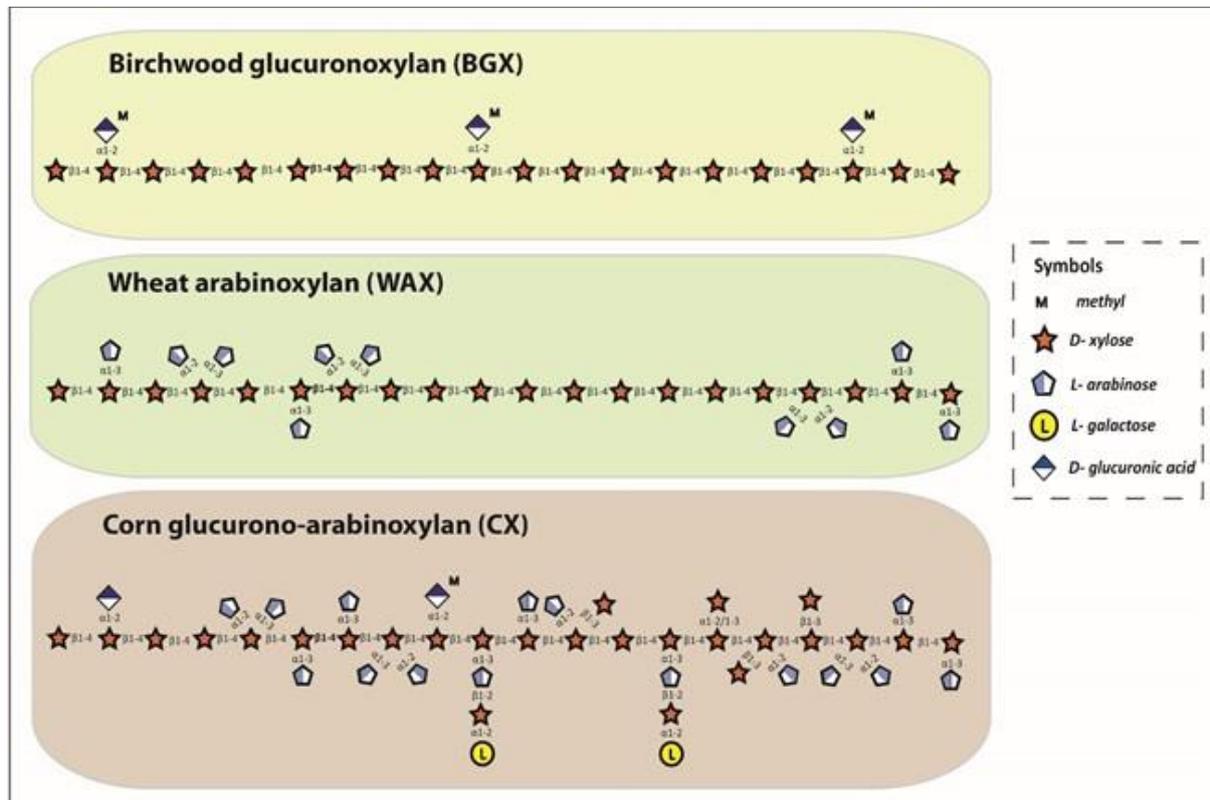


Figure 3.1 Schematic structure of the main classes of xylan used in this chapter. The monosaccharide and linkages of the glucuronoxylin, arabinoxylin and glucuronoarabinoxylin represented in their Consortium for Functional Glycomics format (Raman *et al.*, 2006). The xylan polysaccharides used in this chapter were from Birchwood (Birch glucuronoxylin, BGX), Wheat flour (Wheat arabinoxylin, WAX) and Corn bran (Corn glucuronoarabinoxylin, CX). Taken from Rogowski *et al.* (2015).

The prominent human gut bacterium, *Bacteroides ovatus*, utilizes a range of xylans (Martens *et al.*, 2011). The *B. ovatus* PULs that orchestrate xylan degradation were identified by Martens *et al.* (2011) (Figure 3.2) through transcription data from growth of the bacterium on xylans as a sole carbon source. The two PULs encode a number of putative enzymes that are predicted to possess xylanase/xylosidase activities. Along with the CAZymes identified were genes sharing significant sequence identity with SusC-like and SusD-like protein families, indicating that the loci were indeed

PULs (Martens *et al.*, 2011). The Xylan utilisation system in *B. ovatus* is organised into two distinct loci (a small PUL defined as PUL-XylS, and a large locus termed PUL-XylL) (Figure 3.2), which are upregulated independently by different structures found in xylans. Exposure to more complex substrates, CX or WX, leads to upregulation of PUL-XylL consistent with the wide range of xylan degrading activities encoding by this locus (Figure 3.2a). Exposure to simpler xylans with fewer decorations upregulate PUL-XylS, which encodes a limited number of enzymes (Figure 3.2b) (Rogowski *et al.*, 2015).

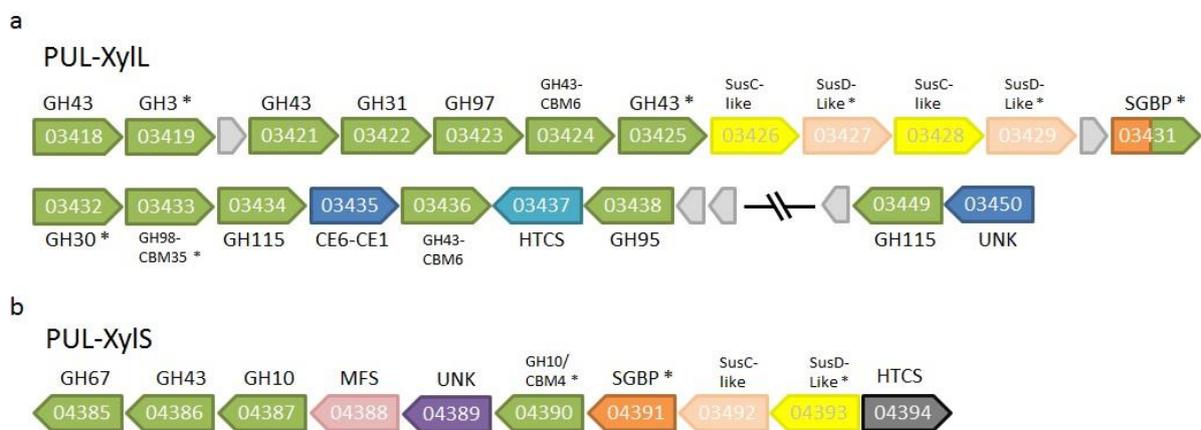


Figure 3.2 Schematic of the *B. ovatus* PULs. Schematic of the large xylan PUL (a) and the small xylan PUL (b) of *B. ovatus*. Genes are drawn to scale with the arrow head indicating gene orientation. Numbers below each gene correspond to their locus tag (*bacova_XXXXX*). Where possible genes are colour coded to known or predicted function and, where appropriate, are also annotated with their CAZY family number : glycoside hydrolase (GH, green), carbohydrate esterase (CE, purple), carbohydrate-binding module (CBM), surface binding proteins (SGBP, orange), unknown with distant similarity to CE6 carbohydrate esterase family (UNK, purple), SusD-homologues (light tan), SusC homologues (yellow), hybrid two component system (HTCS, light/dark blue), transporter of the major facilitator sub family (MFS, pink) and unknown function (grey). Proteins located at the cell surface are marked with an asterisk (*). The defining feature of PULs are the presence of *susCD* homologue pairs which encode specific binding and transport proteins in the *Bacteroides* outer membrane. Based on a figure from Rogowski *et al.* (2014).

PUL-XylS encodes two enzymes belonging to glycoside hydrolase family (GH) 10. The vast majority of characterized GH10 enzymes were shown to be endo- β -1,4-xylanases although the family also contains some endo- β -1,3-xylanases (Hernandez *et al.*, 2008). As with all clan GH-A families, GH10 enzymes exhibit a retaining mechanism that utilise a pair of glutamic acid residues as the catalytic nucleophile and general acid/base, respectively. The general acid/base is typically preceded by an

asparagine in a NEP motif. Three-dimensional structures of GH10 enzymes show a classical $(\alpha/\beta)_8$ TIM barrel fold. GH30 enzymes are also in clan GH-A, and hence have many traits in common with GH10, including the same structural fold and catalytic residues and mechanism of action (Naumoff, 2011). Several activities are attributed to GH30 enzymes including glucuronoxylanases. Many GH30 enzymes were previously characterised as GH5s but were reassigned to GH30 based on sequence analysis and tertiary structure analysis (St John *et al.*, 2010).

The pair of GH10 enzymes present in the PUL-XylS that are likely to be endo-acting xylanases. Detailed analysis of these enzymes affords the opportunity to study the adaptations imposed on CAZymes expressed by human gut bacteria that maximise glycan degradation and subsequent utilisation. For efficient glycan utilisation *B. ovatus* must degrade polysaccharides at the cell surface that minimise the loss of oligosaccharides into the environment. To achieve this the surface GH10 xylanase, BACOVA_04390, should display relatively low activity. The corresponding periplasmic GH10 enzyme, BACOVA_04387, is predicted to rapidly degrade oligosaccharides as the resultant sugars would be contained within the bacterial cell.

The range of different xylan structures utilised for growth by *B. ovatus* allows study of the different strategies employed by the bacterium to effectively utilise related glycans of varying complexity. Resource allocation during glycan utilisation is particularly important for gut bacteria which exist within a highly complex community as part of the human gut microbiota (HGM). Release of oligosaccharides into the environment can lead to the enrichment of certain bacterial species if the bacteria can utilise the oligosaccharides, and hence shape the HGM.

3.1.2 Aims

- To characterise endo-acting glycoside hydrolases of the *B. ovatus* xylan degrading PULs including a pair of GH10 enzymes, one predicted to be located at the cell surface and one in the periplasm.
- To investigate oligosaccharide release at the cell surface during growth of *B. ovatus* on a range of different xylan structures.
- To evaluate whether xylooligosaccharides, released at the cell surface during *B. ovatus* xylan utilisation, are capable of sustaining growth of a second bacterium, which exclusively uses oligosaccharides as growth substrate.

3.2 Results

3.2.1 Expression and purification of BACOVA_04390, BACOVA_04387, BACOVA_03432 and BACOVA_03433

The primary sequence of BACOVA_04390, encoded by *B. ovatus* PUL-XylS includes a GH10 domain bisected by the insertion of two Carbohydrate Binding Module (CBM) distantly related to CBM family 4 (Figure 3.4). The locus also encodes a second GH10 enzyme, BACOVA_04387, which has a simpler structure lacking any accessory modules. PUL-XylS encodes two enzymes that play a role in degrading the xylan backbone. These enzymes, BACOVA_03432 and BACOVA_03433, lack any accessory domains/modules and are members of GH30 and GH98, respectively. Analysis of the N-terminal secretion sequence of BACOVA_04390, BACOVA_03432 and BACOVA_03433, using LipoP 2.0 (<http://www.cbs.dtu.dk/services/LipoP/>), revealed type II signal peptides (See appendix A.1 for full explanation of LipoP data) indicating the proteins were most likely lipoproteins attached to the outer membrane of *B. ovatus* via a covalent bond with cysteine (Juncker *et al.*, 2003). Signal sequence analysis of BACOVA_04387 (LipoP 2.0) shows a likely type I signal sequence, indicating localisation to the periplasm (Juncker *et al.*, 2003). Truncated *bacova_04390*, *bacova_04387*, *bacova_03432* and *bacova_03433* encoding mature BACOVA_04390, BACOVA_04387, BACOVA_03432 and BACOVA_03433, respectively, (lacking the N-terminal 20 residue signal peptide) were cloned by Hongjun Zheng into the *Escherichia coli* expression vector pET21a using appropriate restriction sites such that translation and transcription was initiated by vector sequence, which also supplied the recombinant protein with a C-terminal His₆-tag.

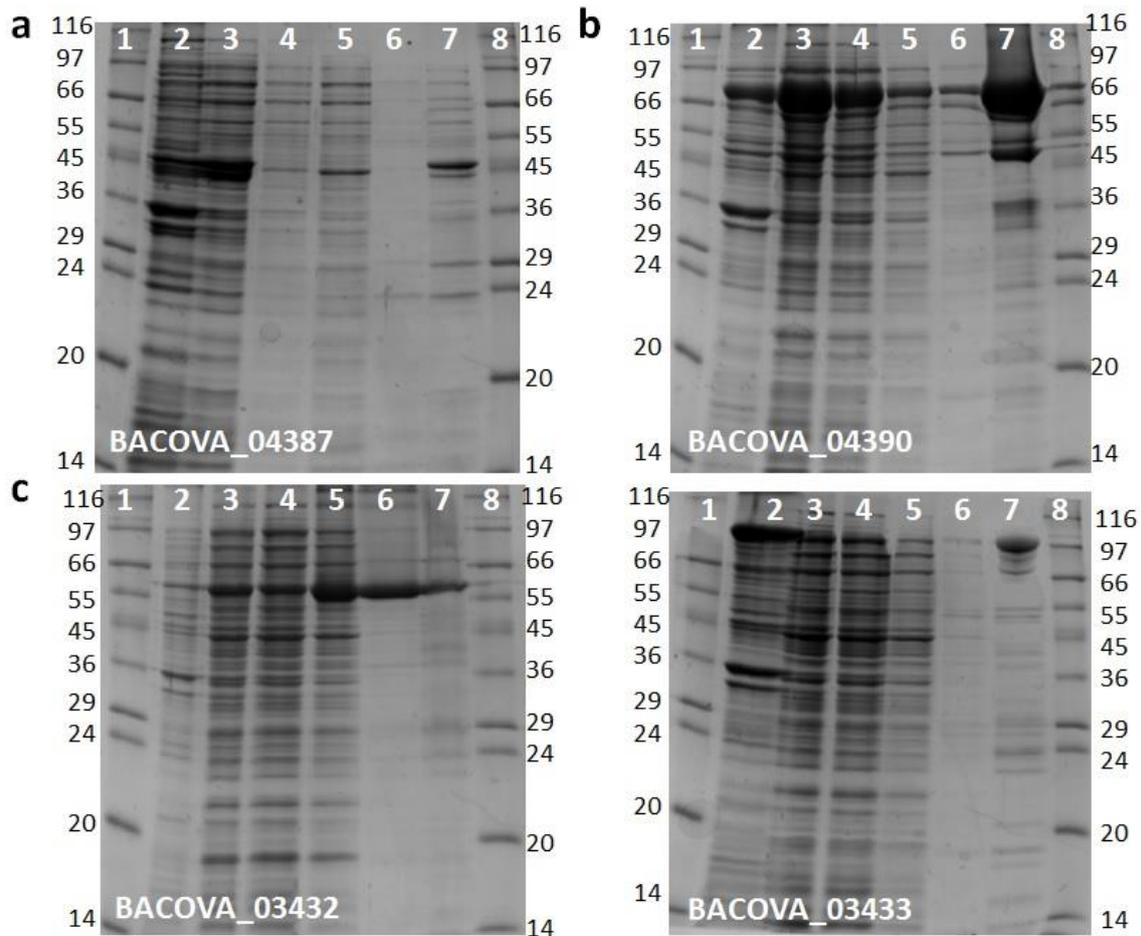


Figure 3.3 SDS-PAGE Protein gels of Xylanase expression and TALON purification elution fraction. BL21 *E. coli* were transformed with vectors containing copies of xylanase genes from the large and small *B. ovatus* xylan PULs for inducible overexpression with IPTG. Recombinant proteins were purified from cell lysate using fused His-tag and TALON IMAC columns. Proteins are eluted from the cobalt columns by washing with TALON buffer with 5 mM or 100 mM imidazole. Elution fractions of BACOVA_04387 (a), BACOVA_04390 (b), BACOVA_03432 (c) and BACOVA_03433 (d) are shown. Samples are run on 12 % (v/v) agarose gels. Lanes are, 1 and 8) Molecular weight markers, 2) Cell pellet, 3) Cell lysate, 4) flow through, 5) TALON buffer wash, 6) 5 mM imidazole elution, 7) 100 mM imidazole elution.

BACOVA_04390, BACOVA_04387 and BACOVA_03432 were expressed in appropriate *E. coli* strains BL21, while BACOVA_03433 was expressed in *E. coli* TUNER cells and purified by Immobilised Metal-Ion Affinity Chromatography (IMAC) using a TALON-cobalt resin. Proteins with a fused His₆-tag are bound by immobilised cobalt ions within the TALON resin. When washed with higher concentrations of imidazole, imidazole mimics histidine ring structure and competes with the protein for binding sites in the TALON resin, hence eluting the bound protein (Lilius *et al.*, 1991). The recombinant

proteins were eluted from the cobalt column by competition with successive increasing concentrations of imidazole. Fractions were evaluated by SDS-PAGE using 12 % (w/v) polysaccharide gels (Figure 3.3). By comparison to known standards a band was identified at 45 kDa (Figure 3.3a), 64 kDa (Figure 3.3b), 58 kDa (Figure 3.3c) and 102 kDa (Figure 3.3d), which corresponds to the theoretical mass of BACOVA_04387, BACOVA_04390, BACOVA_03432 and BACOVA_03433, respectively. The elution with highest degree of purity was selected for each recombinant protein and dialysed into 20 mM sodium phosphate buffer, pH 7.5, containing 150 mM NaCl, which was used in all subsequent assays.



Figure 3.4 Modular structure of BACOVA_04387 and BACOVA_04390. The contrasting domain structure of BACOVA_04387 and BACOVA_04390, green and red represent glycoside hydrolase and CBM (carbohydrate binding module) modules, respectively, as determined by pfam v29.0 (<http://pfam.xfam.org/>).

3.2.2 BACOVA_04387 and BACOVA_04390

The PUL-XylIS, BACOVA_04387 and BACOVA_04390 encoded proteins both belong to GH10 and most likely display endo-xylanase activity. Although such redundancy is not usually observed in glycan, in *Bacteroides* this process occurs in different locations (outer membrane and periplasm), hence the cellular context of proteins in the system can dictate activity or function.

3.2.2.1 GH10 Product Profiles

Initial activity tests and product profiles of BACOVA_04390 and BACOVA_04387 were evaluated using TLC to separate xylooligosaccharides which were identified by their chromatographic migration compared to the standards, xylose (X1), xylobiose (X2), xylotriose (X3), xylotetraose (X4) and xylopentaose (X5). BACOVA_04390 and BACOVA_04387 both showed activity on BX and WX but not

CX (Figure 3.5a,c). BACOVA_04390 initially generated only long oligosaccharides seen as a smear from the origin (Figure 3.5a). In contrast, BACOVA_04387 produced smaller products ranging from X2-X5 (Figure 3.5c). After 16 h BACOVA_04387 generated mostly X2 some xylose and X3 as limit products (i.e the terminal reaction products that were not further degraded). BACOVA_04390 generated limit products after 16 h, which comprised X3-5 with some X2 present (Figure 3.5a). Both BACOVA_04387 and BACOVA_04390 were active on WX although less of the substrate was assessable, as seen by intense spots at the origin, indicative of high molecular weight oligosaccharides (Figure 3.5a,c).

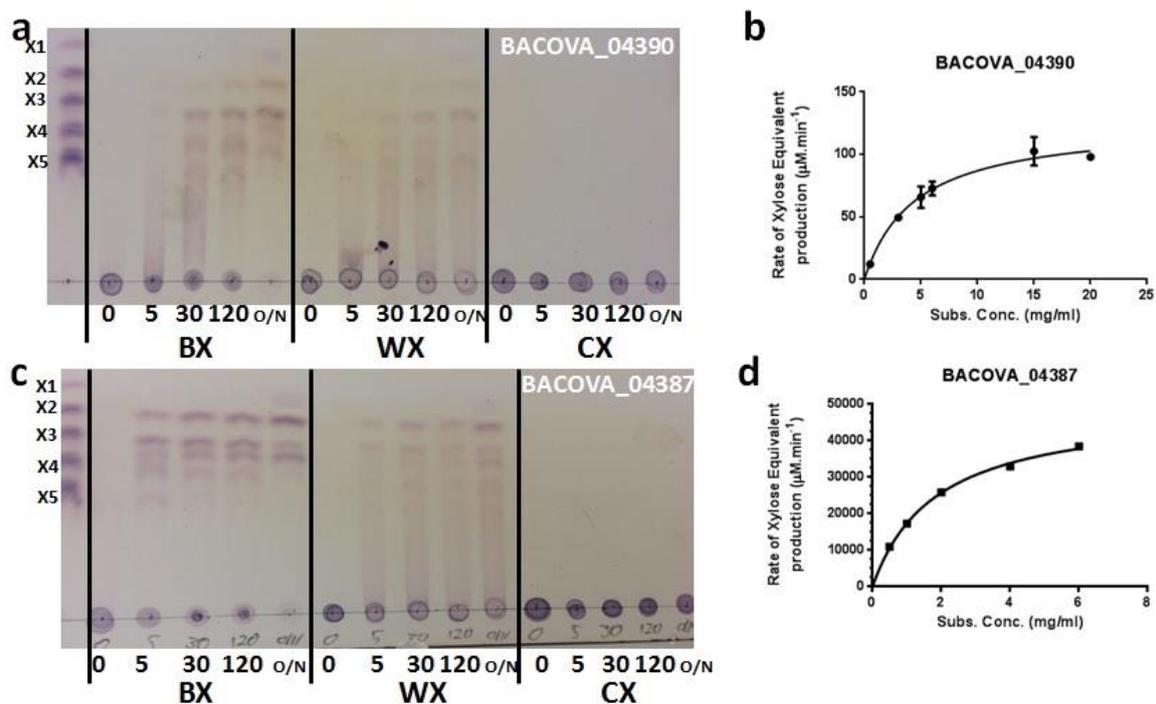


Figure 3.5 Activity of BACOVA_04387 and BACOVA_04390. TLC of 1 mg/ml BX, WX and CX digestion with 1 μM BACOVA_04390 (a) and BACOVA_04387 (c) over a time-course 0-120 min and 16 h (O/N), samples were boiled to inactivate the enzymes. Michaelis-Menten plots of BACOVA_04390 (b) and BACOVA_04387 (d) activity on BX, as determined by DNSA reducing sugar assay were performed at 37 °C in 20 mM sodium phosphate 150 mM NaCl pH 7.5.

3.2.2.2 BACOVA_04390 and BACOVA_04387 Activity on BX

Activity of BACOVA_04390 and BACOVA_04387 was measured using DNSA assays to quantify the formation of terminal reducing sugars, indicative of glycosidic bond cleavage. The kinetic data

showed significant differences in activity of the two GH10 enzymes on BX (Figure 3.5b,d). BACOVA_04390 gave a catalytic efficiency (k_{cat}/K_M) of 280 $\text{min}^{-1} \text{mg}^{-1} \text{ml}$, while the corresponding kinetic value, 8777 $\text{min}^{-1} \text{mg}^{-1} \text{ml}$, is much higher for BACOVA_04387. The 30-fold difference in activity reflects differing functions of the GH10 enzymes in the utilisation of xylan by *B. ovatus*. BACOVA_04390 demonstrates a much higher K_M than BACOVA_04387 on BX while only minor differences in k_{cat} are observed, indicating the vast difference in catalytic efficiency between the two GH10 enzymes stems from the requirement of BACOVA_04390 from much higher substrate concentration for activity (Table 3.1).

Enzyme	Vmax ($\mu\text{M min}^{-1}$)	K_M (mg ml^{-1})	k_{cat} (min^{-1})	k_{cat}/K_M ($\text{ml mg}^{-1} \text{min}^{-1}$)
BACOVA_04390	125.8 \pm 6.1	4.5 \pm 0.6	1258 \pm 61	280
BACOVA_04387	61.6 \pm 2.0	0.7 \pm 0.09	6158 \pm 203	8777

Table 3.1 BACOVA_04390 and BACOVA_04387 activity on BX. Reactions were performed in triplicate.

3.2.2.3 HPAEC analysis of GH10 activity on Xylooligosaccharides

HPAEC was used to determine activity of BACOVA_04390 and BACOVA_04387 on xylooligosaccharides X2 to X6. In each reaction 50 μM of the oligosaccharide was used and enzyme concentration was varied to give optimal degradation in the time frame of the reaction. As the substrate concentration used in the reactions was most likely below the K_M of the enzymes (Appendix A.2) the exponential loss of the substrate was used to determine catalytic efficiency (k_{cat}/K_M).

Enzyme	Substrate	Enzyme Concentration (mM)	k_{cat}/K_M ($\text{min}^{-1} \text{mM}^{-1}$)
BACOVA_04390	X2	0.02	ND ¹
	X3	0.015	NQ ²
	X4	0.01	1.0 ± 0.06
	X5	0.001	12.5 ± 1.1
	X6	0.0002	110.1 ± 19.4
BACOVA_04387	X2	0.02	NQ ²
	X3	0.01	2.9 ± 0.29
	X4	0.0003	509.7 ± 32.7
	X5	0.00005	3368 ± 101.2
	X6	0.00001	3247 ± 392

Table 3.2 BACOVA_04390 and BACOVA_04387 activity on xylooligosaccharides. ¹ND, not detectible; ²NQ, activity too low for reliable quantification, reactions were performed in duplicate due to restricted substrate.

Substrate depletion during the reaction was measured by calculating the area of the peaks that correspond to the oligosaccharide substrate (Figure 3.6a,b). The rate of depletion was calculated using Equation 3.1. Examples of $\ln([S_0]/[S_t])$ versus time plots for BACOVA_04390 and BACOVA_04387 on oligosaccharides X3-6 are given in Figure 3.6c and 3.6d.

$$k \cdot t = \ln\left(\frac{[S_0]}{[S_t]}\right)$$

(Equation 3.1)

Where: $k = k_{cat}/K_M$, t = time, and $[S_0]$ and $[S_t]$ represent the substrate concentration at time 0 and t , respectively. This relationship is only valid when the concentration of enzyme, $[E]$, is \ll than substrate concentration, $[S]$, $\ll K_M$ (Matsui *et al.*, 1991).

3.2.2.4 Activity of BACOVA_04390 and BACOVA_04387 on Xylooligosaccharides

BACOVA_04390 demonstrated preference for longer xylooligosaccharides with no evidence of plateauing of activity up to X6, suggesting that the substrate binding cleft has the capacity to bind glycans with a DP >6 (Figure 3.6c, Table 3.2). BACOVA_04387 was able to hydrolyse substrates X2-X6 (Table 3.2), However an exact rate could not be determined on X2 as activity was extremely low. K_{cat}/K_M of BACOVA_04387 increased 175-fold from X3 to X4, however activity plateaued at X5, (Table 3.2 and Figure 3.6d), indicating the presence of five xylose binding subsites. BACOVA_04387 showed

activity on smaller oligosaccharides than BACOVA_04390, consistent with its periplasmic location and thus its role in the degradative process downstream of the surface xylanase.

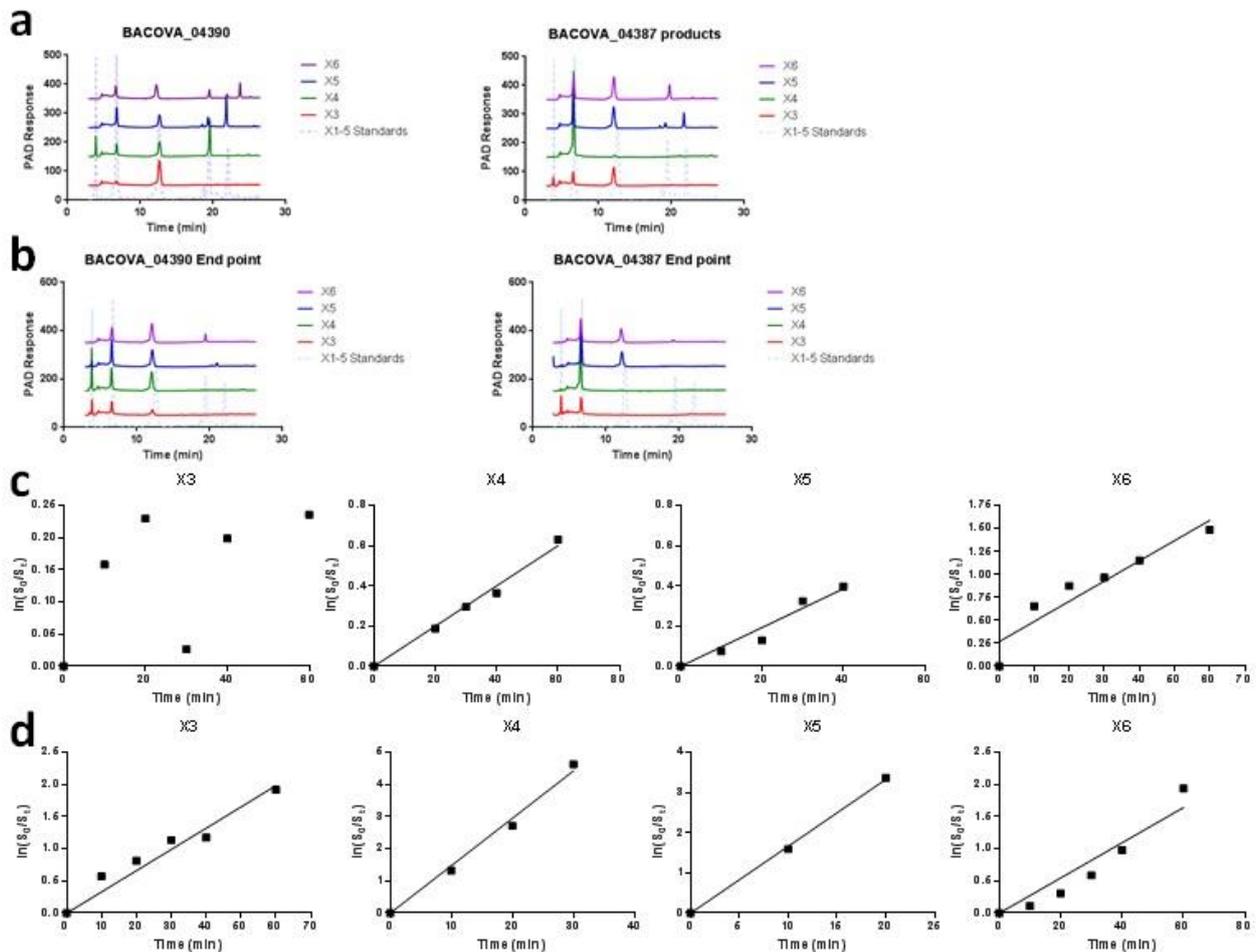


Figure 3.6 HPAEC product profiles and rates of activity of BACOVA_04390 and BACOVA_04387 on xylooligosaccharides. Product profiles of BACOVA_04390 and BACOVA_04387 degradation of xylooligosaccharides X3-6 mid-reaction. Peaks were determined by comparison to known standards (light blue dashed line), peaks (left to right) X1, X2, X3, X4 and X5 (a). End products of xylooligosaccharide digestion by BACOVA_04390 and BACOVA_04387, samples were incubated overnight to give final products of xylanase digestion and run on HPAEC (b). BACOVA_04390 (c) and BACOVA_04387 (d) activity on xylooligosaccharides X3-6 as determined by substrate depletion. Rates are determined by integration of the rate of change in substrate concentration from 0 min to each time point giving a positive rate of substrate depletion. Reactions were performed in 20 mM Sodium Phosphate 150 mM NaCl pH 7.5, enzyme in each sample was inactivated by boiling.

3.2.2.5 Product profiles of BACOVA_04390 and BACOVA_04387 during Xylooligosaccharide hydrolysis

BACOVA_04387 generated xylose and X2 in equal amounts from X3 (Figure 3.6a). X4 hydrolysis resulted in only X2, while cleavage of X5 produced mostly X2 and X3. X2, X3 and X4 were the major products of xylohexaose (X6) hydrolysis (Figure 3.6a & b). The dominance of X2 production from X4 indicates that binding at the +2 subsite is much tighter than at -3, as a dominant role for the distal negative subsite would have generated X3 and X1. Thus, the five subsites in BACOVA_04387 could be distributed -3 to +2 or -2 to +3 (Figure 3.7). Unfortunately, as the oligosaccharides were not labelled at the reducing end it is not possible to determine which of the proposed subsite topologies is correct.

BACOVA_04390 limit products of the oligosaccharide digests were mostly X2 and X3 with some xylose (Figure 3.6b). Mid-digestion products of X4 included mainly xylose and X3 with small amounts of X2 (Figure 3.6a). These data demonstrate a strong involvement of the -3 subsite in productive substrate binding (Figure 3.6a). Intermediate digestion of both X5 and X6 generated significant quantities of X4 indicating a functional -4 subsite in the case of the pentaose, while the hexaose degradation pattern may indicate a functional -4 and/or +4 subsite. As stated above the lack of labelled oligosaccharides prevented the evaluation of these possibilities, although the use of mass spectrometry with O¹⁸ water could have resolved the binding mode (McGregor *et al.* 2016).

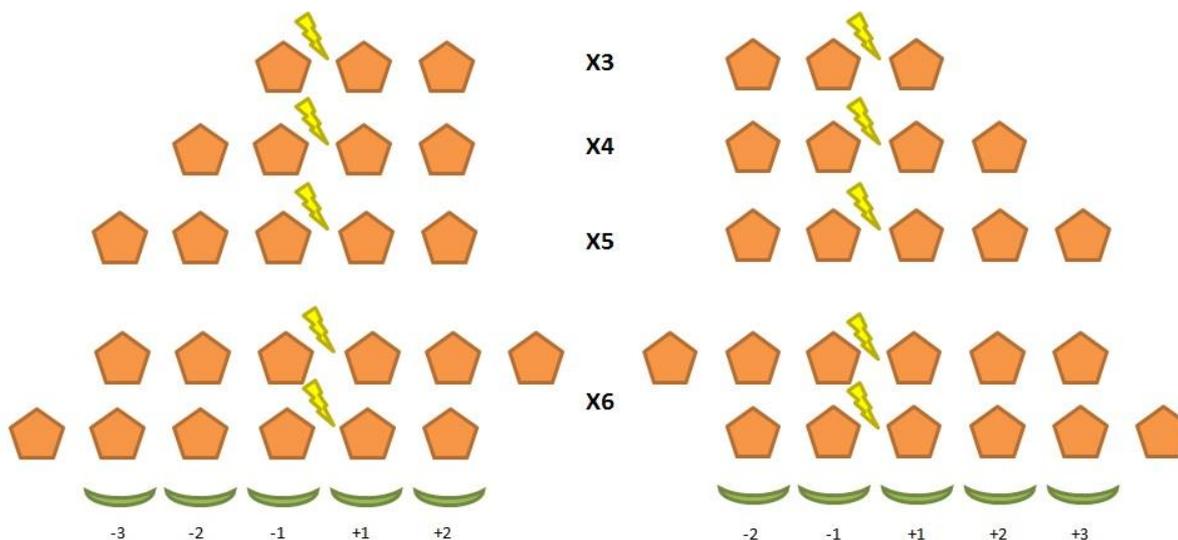


Figure 3.7 Preferential binding modes of oligosaccharides in BACOVA_04387 active site. A schematic representation of preferential productive binding of xylooligosaccharides in the active site of BACOVA_04387 as deduced by relative abundance of products of defined oligosaccharide catalysis using HPAEC to detect products. Pentagons = xylose in oligosaccharides (labelled X3, X4, X5 and X6), crescents = subsites (labelled below), lightning bolt = where in the oligosaccharide hydrolysis occurs.

Product profiles can be used along with k_{cat}/K_M to calculate the contribution to substrate binding of some individual subsites. Here, comparison of the cleavage of X4 and X3 enables binding energy at +2 to be calculated, assuming X3 binds productively from subsites -2 to +1 and X4 from -2 to +2. Thus the difference in activity against the two oligosaccharides reflects binding at the +2 subsite in the case of X4 (Figure 3.6b). Binding energy is calculated using Equation 3.2.

$$\Delta G(kcal/mol) = R.T.\ln\left(\frac{K_{cat}/K_M(x_a).BCF(x_a)^b}{K_{cat}/K_M(x_{a-1}).BCF(x_a)^b}\right)/c/d$$

(Equation 3.2)

Where: $K_{cat}/K_M(x_a) = K_{cat}/K_M$ for a xylooligosaccharide of dp a . $BCF(x_a)^b$ = the bond cleavage frequency for glycosidic bond b of a xylooligosaccharide of dp a . R = the gas constant 8.314. T = temperature (Kelvin). $c = 4.184$, (1 cal = 4.184). $d = 1000$ (converting cal to kcal).

3.2.3 BACOVA_03432

bacova_03432 is within PUL-XylL of *B. ovatus*. BACOVA_03432 showed significant identity to GH30 glucuronoxylanases exemplified by the glucuronoxylan endo-1, 4-beta-xylanase from *Erwinia chrsanthemi* (Urbanikova *et al.*, 2011).

3.2.3.1 BACOVA_03432 Activity

BACOVA_03432 was incubated with 1 mg/ml BX, WX and CX at 37 °C for 16 h. TLC analysis revealed that a series of high molecular weight oligosaccharides were generated from BX but not from WX or CX (Figure 3.8a). As only BX contains GlcA side chains, these data are consistent with the proposal that BACOVA_03432 is a glucuronoxylanase. The activity of the BACOVA_03432 on BX, a glucuronoxylan, was quantified using DNSA reducing sugar assay. Rates from this assay were all at the linear phase of the Michaelis-Menten plot, indicating the substrate concentrations used were below the K_M on this substrate (Figure 3.8b). The data revealed a k_{cat}/K_M value of 47.93 $\text{min}^{-1} \text{mg}^{-1} \text{ml}$.

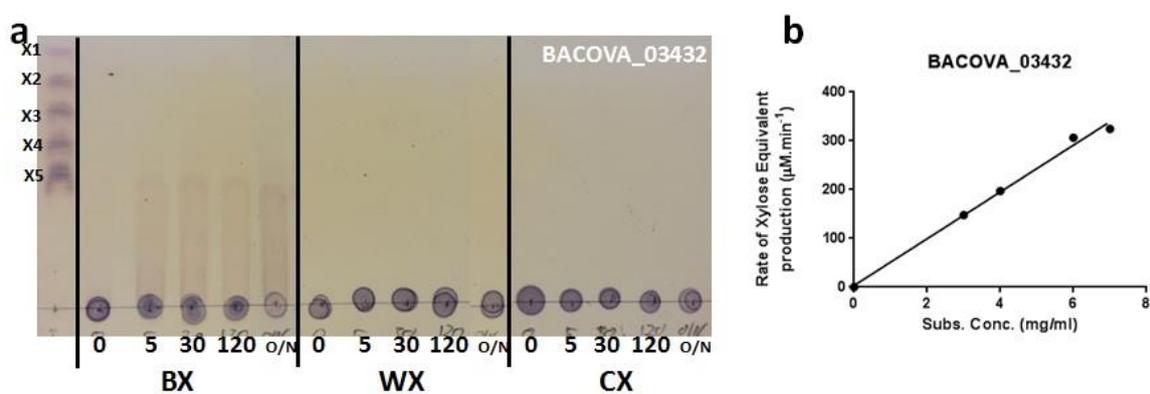


Figure 3.8 Activity of BACOVA_03432. TLCs of BACOVA_03432 product profiles on BX, WX and CX using 0.5 μM enzyme and 1 mg/ml substrate, aliquots were taken at 0, 5, 30 120 min and after an overnight incubation (O/N) and boiled prior to TLC (a). Rate of BACOVA_03432 activity on BX deduced by DNSA reducing sugar assay (b). Assays were performed at 37 °C in 20 mM sodium phosphate 150 mM NaCl pH 7.5.

3.2.4 BACOVA_03433

Present in PUL-XyLL was a gene showing significant identity to GH98 enzymes, a family associated with endo- β -galactosidases targeting mammalian cell antigens (Anderson *et al.*, 2005; Shaikh *et al.*, 2009).

3.2.4.1 BACOVA_03433 Activity

The purified recombinant 102 kDa enzyme was assayed against CX, BX and WX at 37 °C. A smear of high molecular weight oligosaccharides were generated from CX but no activity against BX or WX

was detected (Figure 3.9a). The poorly defined CX-derived oligosaccharides reflects the highly complex and variable structure of the side chains that decorate the xylan backbone of this polysaccharide. These data show that BACOVA_03433 utilizes specific side chain(s) in CX, which are absent in BX and WX, as critical specificity determinants.

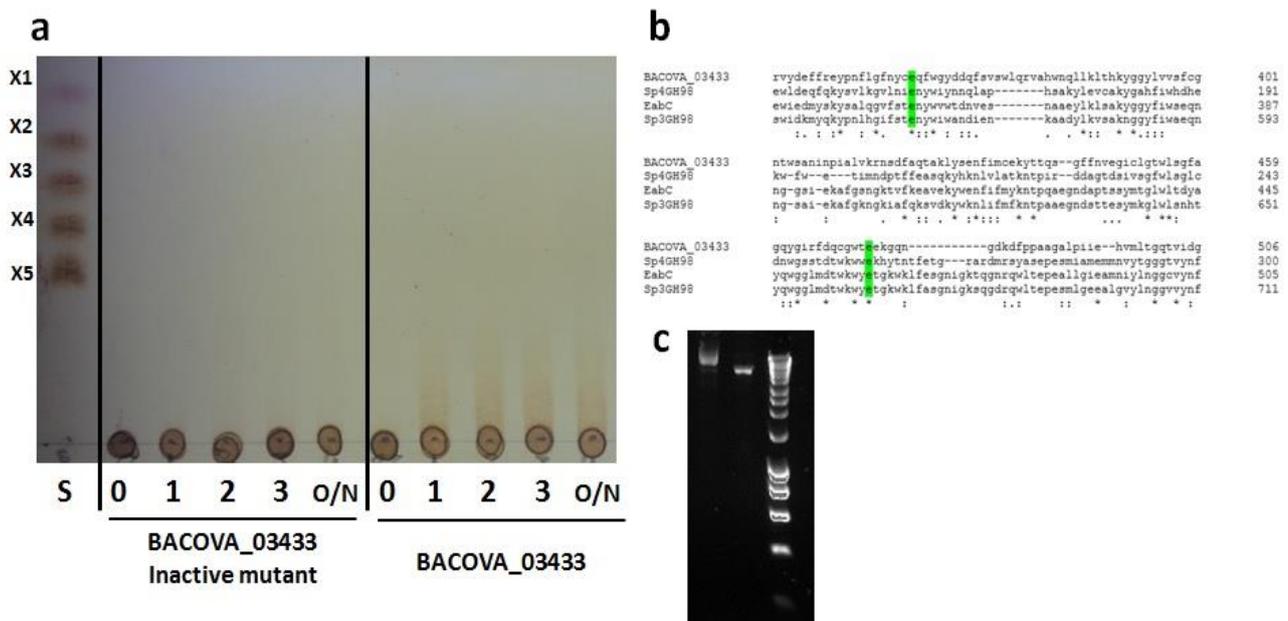


Figure 3.9 Activity of BACOVA_03433 and BACOVA_03433 inactive mutant on CX. TLC of BACOVA_03432 product profiles on CX using 1 μ M enzyme and 0.5 mg/ml substrate, aliquots were taken at 0, 1, 2, 3 h and after an overnight incubation (O/N) and boiled prior to TLC. The BACOVA_03433 inactive mutant is a double substitution mutant of the catalytic residues, E361 and D467 to alanine. Assays were performed at 37 $^{\circ}$ C in 20 mM sodium phosphate 150 mM NaCl pH 7.5 (a). Identification of BACOVA_03433 catalytic residues by sequence alignment against Sp4GH98 and Sp3GH98 from *Streptococcus pneumoniae* and EabC from *Clostridium perfringens* ATCC 10543, with catalytic residues highlighted in green (b). Agarose gel (0.8%) of the mutagenesis reaction mixture before and after DpnI digest (c).

Through sequence alignment of BACOVA_03433 with previously characterised GH98 enzymes potential catalytic residues were identified as E361 and D467 (Figure 3.9b) by alignment with Sp3GH98 and Sp4GH98 from *Streptococcus pneumoniae* (Higgins *et al.*, 2009) and EabC from *Clostridium perfringens* (Shaikh *et al.*, 2009). These mutations were introduced into the coding sequence of *bacova_03433-pet21a* by site-directed mutagenesis (Chapter 2.5.6) using PCR with primers with the desired mutations and sequences homologous to the region surrounding the

targeted bases. The reaction mixture was incubated with DpnI a restriction nuclease that selectively degrades methylated DNA which, in this reaction, was the template DNA (Figure 3.9c). The mixture was used to transform TOP10 *E. coli* cells and positive mutations were selected by sequencing the resulting *mut-bacova_03433-pet21a*. The resultant E361A and D467A mutants of the GH98 enzyme were unable to hydrolyse 0.5 mg/ml CX at 1 μ M enzyme at 37 °C in 16 h reactions (Figure 3.9a). These data confirm that E361 and D467 are the catalytic residues of BACOVA_03433.

3.2.5 *B. ovatus* xylan Cross-feeding

B. ovatus xylan utilisation gives an opportunity to study the extent of cross-feeding during utilisation of xylan polysaccharides with a range of xylan structures. Cross-feeding has been previously reported among different *Bacteroides* species (Rakoff-Nahoum *et al.*, 2014; Rakoff-Nahoum *et al.*, 2016) between *Bacteroides* and other members of the HGM (Van der Meulen *et al.*, 2006) but never observed during growth on xylan polysaccharides. Other previously studied utilisation systems either demonstrate a selfish strategy, where product utilisation is restricted to the polysaccharide degrading organisms, or an altruistic mechanism in which monosaccharides and/or oligosaccharides released into the growth supernatant are used by other bacteria in the same culture (Rakoff-Nahoum *et al.*, 2014; Cuskin *et al.*, 2015).

3.2.5.1 Supernatant Oligosaccharides

Expression of both xylan degrading PULs during growth allows *B. ovatus* to successfully utilise a range of xylans including BX, WX and CX (Rogowski *et al.*, 2015). Utilisation, however, may not be complete as oligosaccharides generated at the cell surface may diffuse into the culture medium rather than being sequestered by *B. ovatus*. Here, samples of culture were taken to assay for presence of oligosaccharides in the cell supernatant. *B. ovatus* was grown in 5 ml tube cultures with 0.5 % BX, WX or CX as the sole carbon source. At the different phases of aliquots were taken and the culture supernatants were subjected to TLC to evaluate the oligosaccharides released by the bacteria into the media (Figure 3.10). During growth of *B. ovatus* on BX a smear of oligosaccharides

was present at early to mid-exponential phase. More defined oligosaccharides with lower DPs also began to accumulate; these molecules correspond with X4 and an unidentified oligosaccharide (Oligo X) that migrated between X3 and X4 (Figure 3.10a). At late exponential phase the lower molecular weight discrete oligosaccharides were more evident, while the quantity of material at the origin, corresponding to the polysaccharide growth substrate, was greatly reduced. At stationary phase, when growth had ceased, there was no polysaccharide visible at the origin, Oligo X, however, was the only oligosaccharide remaining (Figure 3.10a). Indeed Oligo X was present in stationary culture supernatants for each of the xylans used, indicating that it is a generic feature of this class of hemicelulosic polysaccharide.

There was a similar pattern of oligosaccharides generated by *B. ovatus* grown on WX (Figure 3.10b), with the exception that the oligosaccharide smear appeared during lag phase, and the starting material was still present later in growth. Again, OligoX was apparent in the products.

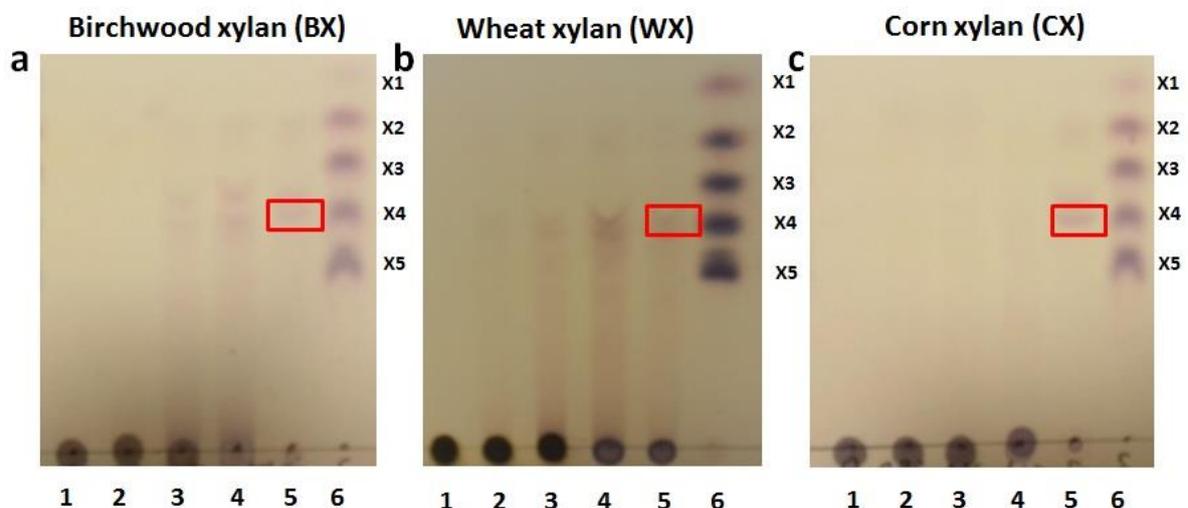


Figure 3.10 TLC showing oligosaccharides released by *B. ovatus* during growth on xylans. TLC of supernatant of *B. ovatus* grown on birchwood xylan (a), wheat xylan (b) and corn xylan (c). Samples were taken during each phase of growth, as determined by OD measurements, cells were removed and the remaining supernatant was blotted onto TLC plates. Each lane corresponds to different phases of growth, 1) Start/inoculation, 2) Lag phase, 3) Early-mid exponential, 4) Mid-late exponential, 5) Stationary and 6) xylooligosaccharide standards. The position of OligoX is indicated by a red box.

The pattern of oligosaccharides present in the CX growth medium was entirely different from the other xylan cultures; only very faint smears of high molecular weight oligosaccharides were present during late exponential phase. At stationary phase OligoX was present, while the starting material was almost completely absent (Figure 3.10c).

Indeed, *detectable* xylooligosaccharides were released into the growth medium during *B. ovatus* utilisation of xylans, including the highly complex CX, although only at low concentrations. The simpler xylans, BX and WX, were utilised in a less stringent manner with higher levels of oligosaccharides being released throughout the growth phases. Unfortunately further investigation using HPAEC-PAD was not possible as cell debris from cell cultures can cause damage to the detection electrode of the HPAEC machine, hence it was decided to just use TLC to visualise supernatant oligosaccharides (Figure 3.10)

3.2.5.2 *Bifidobacterium adolescentis* (*Bi. adolescentis*) growth on xylans and xylooligosaccharides

Released xylooligosaccharides from *B. ovatus* growth on BX, WX and CX could potentially be used by a second bacterium that is able to utilise oligosaccharides derived from the xylans. *Bifidobacterium adolescentis* was identified as a potential xylooligosaccharide user. Here, *Bi. adloescentis* growth was tested on BX WX and CX oligosaccharides, mimicking those released by *B. ovatus*.

Bi. adolescentis was grown overnight in Clostridial media (CM) with hematin to stationary phase. After centrifugation the bacterial pellet was washed and diluted 1 in 3 in PBS. The *Bi. adolescentis* cells were inoculated into BiMM (Bifidobacterium minimal media)(Van der Meulen *et al.*, 2006) with hematin supplemented with 0.5 % BX oligosaccharide, WX oligosaccharide, CX oligosaccharide, and the respective polysaccharides, xylose, arabinose or glucose in separate positions on a microtitre dish. Growth was measured by an automatic plate reader ($A_{600\text{ nm}}$) taking readings every 15 min in an anaerobic cabinet at 37 °C.

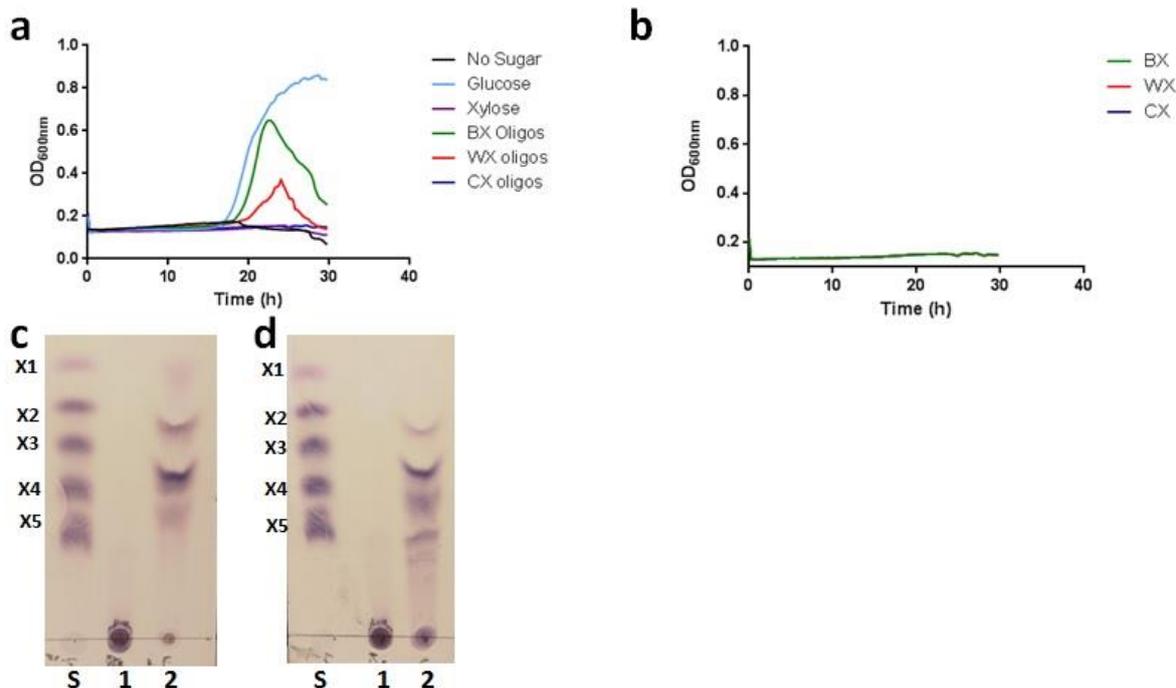


Figure 3.11 TLC of xylan oligosaccharides and resulting *Bi. adolescentis* growth. The xylan oligosaccharides generated were used as a growth substrate for *Bi. adolescentis* in 200 μ l plate reader growths along with glucose, xylan derived oligosaccharides and a no sugar control (a), and on untreated BX, WX and CX (b). *Bi. adolescentis* was grown in BiMM with 0.5 % growth substrate. TLC of xylan oligosaccharides generated from digestion of 10 mg/ml BX (c) and WX (d) with 5 μ M recombinant BACOVA_04390 for 3 h. S = xylooligosaccharide standards X1-5, 1 = 10 mg/ml xylan, 2 = 10 mg/ml xylan with BACOVA_04390.

To test growth on xylooligosaccharides 10 mg/ml BX and WX was digested with 5 μ M

BACOVA_04390 for 3 h to generate oligosaccharides derived from each of the starting xylooligosaccharides (Figure

3.11a), giving a range of oligosaccharide lengths. CX oligosaccharides were made by digestion of 10

mg/ml CX with 2 μ M BACOVA_03433 for 5 h. *Bi. adolescentis* grew on glucose, BX and WX

oligosaccharides (figure 3.11a), however, the bacterium could not be cultured on untreated BX, WX

and CX (figure 3.11b). Growth on BX oligosaccharides was similar to glucose, but had a 1-2 h longer lag phase and only achieved a maximum OD of 0.65 compared to an OD of 0.85 when *Bi.*

adolescentis was cultured on the hexose sugar. Growth on WX oligosaccharides showed a similar lag

phase to the BX oligosaccharide culture and achieved a maximum OD of 0.4 at 24 h. There was no

growth observed for CX oligosaccharides, arabinose or xylose. Maximal growth rate of *Bi.*

adolescentis on BX and WX oligosaccharides was calculated as 0.18 and 0.03 OD.h⁻¹, respectively.

These data confirm that ability of *Bi. adolescentis* to utilise oligosaccharides derived from BX and WX but not those generated from BACOVA_03433 digestion of CX.

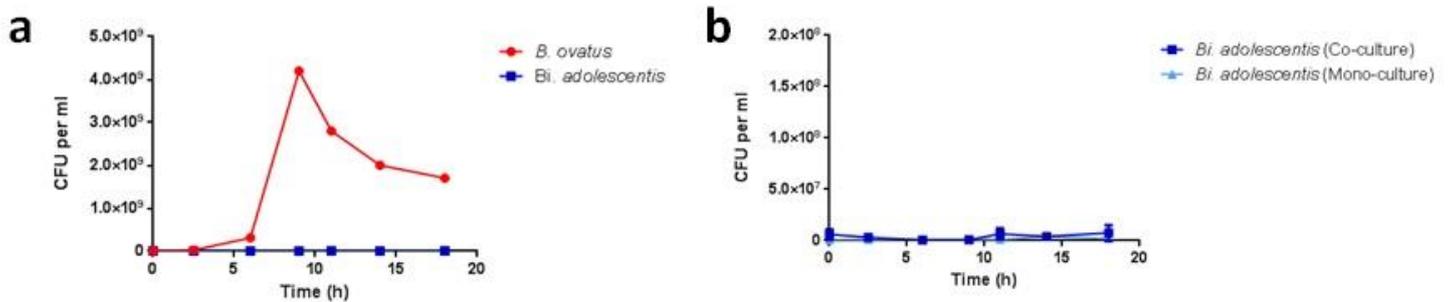


Figure 3.12 Co-culture of *B. ovatus* and *Bi. adolescentis* on xylose. Growth measured by CFU/ml of *B. ovatus* in co-culture with *Bi. adolescentis* (a) and *Bi. adolescentis* in co-culture and mono-culture on BiMM with 0.5 % xylose (b) calculated by colony counts of serial dilution of continuous growths plated on to BHI agar with hematin.

3.2.5.3 *B. ovatus* – *Bi. adolescentis* Crossfeeding on xylans

B. ovatus and *Bi. adolescentis* were grown separately until stationary phase when they were centrifuged, washed in PBS to remove the complex growth medium, and then resuspended in PBS.

The cells suspensions were mixed in a 50:50 ratio and 200 μ l was used to inoculate media with 0.5 % BX, WX and CX. The mixed cells were grown at 37 °C under anaerobic conditions. Samples of the cultures were plated onto CM-agar plates for CFU determination. Optical density (OD) of the cultures were measured over the entire growth phases. Mono-cultures were used as controls for *B. ovatus* and *Bi. adolescentis* on BX, WX and CX. The CFUs for these mono-cultures were also determined. In the mixed cultures species were distinguished on colony morphology and confirmed by re-streaking colonies on to BHI plates with 200 μ g/ml gentamicin (Figure 3.13); *Bi. adolescentis* but not *B. ovatus*, is sensitive to the antibiotic (Lim *et al.*, 1993). *Bi. adolescentis* colonies are differentiated from *B. ovatus* colonies by size, colour and shape of colony margin (Figure 3.13b), where *Bi. aoldescentis* colonies are small and possess a rougher edge than the larger smoother *B. ovatus* colonies (Figure 3.13a-c).

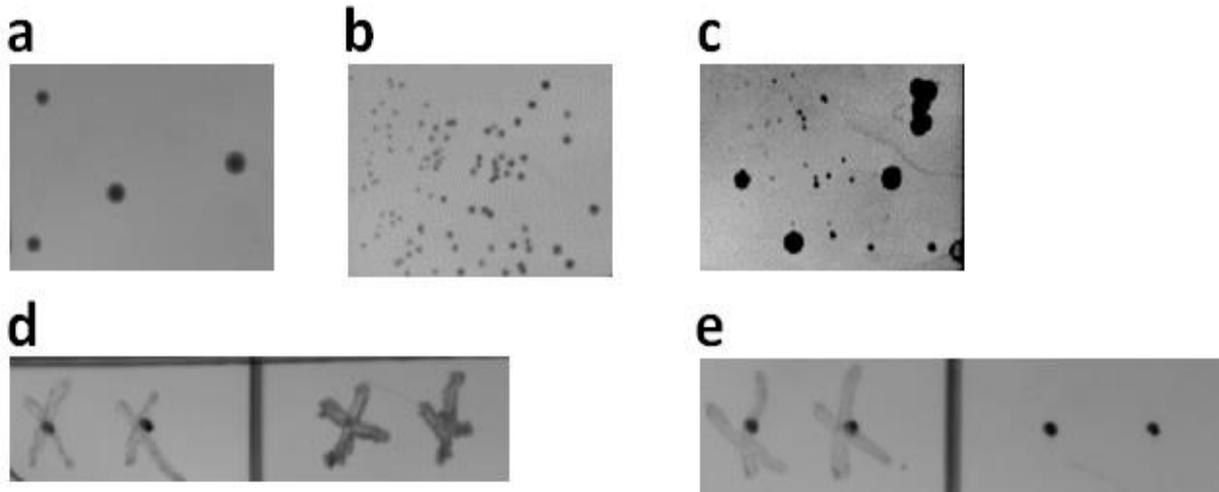


Figure 3.13 Selection of *Bi. adolescentis* and *B. ovatus* during colony counts. Colonies of only *Bi. adolescentis* (a) and *B. ovatus* (b) grown on CM-Agar. Mixed inoculum of *B. ovatus* and *Bi. adolescentis* grown in CM media to late exponential phase and plated onto CM-agar to demonstrate difference in colony morphology between the two bacteria (c). To confirm identity, many colonies from each plated sample were picked and used to inoculate both CM-agar plates without (d) and with (e) 200 µg/ml gentamycin, where *B. ovatus* was spread as Xs on the left side while *Bi. adolescentis* was on the right of panels d and e, 'dots' are pen markers to help plate inoculation.

In mono-culture *Bi. adolescentis* was unable to utilise xylose as a growth substrate, however *B. ovatus* metabolises the free monosaccharide. To ensure growth of *Bi. adolescentis* in the polysaccharide co-cultures was dependent on oligosaccharides released into the media by *B. ovatus*, not other metabolites or the media itself, a co-culture was performed using xylose as the growth substrate. In the xylose mixed cultures *B. ovatus* grew to 4.1×10^9 CFU (Figure 3.12a), while no increase in the number of *Bi. adolescentis* cells was observed (Figure 3.12b).

In the BX co-culture the ratio of *B. ovatus* to *Bi. adolescentis* began at 60:40 but by the end of log phase 98-99% of the cells were *B. ovatus*. This ratio was maintained throughout log phase and into stationary (Figure 3.14a). Despite this ratio of the two bacterial species, the CFU data showed that the level of *Bi. adolescentis* in the co-culture was 50-fold greater than when the bacterium was grown in mono-culture on BX (Figure 3.14a). A similar growth pattern was observed in the WX co-culture, although with a slightly extended lag phase and *Bi. adolescentis* only achieved a 20-fold increase in CFUs compared to the bacterium in the WX mono-culture (figure 3.14b). The CX co-culture, however, was different. *Bi. adolescentis* in the CX co-culture was similar to the CX mono-culture control. This was also reflected in the ratio of *B. ovatus* to *Bi. adolescentis*, which was

maintained at above 10000:1 throughout exponential phase (Figure 3.14c). *B. ovatus* growth profile in the co-culture was analogous on BX and WX, however, CFU doubled on CX compared to the simpler xylans, despite displaying similar OD readings at equivalent time points (Figure 3.14d).

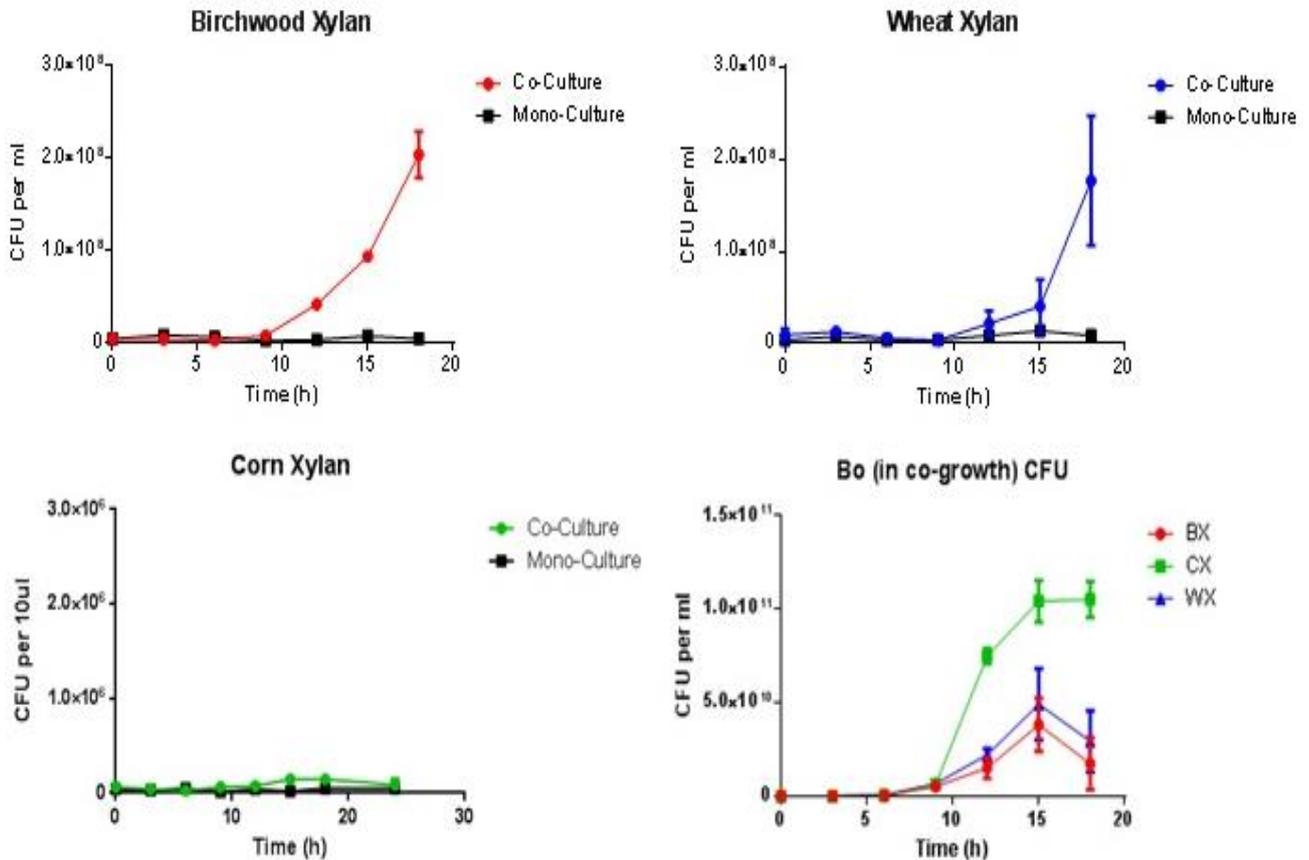


Figure 3.14 Growth curves of *Bi. adolescentis* and *B. ovatus* in xylan co-cultures. Growth of *Bi. adolescentis* in BiMM with 0.5 % BX (a), WX (b) and CX (c) at 37 °C in anaerobic conditions in the same tube as *B. ovatus* (Co-Culture) and without *B. ovatus* (Mono-Culture). Growth of *B. ovatus* in BiMM with 0.5 % BX, WX and CX Co-cultured with *Bi. adolescentis* (d). CFU was calculated using serial dilutions of aliquots of growth culture plated onto BHI agar with heamtin. Colony type was confirmed by replica plating onto BHI with gentamycin.

3.2.4.4 *B. ovatus* – GH98 mutant Cross-feeding

To explore the possibility that there were utilisable CX oligosaccharides produced during *B. ovatus* growth on the xylan, a *B. ovatus* mutant was constructed in which the catalytic apparatus of the key GH98 surface xylanase, BACOVA_03433, was inactivated. The candidate catalytic residues, Glu361 and Asp467, identified by alignments against previously characterised enzymes belonging to GH98 (Figure 3.9b), were mutated to alanine. Initial activity tests were performed on the E361A/D467A

double mutant of BACOVA_03433 (Figure 3.9a). The mutant enzyme was completely inactive on CX, indicating the mutations had the desired result. The wild type copy of *bacova_03433* was exchanged for the mutant of this gene, defined as $\Delta GH98$, encoding an inactive variant of BACOVA_03433.

Figure 3.15 shows schematic diagrams of key stages of the mutagenesis process to create a genomic mutant. *Mut-bacova_03433-pet21a* was used as template of the inactive BACOVA_03433 to generate *mut-bacova_03433* DNA to be cloned into pExchange-tdk. Positive transformants of *E. coli* s17 λ pir cells harbouring *mut-bacova_03433-pexchange-tdk* were grown on BHI-agar plates along with *B. ovatus tdk-* (Figure 3.15a). Aerobic *E. coli* grows first creating an anaerobic environment for *B. ovatus tdk-* to grow, encouraging conjugation and passage of DNA between *E. coli* and *B. ovatus tdk-* (Figure 3.15b). Selection for *B. ovatus tdk-* harbouring *mut-bacova_03433-pexchange-tdk* was performed on BHI-agar with 25 μ g/ml and 200 μ g/ml, gentamycin and erythromycin, respectively. Integration of the target gene sequence (Figure 3.15c) into *B. ovatus tdk-* genome without the vector sequence was selected for by growth on BHI-agar with 200 μ g/ml 5-Fluro-2'-deoxyuridine (FudR). As thymidine kinase (tdk) generates toxic products from FudR, mutants which have lost the vector DNA will grow and should also possess the *mut-bacova_03433* DNA. The resulting colony genomes were then sequenced at the target gene and the flanking regions (Figure 3.15d) to ensure the mutation was successful. The introduction of a gene that encodes an inactive form of the enzyme is preferable to deletion of the cognate gene mutation, as any protein-protein interactions that may contribute to the utilisation apparatus at the cell surface will remain intact. The mutant displayed no growth on MM+ 0.5 % CX, however, the strain grew on CX that had been pre-digested with recombinant wild type BACOVA_03433 (Figure 3.17a). Indeed, the growth profile of $\Delta GH98$ on CX oligosaccharides was analogous to wild type *B. ovatus*, with a slightly elongated lag phase (Figure 3.17a). This is most likely due to the lack of a functioning BACOVA_03433, which was unable to generate oligosaccharides at the cell surface; instead the mutant relies on diffusion for the glycans to bind to the SusCD–import apparatus.

To distinguish between wild type *B. ovatus* and the $\Delta GH98$ variant in co-cultures, a unique signature-tag was introduced into the genome of both strains as described in Chapter 2.9.4. The 24 bp DNA sequences *tag1* and *tag11* (See Appendix A.3) were inserted into a mobile transposon, NBU2 (Wang *et al.*, 2000), within a modified suicide vector (Martens *et al.*, 2008). NBU2 displays homology with the *att1* or *att2* sites within the *B. ovatus* genome, leading to homologous recombination insertion of the transposon along with the relevant nucleotide sequence tag. During growth aliquots were taken and genomic DNA was extracted for relative quantification of signature-tags present in the genomic DNA mixture by qPCR (Koropatkin *et al.*, 2008). Analysis of amplification curves demonstrates differential concentrations of each gDNA sample (Figure 3.16a). Melt curve analysis shows products of amplification were the same length indicating the fluorescence recorded for each sample could be compared directly (Figure 3.16b). Samples diverging from the mean melting temperature were excluded from the data. Samples of Known gDNA concentration were used to give a standard curve to convert quantitation cycle (Cq) values, the cycle in which fluorescence is detected, into relative concentration (Figure 3.16c). The same aliquots were used to determine total CFU during the growth, allowing the ratio of mutant to wild type to be used along with the total CFU to determine the CFU of each in the growth culture given in Figure 3.17b.

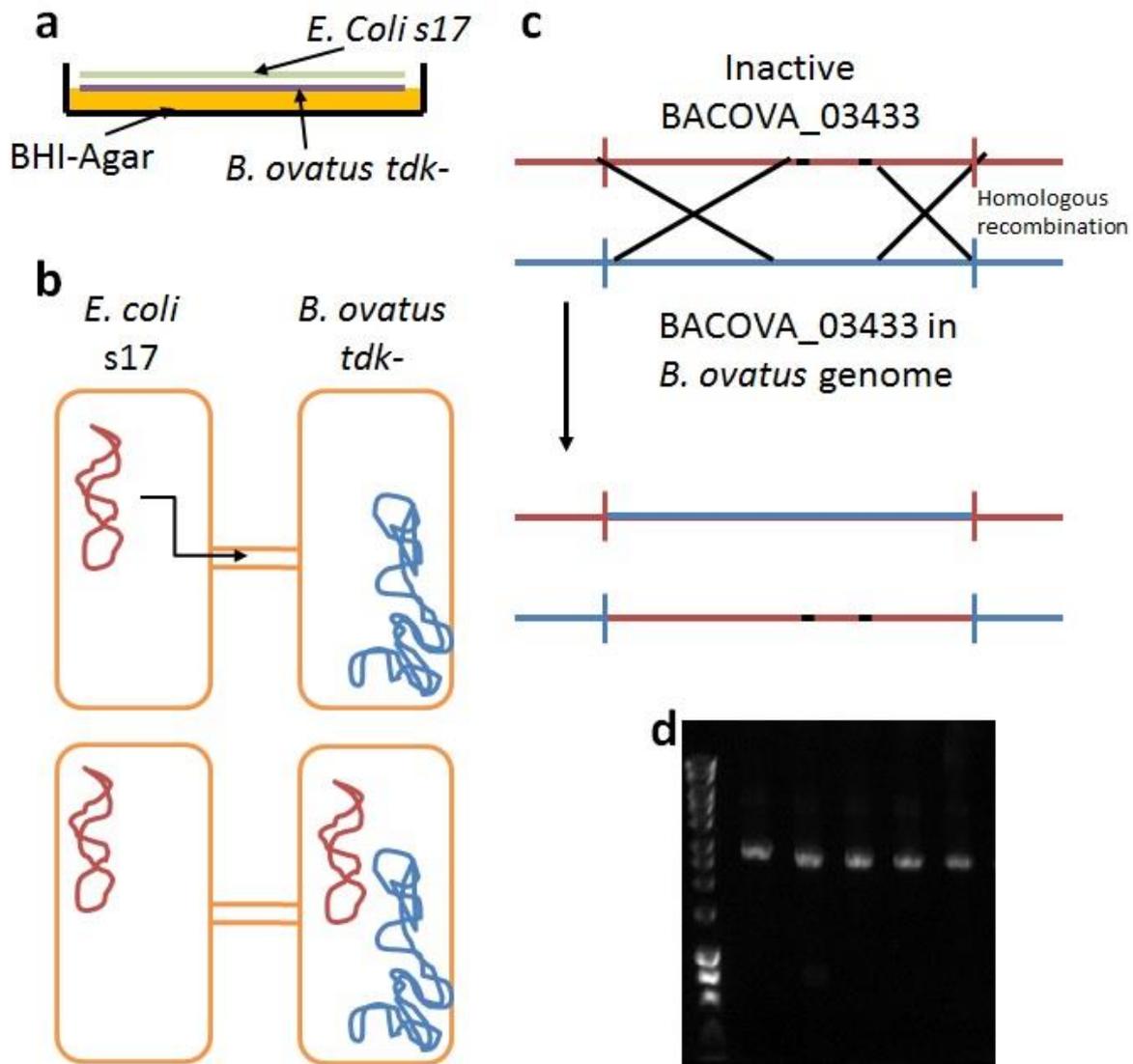


Figure 3.15 *B. ovatus* mutant generation procedure. Diagram of plating technique which allows for conjugation between *E. coli* s17 cells and *B. ovatus* tdk- on BHI-Agar plates (a). Simplified schematic of transmission of pExchange plasmid harbouring mutated *bacova_03433* from an *E. coli* s17 cell to a *B. ovatus* tdk- cell (b). Simplified schematic diagram of crossover events occurring between homologous regions of *mut-bacova_03433-pExchange* and wild type *bacova_03433* present in the *B. ovatus* genome (c). Where, red is pExchange plasmid DNA and blue is *B. ovatus* tdk- genomic DNA. Agarose (0.8 %) electrophoresis gel of PCR products of *bacova_03433*, run alongside known size markers, generated to screen for mutation of *bacova_03433* in *B. ovatus* genome by sequencing (d).

Throughout growth the ratio of wild type *B. ovatus* and $\Delta GH98$ remained around 70:30 (Figure 3.17c). As stated above $\Delta GH98$ was unable to grow on CX in mono-culture; the CFUs remained similar to the inoculation CFU. In the co-culture experiment, however, $\Delta GH98$ grew to 6×10^8 CFUs/ml, a 150-fold increase in the number of cells (Figure 3.17b). These data demonstrate that fermentable CX oligosaccharides were generated and released into the growth media by *B. ovatus*.

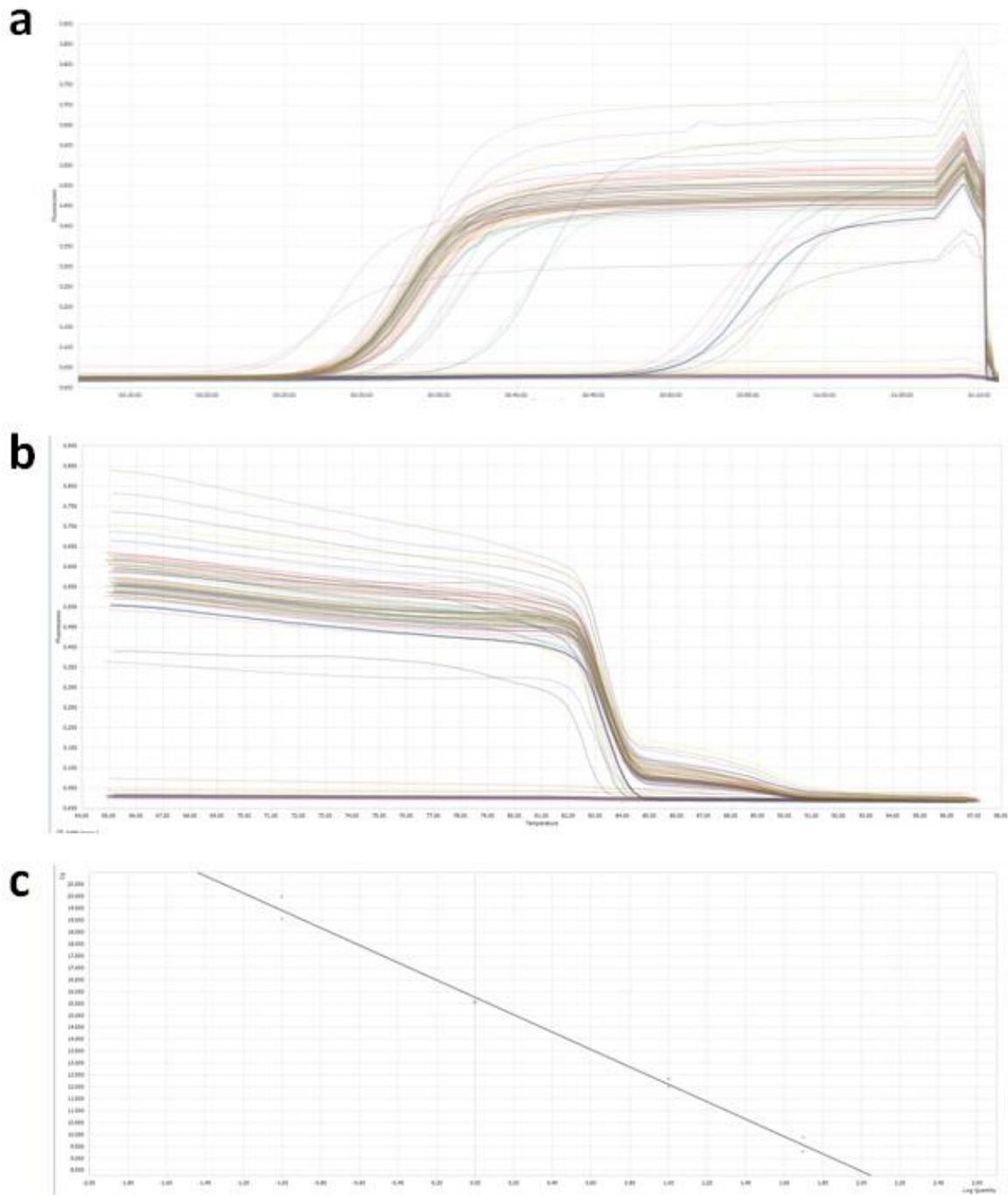


Figure 3.16 Analysis of qPCR data using Roche lightcycler96 software. Amplification curves of fluorescence against cycle number, used to calculate Cq values (a). Melt curve analysis showing homologous product in the reaction (b). Standard curve of known genomic DNA concentrations used to calculate relative concentration of sample DNA (c).

3.3 Discussion

Xylan degradation and utilisation is increasingly being recognised as an important process in the human gut. Current research is amassing evidence of xylan and xylooligosaccharides as prebiotics promoting growth of select members of the gut microbiota associated with a healthy gut (Aachary and Prapulla, 2011; Chapla *et al.*, 2012; Yu *et al.*, 2015).

3.3.1 GH10 xylanases from the small xylan PUL

The data presented in this chapter explore the activity of a pair of GH10 enzymes, BACOVA_04387 and BACOVA_04390, encoded by PUL-XylS of *B. ovatus*. Previous studies on *Bacteroides* spp. glycan utilisation systems have shown the importance of protein localisation in relation to the degradative apparatus (Cartmell *et al.*, 2011; Cuskin *et al.*, 2015). Data from Dr Rogowski confirmed the LipOP prediction of the cellular location of BACOVA_04390 and BACOVA_04387. Rogowski, using fluorescent microscopic analysis of cells labelled with specific antibodies and whole cell enzyme assays (Rogowski *et al.*, 2015), showed that BACOVA_04387 and BACOVA_04390 were located in the periplasm and cell surface, respectively.

The catalytic efficiency of BACOVA_04390 against BX was 30-fold lower than BACOVA_04387 (Figure 3.5b). Reduced activity of the surface BACOVA_04390 is most likely caused by the insertion of two CBM4 sequences into the GH10 catalytic module (Figure 3.4). The CBM4 insertion in the GH10 module appears to modify the activity of BACOVA_04390, which may explain not only its lower catalytic efficiency but its inability to degrade oligosaccharides with a low DP. Thus, the slow generation of large xylooligosaccharides by BACOVA_04390 reduces the possibility of saturation of the associated SusCD transport machinery.

3.3.1.1 Investigating number of subsites and preferred productive binding modes

HPAEC data (Figure 3.6) was used to measure substrate depletion and thus the activity of each enzyme on the xylooligosaccharides X2-6. k_{cat}/K_M of both BACOVA_04390 and BACOVA_04387 GH10

enzymes acting on each of the xylooligosaccharides was determined using Equation 3.1, with the assumption that substrate concentration in each reaction was below the K_M of the enzyme, and recorded in Table 3.2.

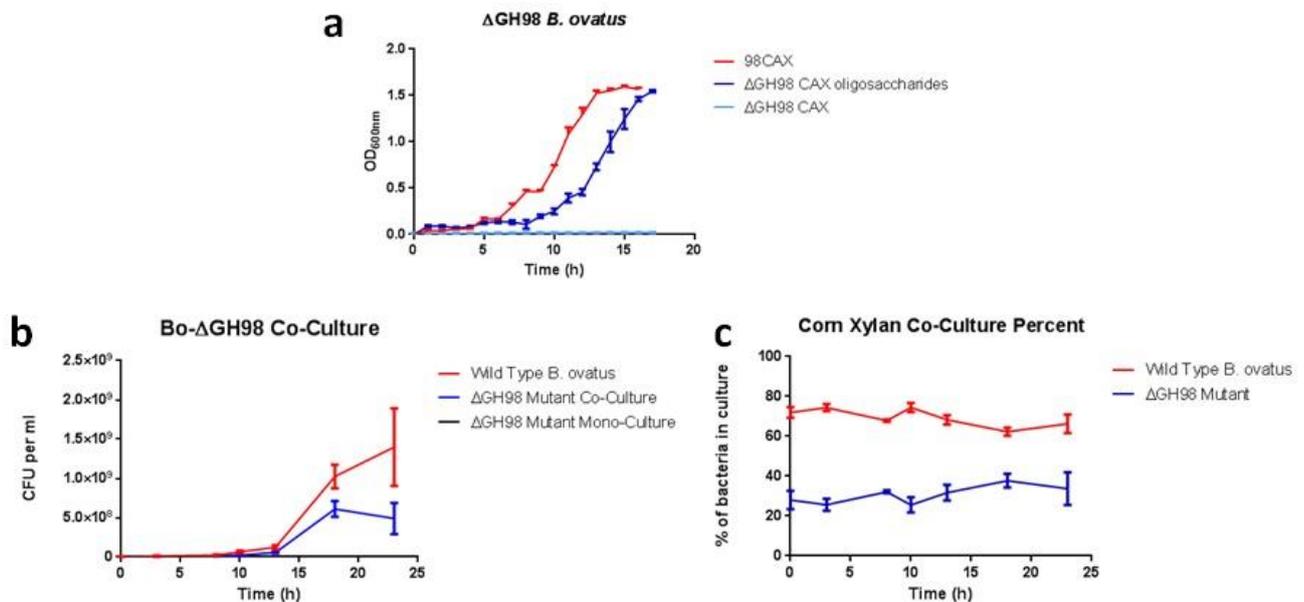


Figure 3.17 Growth of $\Delta GH98 B. ovatus$ on CX in mono-culture and in co-culture with wild type *B. ovatus*. Growth measured by optical density of the $\Delta GH98$ mutant *B. ovatus* in MM with 0.5 % CX and CX oligosaccharide generated by digestion with recombinant BACOVA_03433 (a). Co-culture of $\Delta GH98$ and *B. ovatus* in MM with 0.5 % CX measured by CFU/ml determined by ratio of wild type to $\Delta GH98$ and total CFU/ml (b). Ratio of wild type *B. ovatus* to $\Delta GH98$ in the co-culture. Ratio was deduced by purification of genomic DNA at time points in the co-culture and quantification of specific tags inserted into the genomic DNA of *B. ovatus* and $\Delta GH98$ by qPCR (c).

BACOVA_04390 activity on X4-6 was determined, as the length of substrate was increased the rate increased in 10-fold increments between X4, X5 and X6 (Table 3.2), indicating a preference for longer substrates. This is further supported by lack of quantifiable activity on xylooligosaccharides with a dp < 4 (Figure 3.6c). Unfortunately, it is not possible to deduce the precise number of subsites in the substrate binding cleft of BACOVA_04390 active site from these data. Insertion of CBM4 modules most likely extended the length of the binding site potentially compensating for weak binding at subsites proximal to the active site, as evidenced by the lack of activity against X3, while allowing for increased activity on long oligosaccharides and xylan polysaccharide. (Table 3.2). Thus,

by relying on subsites distal to the core of the substrate binding cleft (subsites -2 to +1), the xylanase specifically targets polysaccharides mediating infrequent cuts to the xylan backbone, reducing the amount of oligosaccharides generated at the cell surface during the degradative process.

BACOVA_04387 showed activity on xylooligosaccharides X2-6, although activity on X2 was not quantifiable. From X3 to X4 there was a 175-fold increase in activity (Table 3.2). X5 was hydrolysed 10-fold more efficiently than X4, while activity on X6 was similar to X5 (Figure 3.6, Table 3.2). These data demonstrate that BACOVA_04387 has evolved to digest mid to long oligosaccharides to generate products with lower DP. The high activity of BACOVA_04387 on BX is tolerated due to its location in the periplasm; here it generates short oligosaccharides which are sequestered in the periplasm without loss to the environment. However, as discussed later, this system does allow complex oligosaccharides to be released into the environment prior to sequestration into the periplasm where they may be used by competing microorganisms.

BACOVA_04387 xylooligosaccharide product profiles can be used to infer some information on the preferred binding mode of selected substrates (Figure 3.7). X3 was converted to X2 and X1, albeit very slowly, requiring overnight incubation at high enzyme concentrations (Figure 3.6a,b). X4 was converted exclusively to X2 suggesting the importance of the +2 subsite, and indicating the absence of a significant -3 subsite, which would have promoted the generation of X3 and X1. Indeed the difference in activity against X3 and X4 reflects substrate binding at +2, which was $3.19 \text{ kcal mol}^{-1}$. X5 was converted to X2 and X3. The limit products generated from X6 were X4, X3 and X2, implying two potential binding modes with hydrolysis resulting in two X3 or an X2 and X4. The ratio of products suggests both modes are equally likely (Figure 3.6a).

Both GH10 enzymes described here have acquired adaptations to the function required during xylan utilisation. The Surface GH10, BACOVA_04390 has undergone insertion events altering specificity of the enzyme to target longer substrates with a longer substrate binding site, generating long oligosaccharide products. While the periplasmic BACOVA_04387 remains closer to the more typical

GH10 xylanase by quickly degrading the oligosaccharide substrates in the periplasm releasing short oligosaccharides which are then degraded to monosaccharides by xylosidases encoded in the xylan PULs (Rogowski *et al.*, 2015).

3.3.2 Two surface endo-xylanases demonstrating different specificities from the large Xylan PUL

The PUL-XyLL was upregulated by arabinose containing xylans, WX and CX (Martens *et al.*, 2011). The locus encodes a larger repertoire of enzymes than the PUL-XyLS, which are required to fully digest and utilise the complex xylans (Rogowski *et al.*, 2015). BACOVA_03432 and BACOVA_03433 were both confirmed as cell surface enzymes by Dr Rogowski (Rogowski *et al.*, 2015). The data presented here demonstrate that both proteins are endo-acting xylanases (Figure 3.8a and 3.9a).

BACOVA_03432 is active on BX, a glucuronoxylan, indicating glucuronic acid substitutions are vital to substrate recognition and hydrolysis. As a typical GH30 glucuronoxylanase, activity is conferred by presence of a GlcA residue appended to the xylose occupying the -2 subsite of the enzyme prior to hydrolysis, hence the specificity for glucuronoxylans (St John *et al.*, 2010; Urbanikova *et al.*, 2011). Lack of activity on WX and CX, is consistent with the absence of GlcA in WX and the complex context of the uronic acid in CX. Assaying BACOVA_03432 activity on BX, within the constraints of substrate solubility and an observed high K_M revealed that initial rates were in the linear phase of the Michaelis-Menten kinetic curve. Thus, only k_{cat}/K_M could be deduced and not the individual kinetic parameters. The high observed K_M reflects the relatively low concentration of glucuronic acid substitutions to the xylan backbone.

BACOVA_03433 is active on CX but not BX or WX (Figure 3.9a), indicating a particular motif unique to CX is vital to activity. BACOVA_03433 is a GH98 enzyme, the first known member of this family to display xylanase activity. Interestingly, previously characterised GH98 enzymes are endo-galactosidases active on glycans presented on the surface of red blood cells, and play a role in bacterial pathogenesis (Anderson *et al.*, 2005; Shaikh *et al.*, 2009). Oligosaccharides generated by BACOVA_03433 activity on CX appear as an undefined smear continuous with the polysaccharide

spot at the origin on the TLC plate (Figure 3.9a). This indicates that the reaction products are high molecular weight oligosaccharides suggesting that the specificity determinant(s) for the xylanase occur infrequently in CX. The infrequent specificity determinants available to BACOVA_03433 and BACOVA_03432, ensures that these xylanases also only generate large products. This is consistent with the concept that infrequent cleavage of the xylan backbone ensures that the generation of a small number of large products is unlikely to saturate the import apparatus, as discussed above.

3.3.3 *Bi. adolescentis* growth on xylans and xylan derived oligosaccharides

Bi. adolescentis was unable to utilise xylose BX, WX or CX as growth substrates. However, when BX and WX were hydrolysed with BACOVA_04390, and CX with BACOVA_03433 (both enzymes are xylanases), the resultant xylooligosaccharides (XOS), arabinoxylooligosaccharides (AXOS) and glucuronoarabinooligosaccharides (GAXOS) were then used as growth substrates by *Bi. adolescentis* (Figure 3.11a & b). These data suggest *Bi. adolescentis* can only import or utilise xylose in the context of short xylooligosaccharides, explaining the relatively low growth on WX and BX oligosaccharides as only a fraction of these molecules can be used by *Bi. adolescentis*. This contradicts findings of Pastell *et al.*, (2009), which showed *Bi. adolescentis* grew on xylose. Both the work described here and that of Pastell *et al.*, (2009) used *Bi. adolescentis* ATCC 15703, however, the growth medium differed. Pastell included 1 g/L yeast extract, which was not used here as this would have encouraged inappropriate growth of *B. ovatus*, which is able to grow on fungal glycans (Martens *et al.* 2011). The difference observed in growth may stem from the longer growth period used in the Pastell study where absorbance was measured at 60 and 140 h (Pastell *et al.*, 2009), whereas the cultures here were only allowed to grow for 30 h (Figure 11a), perhaps growth on xylose requires longer incubation time. This does not affect the results of this study, in which xylose was used as a control substrate to ensure growth observed for *Bi. adolescentis* in co-culture is the result of XOS/AXOS utilization and not another *B. ovatus* metabolite. To ensure *Bi. adolescentis* were indeed utilising XOS and AXOS generated by *B. ovatus* and not any other *B. ovatus* metabolites which may be

released into the media, the two species were co-cultured in media containing xylose, a sugar which is inaccessible to *Bi. adolescentis* but not *B. ovatus*, as a sole energy source (Figure 3.12a). *B. ovatus* grew while *Bi. adolescentis* did not (Figure 3.12a), confirming the growth observed in the xylan co-cultures was a result of oligosaccharide cross-feeding and not a result of *Bi. adolescentis* utilising another *Bacteroides* metabolite.

Bi. adolescentis possesses glycoside hydrolases potentially active on XOS or AXOS, including β -xylosidases and α -arabinofuranosidases (van den Broek *et al.*, 2005; Lagaert *et al.*, 2010). The respective genes are in close proximity to genes encoding a putative extracellular surface binding protein (ESBP), ABC transporter and LacI-type regulator, indicating a locus upregulated in response to XOS or AXOS that depolymerize these oligosaccharides. The ESBP shows sequence identity to a characterised ESBP, BIXBP, expressed by *Bi. animalis*. BIXBP is a XOS binding protein able to tolerate some decorations to the xylose chain. ESBPs bind oligosaccharide with very high affinity outside the *Bifidobacterium* cell and brings the ligand to its corresponding ABC transporter embedded in the cell membrane for transport into the cytoplasm for subsequent digestion (Ejby *et al.*, 2013). This high affinity oligosaccharide binding protein could be the basis of the highly efficient *Bifidobacterium* oligosaccharide utilisation system, forging a specific niche within the highly competitive human gut microbiota.

3.3.4 *B. ovatus* generates oligosaccharides at the cell surface during growth on Xylans

The xylan degradation apparatus expressed by PUL-XyIS and PUL-XyIL are optimized to maximise the intracellular breakdown and uptake of hemicellulose, similar to previously characterised glycan utilisation systems of *Bacteroides* (Martens *et al.*, 2009). The *B. thetaiotamicron* yeast mannan utilisation system adopts a selfish mechanism, where all breakdown products generated by the surface mannanases are transported into the periplasm, resulting in an absence of breakdown products in the growth supernatant (Cuskin *et al.*, 2015). Here, when grown on BX, WX and CX, *B. ovatus* generates oligosaccharides in the growth supernatant during different phases of growth

(Figure 3.10). The TLC data show products that match the XOS standards and an additional oligosaccharide which does not migrate with xylooligosaccharides used here (Figure 3.10). The unknown oligosaccharide remains in the media at stationary phase indicating it is not used by *B. ovatus* for growth. In plant cell walls the xylan backbone in the hemicellulose fraction is built upon an oligosaccharide of 4-β-D-Xylp-(1→4)-β-D-Xylp-(1→3)-α-L-Rhap-(1→2)-α-D-GalpA-(1→4)-D-Xylp at the reducing end of the chain (Pena et al., 2007). During synthesis of xylan in plant cell walls this oligosaccharide acts as a primer from which the backbone is built (Pena et al., 2007). As the unknown oligosaccharide is present in the BX, WX and CX culture supernatants, it may be the primer oligosaccharide as this is a common feature of xylans. Utilisation of the xylan primer by *Bi. adolescentis* was not tested here it was found to be not possible to purify in suitable quantities for growths. It is unlikely that *Bi. adolescentis* would be capable of utilising a relatively complex oligosaccharide such as this with four unique linkages requiring specific enzymes for breakdown of a sparse substrate. *Bi. adolescentis* was able to use oligosaccharides generated from a digest of BX and WX with recombinant BACOVA_04390 (Figure 3.11c&d), confirming *Bi. adolescentis* is capable of utilising xylooligosaccharides derived from these polysaccharides (Figure 3.11a). *Bi. adolescentis*, however, was unable to use CX digested with the GH98 xylanase BACOVA_03433 as a growth substrate, suggesting that the resultant oligosaccharides were too complex to be utilised by the *Bifidobacterium* (Figure 3.11a). BACOVA_03433 generates GAXOSs that are extensively decorated, and these side chains may block binding by the ESBP, or perhaps the oligosaccharides are simply too large to fit into the ABC transporter channel used for oligosaccharide import. Bioinformatic and biochemical studies of *B. adolescentis* xylanases show activity on only a relatively narrow range of xylans and/or xylooligosaccharides (van den Broek et al., 2005). *Bi. adolescentis* encodes two GH43, a GH120 β-xylosidase and a GH8 xylanase, which show activity on xylooligosaccharides and are vital to utilisation (Lagaert et al., 2007). The presence of GH43 α-L-arabinofuranosidases suggests *Bi. adolescentis* is capable of utilising AXOS in addition to linear XOS (Lagaert et al., 2010).

In contrast with the selfish hypothesis proposed for mannan utilisation (Cuskin *et al.*, 2015), *B. ovatus* is able to support the growth of *Bi. adolescentis* on BX and WX (Figure 3.13a,b). Despite differences in WX and BX complexity and the presence of arabinose substitutions in WX, similar levels of *Bi. adolescentis* were observed in the co-cultures on the two xylans. During growth *Bacteroides spp* generate metabolites which include short chain fatty acids (SCFAs) as waste products of anaerobic ATP generation (Salt *et al.*, 1985). SCFAs are energy rich molecules that are used as an energy source by colonocytes in the gut (Donohoe *et al.*, 2011).

Interestingly, *B. ovatus* viable counts are much higher in the CX co-culture than either of the BX or WX co-cultures (Figure 3.14d). As *B. ovatus* generates much larger oligosaccharides at the cell surface when cultured on CX, than those produced from WX and BX, each CX oligosaccharide transported contains higher potential energy yield than BX or WX oligosaccharides. Due to the nature of the SusCD transport system each oligosaccharide, regardless of size is imported at the cost of an ATP molecule, giving higher energy to cost ratio for the longer oligosaccharides, which in turn allows greater growth, hence the greater cell numbers observed in the CX co-culture. In this regard it may also be relevant that all the surface xylanases have evolved to cleave the xylan backbone infrequently generating large oligosaccharides.

CX, as a highly complex substrate, requires a suite of enzymes for full breakdown and utilisation by *B. ovatus*, which are not associated with typical xylan degradation systems (Rogowski *et al.*, 2015). The inability of *Bi. adolescentis* to grow on CX oligosaccharides, and in co-culture with *B. ovatus* on CX, may simply reflect the lack of the enzymes or transport systems required to degrade the complex xylan. This view is supported by the observation that a mutant of *B. ovatus*, lacking the surface GH98 xylanase was unable to grow on CX in monoculture, but could utilise the polysaccharide in co-culture with the wild type bacterium (Figure 3.17a). These data demonstrate oligosaccharides are released during *B. ovatus* growth on CX, but these molecules are not utilized by *Bi. adolescentis*. The inability to observe the CX oligosaccharides may be because they are too large

to migrate from the origin on TLC (Figure 10c). It is possible that CX and other highly complex substrates are only used as a last resort during times of nutritional crisis, requiring a high transcriptional/translational investment from the cell to produce the required apparatus for utilisation. For *Bifidobacteria* the initial investment to use such substrates may be too high when host glycans, such as mucins, are readily available (Egan *et al.*, 2014). *B. thetaiotaomicron* glycan utilisation preferences can be organised into a hierarchy showing PULs directed against host glycans and simpler glycans are preferentially upregulated over more complex glycans (Rogers *et al.*, 2013), which may hold true for *B. ovatus* as a relative of *B. thetaiotaomicron*.

3.4 Conclusion

B. ovatus expresses the full repertoire of enzymes required to fully degrade simple (BX, WX) or complex (CX) xylans in the human gut lumen. PUL-XylS orchestrates the degradation of simpler xylans expressing two GH10 xylanases displaying different levels of activity reflecting their localisation to the cell surface, BACOVA_04390, or periplasm, BACOVA_04387, to minimise loss of breakdown products to the competitors in the gut. Despite this adaptation oligosaccharides are released that are used by other members of the gut microbiota. In this chapter, *Bifidobacteria* were shown to use the BX- and WX-derived oligosaccharides generated by *B. ovatus*. In turn the *Bifidobacteria* promote gut health by helping to alleviate symptoms of inflammatory bowel disease (IBD) and modulating inflammation of the gut during immune responses (Kajander *et al.*, 2008; Ishikawa *et al.*, 2011). Although variable in size, it would appear that *B. ovatus* releases xylooligosaccharides into the culture supernatant irrespective of the complexity of the xylan. *B. ovatus* appears to aid *Bi. adolescentis* when co-cultured on xylans, generating accessible substrate from inaccessible xylan polysaccharide, despite lack of an apparent benefit of promoting *Bi. adolescentis* growth to the bacterium itself. This apparent altruism may be rewarded in the gut from currently unknown pathways due to the complexity of the interactions within the HGM. Data presented here demonstrates potential for enrichment of different members of the HGM by

including different fractions of glycan in the human diet. Complex GAX, like CX, promote *B. ovatus* growth in the gut, which in turn produce propionate, a molecule implicated in lipogenesis reduction in the host (Hosseini *et al.*, 2011). Utilisation of simple xylans by *B. ovatus* will promote growth of *Bifidobacteria* to increase butyrate, a molecule known to maintain the health of intestinal epithelium (Wachtershauser and Stein, 2000). Indeed such oligosaccharides have been shown to selectively promote *Bifidobacteria*, suggesting simple xylans and xylooligosaccharides can be used as bifidogenic prebiotics, due to a relatively low dose required to have a positive effect (Finegold *et al.*, 2014). Products of complex xylans are, however only used by a small subset of gut microbes due to the rarity of enzymes displaying relevant activities. Thus, the xylan-*B. ovatus* axis could be manipulated to encourage the growth of selective organisms in the human distal gut microbiota.

Chapter 4: Pectin Utilisation by *Bacteroides thetaiotaomicron*

4.1 Introduction

4.1.1 Background

Pectic polysaccharides are some of the most complex glycans in the human diet. The backbone of pectins, exemplified by homogalacturonan, rhamnogalacturonan I (RGI) and rhamnogalacturonan II (RGII) are rich in galacturonic acid (GalA). However, the extensive side chains that decorate some pectins can be charged, as occurs in RGII, or contain neutral sugars typified by the galactan and arabinan chains that decorate RGI (Figure 4.1). The RGI backbone consists of alternating α 1,2-linked rhamnose (Rha) and α 1,4-linked GalA with occasional acetyl substitutions of the GalA residues (Mohnen, 2008). The galactan and arabinan side chains, which extend from rhamnose in the backbone. Galactan consists of β 1,4-linked galactose units while arabinan consists of α 1,5-linked arabinose with O2 and/or O3 arabinose side chains (Silva *et al.*, 2016). Although lacking GalA, arabinans and galactans are viewed as pectins.

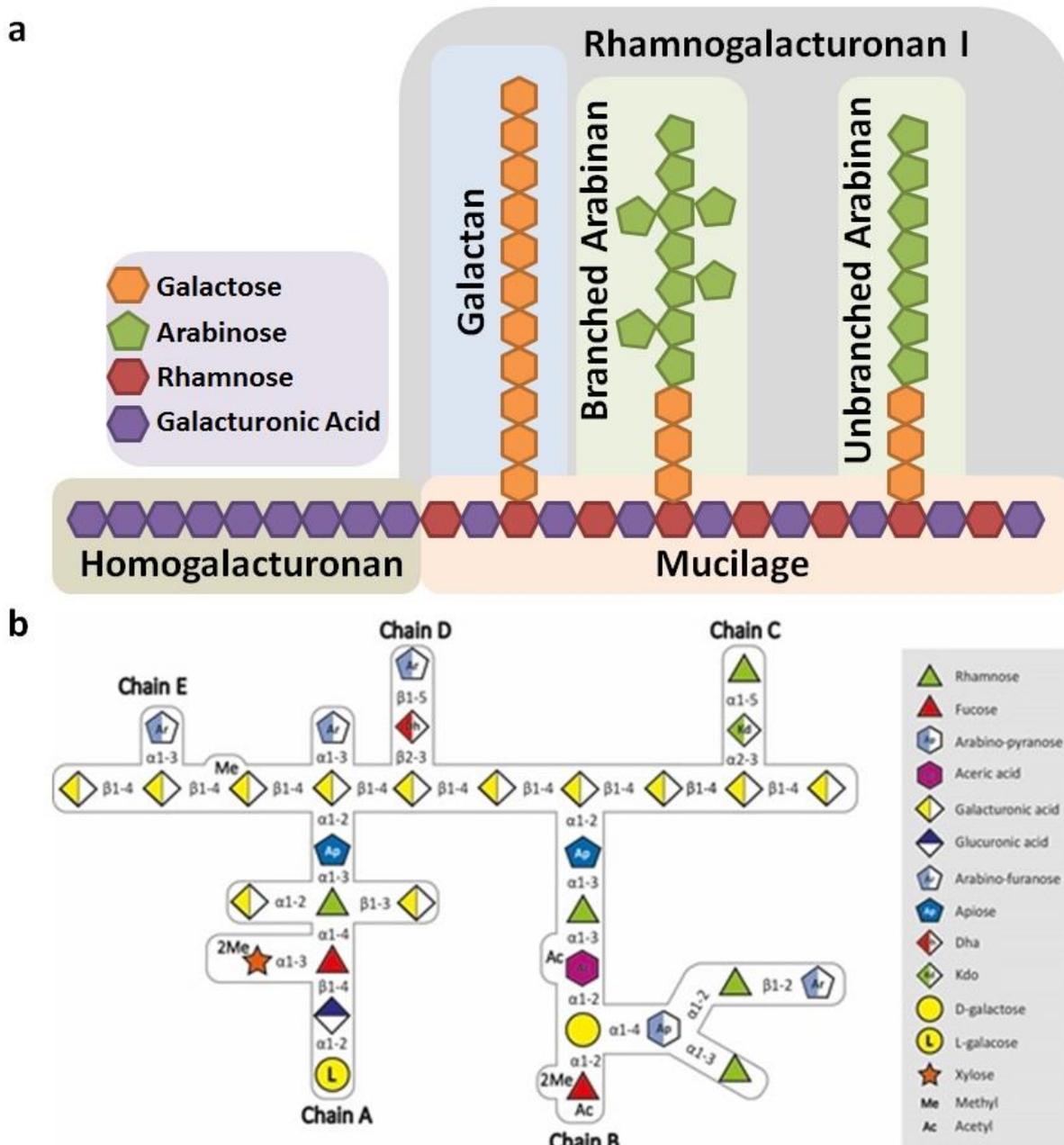


Figure 4.1 Pectin structure. A schematic representation of the Pectic polysaccharides. Schematic representation of RGI along with, galactan, arabinan and homogalacturonan (a), and RGI (b).

Regions of RGI give way to the non-decorated Homogalacturonan (Figure 4.1a) which, as the name suggests consists of GalA residues. The most complex pectin, probably the most complex glycan in the human diet, is RGI. The simpler galactan and arabinan side chains of RGI (Figure 4.1a) form helical structures inhibiting tight packing of pectic polysaccharides. Pectin performs a space filling role providing elasticity to the plant cell wall. During times of growth the amount of pectin increases

to give the cell wall flexibility (Mohnen, 2008), much like cholesterol in mammalian cell membranes. Along with the more irregular structure of pectin side chains forming interactions between strands, borate bridges exist between RGII side chains, giving a degree of strength to the polysaccharide (Pabst *et al.*, 2013).

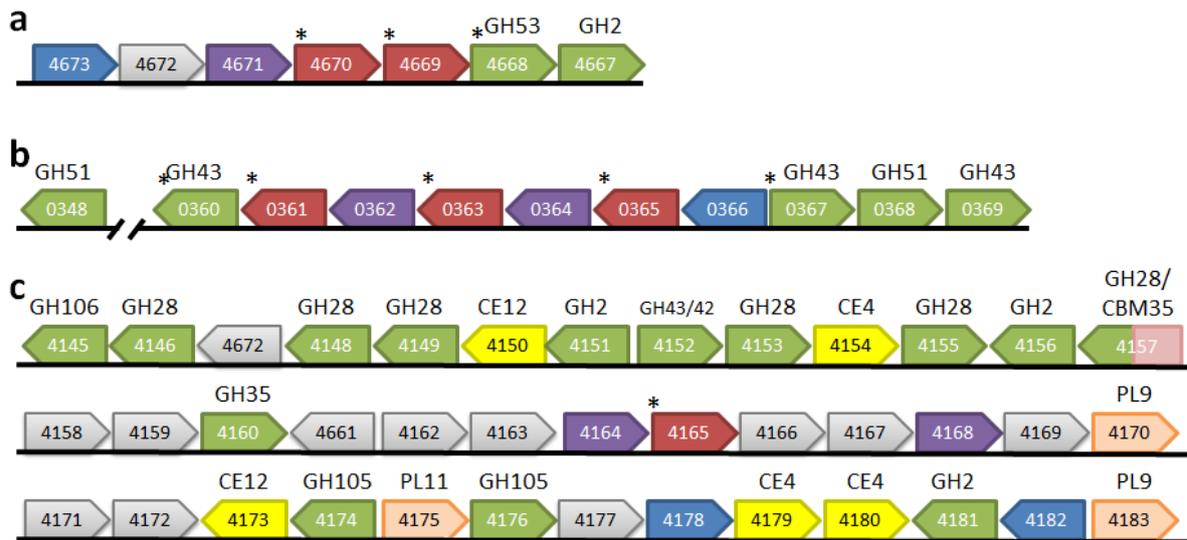


Figure 4.2 Pectin PUL organisation. A schematic representation of the polysaccharide utilisation loci upregulated in response to galactan (a), arabinan (b), rhamnogalacturonan I (c). Gene locus tags are prefixed with *bt_XXXX*, enzyme family is annotated above the gene as GHXX, glycoside hydrolase; PLXX, polysaccharide lyase; CEXX, carbohydrate esterase; CBMXX, carbohydrate binding module. Genes are coded by colour, blue, hybrid two component system (HTCS), grey, gene of unknown function, purple, SusC-homologue, red, surface glycan binding protein or SusD-homologue, green, glycoside hydrolase, yellow, carbohydrate esterase, pink, carbohydrate binding module, orange, polysaccharide lyase. Proteins predicted to be lipoproteins attached to the cell surface are indicated by an asterisk. Abridged version of figure and data from Martens *et al.* (2011).

Due to the heterogeneity of glycosidic bonds found in pectins the human gut bacterium *Bacteroides thetaiotaomicron* express a wide array of carbohydrate active enzymes (CAZymes) to degrade and subsequently utilize these polysaccharides (Martens *et al.*, 2009). These include glycoside hydrolases (GH), pectic lyases (PL) and carbohydrate esterases (CE) (Lombard *et al.*, 2014), and reflect the relative complexity of the target substrates. The majority of human gut microbiota (HGM) members only utilise a subset of pectic polysaccharides. *Bifidobacterium breve* (*Bi. breve*) and *Bi. longum*, for

example, express glycan utilisation systems that target galactan while being unable to grow on RGI or RGI. *B. thetaiotamicron*, however, is able to utilise all known pectins (Martens *et al.*, 2011). As such, *B. thetaiotaomicron* is considered a generalist glycan user whose genome encodes many complex carbohydrate utilisation systems. The *Bacteroides* polysaccharide utilisation systems mediates glycan degradation at the cell surface and in the periplasm (Martens *et al.* 2009). The polysaccharide utilisation loci (PULs) are identified by the presence of *susCD* homologues, which mediated oligosaccharide import across the outer membrane, along with CAZymes and a regulator. *B. thetaiotaomicron* encodes 208 *susC* and *susD* homologues and is predicted to encode 88 PULs, devoting 18% of its genome to glycan utilisation (Martens *et al.*, 2009).

A number of these PULs have been shown experimentally to be upregulated in response to detection of the target glycan (Martens *et al.*, 2011). Based on this criterion several *B. thetaiotaomicron* PULs are dedicated to the breakdown and utilisation of pectic polysaccharides (Figure 4.2). There are discrete PULs dedicated to each pectic polysaccharide, although some co-regulation/crosstalk has been observed between these related PULs (Martens *et al.*, 2011), although this may reflect the purity of these polysaccharides.

4.1.2 Aims

The overarching objective of this chapter is to characterise the *B. thetaiotaomicron* systems that degrade selected pectins. The specific aims are as follows:

- To characterise the galactan and arabinan utilisation systems.
- To characterise the enzymes involved in the release of the galactooligosaccharides that remain appended to the RGI backbone after galactan degradation.
- To investigate the degree and extent to which cross-feeding occurs during growth of *B. thetaiotaomicron* on pectic polysaccharides.

4.2 Results

4.2.1 Cloning, Expression and Purification of Pectin utilisation proteins

This chapter describes the characterisation of *B. thetaiotaomicron* and a *B. ovatus* proteins expressed in response to galactan, arabinan and RGI, leading to the complete degradation and utilisation of each polysaccharide (Table 4.1). The genes in the PULs activated by these pectins (Figure 4.2) were cloned either 'in-house' or externally by NZYTech (Lisbon, Portugal). The regions of the genes were cloned without the associated N-terminal signal peptide coding sequence. This avoids membrane localisation when expressed in *E. coli*, increasing the chance of producing soluble recombinant protein. The genes were cloned into the expression vectors pET28b or pET21a, which results in the fusion of a C-terminal His-tag to the encoded recombinant protein to enable purification using immobilized metal ion affinity chromatography. While most of the proteins studied retained their His-tag, *bacova_05493* (*B. ovatus* gene) and DNA encoding the ligand binding domain (from amino acid 27 to 798 of the HTCS sequence) of the predicted hybrid two component system (HTCS) regulator, BT_4673, were cloned into pET28a to take advantage of thrombin cleavable of the His-tag from the C-termini of the recombinant proteins. The majority of the GHs, PLs and CEs described in this chapter are members of existing CAZy sequence-based families, and are defined as GHXX, PLXX and CEXX, respectively.

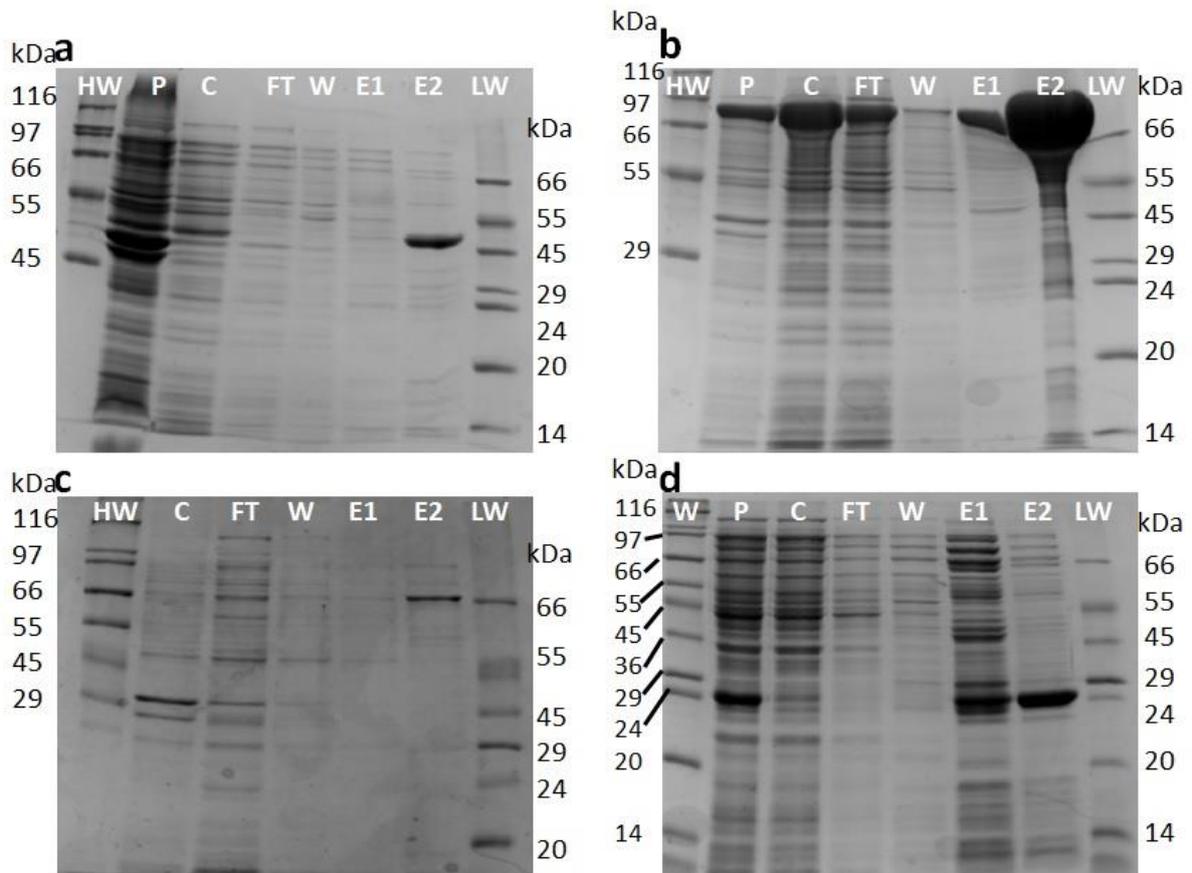


Figure 4.3 Examples of and purification of recombinant proteins involved in pectin degradation. The proteins were purified by IMAC and the fractions were subjected to SDS-PAGE using a 12.5% polyacrylamide gel. BT_4668 (a, medium expression), BT_4667 (b, high expression), BT_0360 (c, low expression), BT_4158 (d, medium expression). HW, high molecular weight marker, P, bacterial cell pellet, C, cell lysate, FT, flow through, W, wash fraction with TALON buffer, E1, elution fraction 1 with 5 mM imidazole in TALON buffer, E2, elution fraction 2 with 100 mM imidazole in TALON buffer and LW, low molecular weight marker. The expression and purification of the proteins are described in Chapter 2.6.2. Full list of protein expression gels can be found in Appendix A.4.

Protein	Predicted Function/ GH Family	Recombinant protein Length (aa)	Recombinant Protein Mass (kDa)	Extinction Coefficient (M ⁻¹)	Dialysis Buffer
RGI PUL					
BT_4151*	GH2	949	108	198590	A
BT_4156*	GH2	954	107	197850	A
BT_4160*	GH35	690	78	129650	A
BT_4181*	GH2	980	111	201570	A
BT_4158*	Carbohydrate Esterase	217	25	35870	A
Galactan PUL					
BT_4667*	GH2	816	89	157525	A
BT_4668*	GH53	345	37	73380	A
BT_4669*	Surface Glycan Binding Protein (SGBP)	566	60	140385	B
BT_4670*	SusD-like protein	536	58	101690	B
BT_4673	Hybrid Two Component System (HTCS)	747	88	176480	B
Bacova_05493	GH	869	99	154940	D
Arabinan PUL					
BT_0348*	GH51	504	57	110240	A
BT_0360	GH43	628	70	185530	C
BT_0361*	SusD-like protein	579	66	120100	B
BT_0363*	SusD-like protein	568	63	125140	B
BT_0365*	Surface Glycan Binding Protein (SGBP)	746	81	162610	B
BT_0367	GH43	486	56	137740	C
BT_0368*	GH51	653	74	121700	A

Table 4.1 Proteins expressed and use in this chapter. * = Cloned by NZYTech, A=20 mM Tris-HCl buffer pH 7.5 containing 150mM NaCl pH 7.5, B=50 mM Na-HEPES buffer pH 7.5 containing 150 mM NaCl, C=20 mM sodium phosphate 150 mM NaCl pH 7.5, D=20 mM sodium phosphate pH 7.5 300 mM NaCl.

Vectors containing the target genes were used to transform *E. coli* strains BL21(DE3) or TUNER (DE3) when the target protein was predicted to be a β -galactosidase (GH2 or GH35) as this *E. coli* strain lacks a functional *lacZ* gene and thus does not express an endogenous β -galactosidase. The expression and purification of these proteins are described in Chapter 2.6. Examples of SDS-PAGE analysis of the expression and purification of the recombinant proteins is shown in Figure 4.3. In

general the proteins were purified to a high level of purity using IMAC, providing initial expression was reasonable.

4.2.2 Galactan utilisation by *B. thetaiotaomicron*

4.2.2.1 Identification of the Galactan PUL

A genetic locus (*bt_4667-73*) was identified in the genome of *B. thetaiotaomicron* that was highly upregulated when the bacterium was grown on galactan (Martens *et al.*, 2011). Further bioinformatic analysis of the genes in this locus predicted the presence of GH2 and GH53 enzymes, *bt_4667* and *bt_4668*, respectively. The locus was identified as a PUL by the presence of the characteristic *susCD*-homologues, *bt_4671* and *bt_4670* (Figure 4.2a) (Martens *et al.*, 2011).

4.2.2.2 Purification of Galactooligosaccharides

Specific galactooligosaccharides were required to characterise the proteins encoded by the PUL-Gal of *B. thetaiotaomicron*. These oligosaccharides were generated by acid hydrolysis of galactan with 75 mM HCl at 100 °C for 4 h. The resulting mixture was neutralised with NaOH and concentrated before being loaded onto a P2-Biogel (BioRad) size exclusion column. Elution fractions were collected and evaluated by TLC (Figure 4.4a). Different oligosaccharide fractions were identified by relative size and comparison to known galactose and galactobiose standards. Each fraction containing pure oligosaccharide was pooled and concentrated to a working concentration for use in subsequent assays. Oligosaccharide purity was confirmed by high-performance anion-exchange chromatography (HPAEC) (Figure 4.4b). Hereafter galactooligosaccharides are referred to as GalX, where X is the degree of polymerisation (DP) of the oligosaccharide. DP was confirmed by comparison to known standards, galactose and Gal2, and extrapolating for the longer oligosaccharides assuming each band corresponds to 1 galactosyl-residue longer than the band which preceded, or peak in the case of HPAEC analysis (Figure 4.4b). The galactobiose standard used was supplied by Megazyme Ltd. was not pure giving two peaks (Figure 4.4b).

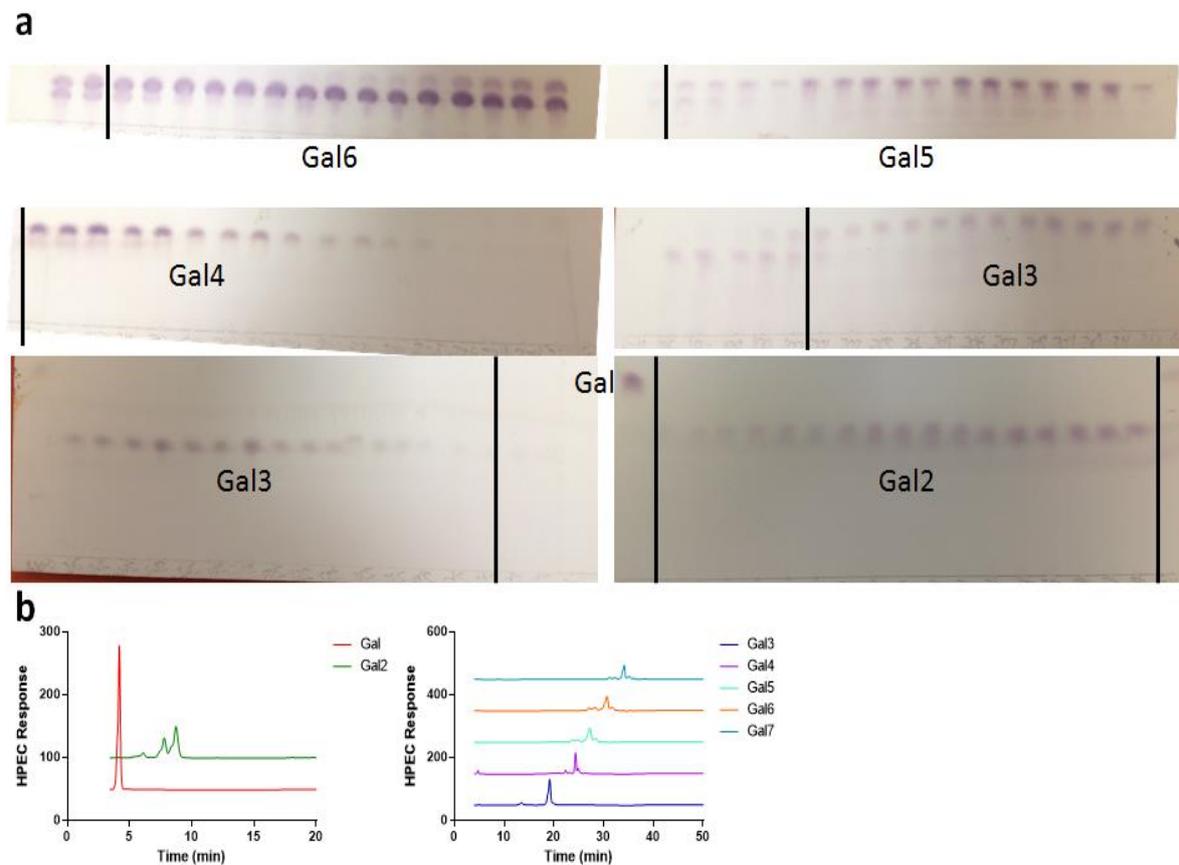


Figure 4.4 Galactooligosaccharide purification. Acid hydrolysed galactan was run over two P2 Bio-gel size exclusion columns. Collected fractions were run on TLC to assess oligosaccharide content (a). HPAEC chromatogram of purified galactooligosaccharide fractions to assess purity (b).

4.2.2.3 Activity of BT_4668 on galactan and galactooligosaccharides

Bioinformatic analysis of BT_4668 indicates the presence of a GH53 catalytic module, a family that to date comprises exclusively endo- β 1,4-galactanases that display a “retaining” mechanism (GHs that display an acid- base double displacement mechanism leading to retention of anomeric configuration) (Braithwaite *et al.* 1997). Sequence alignment of BT_4668 with well characterised GH53 enzymes; GalA from *Cellvibrio japonicus* (Braithwaite *et al.*, 1997), GanA from *Geobacillus stearotherophilus* (Tabachnikov and Shoham, 2013), GalA from *Bi longum* (Hinz *et al.*, 2005) and GalA from *Bi breve* (O'Connell Motherway *et al.*, 2011) identified conserved motifs (Figure 4.5a).

Motifs of note include the catalytic glutamate residues in the characterised GH53 enzymes, which in BT_4668 were Glu-180, the general acid/base, and Glu-292 which aligns with the known nucleophile in the other characterised GH53 enzymes (Figure 4.5a). PCR-based Site-directed Mutagenesis (described in Chapter 2.5.6) was used to introduce E180A and E292A substitutions into recombinant BT_4668 (Figure 4.6a). The resulting mutant was unable to hydrolyse galactan at 10 μ M enzyme incubated with 1 mg/ml galactan overnight at 37 °C (Figure 4.6b). These data confirm that Glu180 and Glu292 are the catalytic residues of BT_4668.

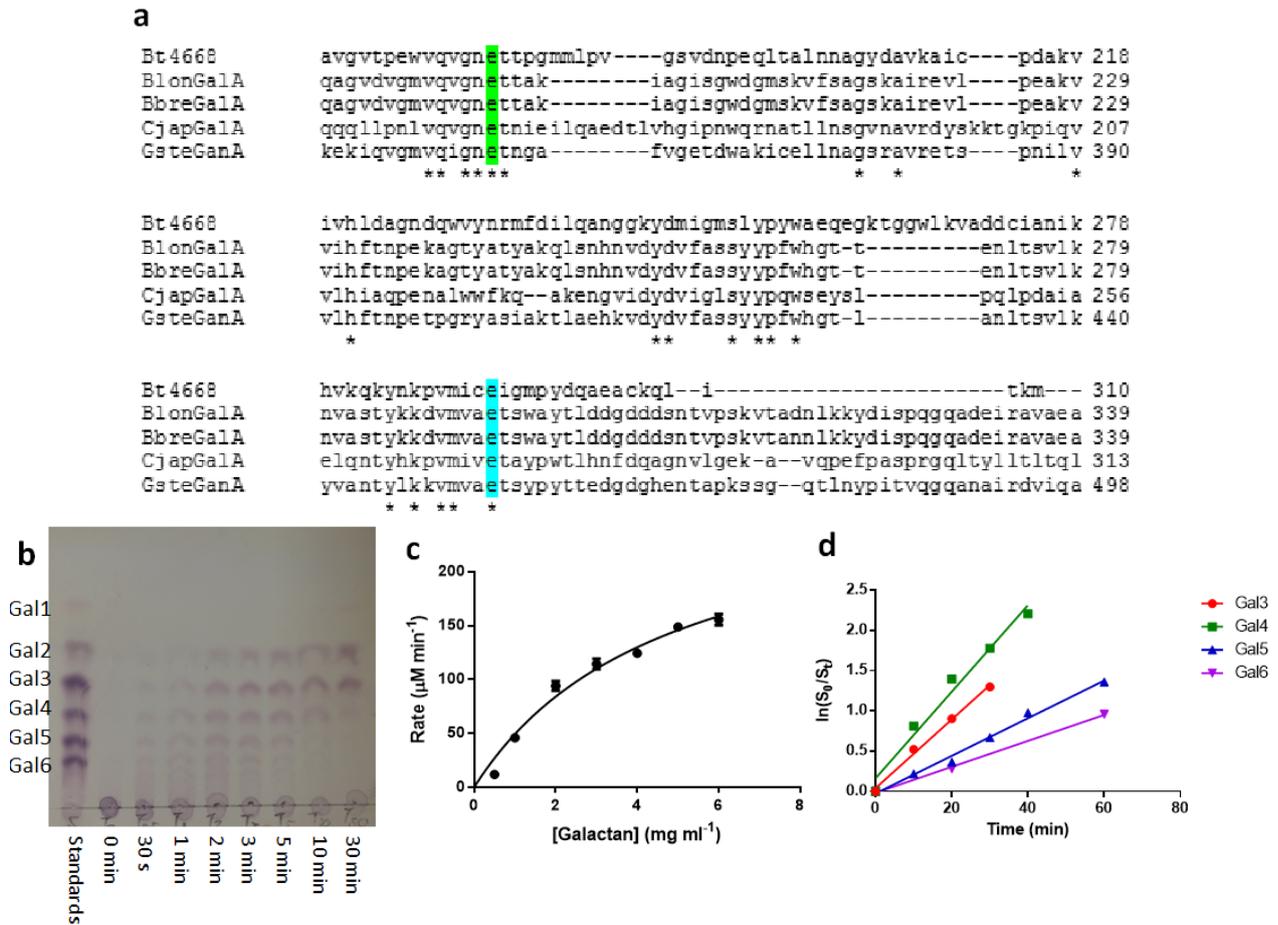


Figure 4.5 Characterisation of BT_4668. Amino acid sequence alignment (using clustal omega server) of BT_4668 with previously characterised GH53 enzymes, GalA from *C. japonicus*, GanA from *G. stearothermophilus*, GalA from *Bi longum* and GalA from *Bi breve* (a). The catalytic nucleophile is predicted to be Glu180 (green) and the general acid/base Glu292 (blue). TLC product profile of 1 mg/ml galactan digested with 1 μ M BT_4668 in 20 mM sodium phosphate 150 mM NaCl pH 7.5 over 30 min. Samples were run alongside known galactooligosaccharide standards Gal1-Gal6 (b). Michaelis-Menten curve of BT_4668 activity on galactan in 20 mM sodium phosphate 150 mM NaCl pH 7.5 at 37 $^{\circ}$ C (c). The inverse rate of substrate depletion of defined galactooligosaccharides when incubated with BT_4668 in 20 mM sodium phosphate 150 mM NaCl pH 7.5 at 37 $^{\circ}$ C to give catalytic efficiency (k_{cat}/K_M , d).

As BT_4668 belongs to GH53, it most likely displays endo-galactanase activity. To test this hypothesis 500 nM BT_4668 was incubated with 1 mg/ml galactan with samples taken and boiled at intervals to stop the reaction. The samples were then subjected to TLC alongside galactan oligosaccharide standards with known DPs as described in Chapter 2.8.4. Long oligosaccharides were present in the earlier time points, 30 s to 1 min. After 1 min a wide range of oligosaccharides were present and the

origin spot began to grow faint, indicating a reduction in long chain galactan in the reaction (Figure 4.5b). In the last sample there was an accumulation of Gal2 and Gal3 (Figure 4.5b). This product profile is typical of endo-activity, confirming BT_4668 as an endo- β -1,4-galactanase enzyme.

The DNSA reducing sugar assay (described in Chapter 2.8.1.1) was used to measure kinetics of hydrolysis of galactan by BT_4668. The resultant Michaelis-Menten curve (Figure 4.5c, Table 4.2) showed that the activity of BT_4668 on galactan was similar to that of GalA, a β 1,4-galactanase expressed by *C. japonicus*, while being four-fold more active than the newly described GH53 from *Bacillus licheniformis* CBMAI 1609 (de Lima *et al.*, 2016). BT_4668 activity on galactan appears to be similar to other characterised GH53 enzymes, confirming that BT_4668 is a typical GH53 galactanase.

Substrate	V _{max} ($\mu\text{M min}^{-1}$)	K _M (mg ml ⁻¹)	k _{cat} (min ⁻¹)	k _{cat} /K _M (ml min ⁻¹ mg ⁻¹)	Enzyme concentration (μM)
Galactan	285.3 ± 29.6	4.78 ± 0.92	285300 ± 29609	59623 ± 5816	0.001
Galactooligosaccharides					
Substrate				k _{cat} /K _M (min ⁻¹ mM ⁻¹)	
Galactobiose				ND	1
Galactotriose				0.28 ± 0.02	0.1
Galactotetraose				4.2 ± 0.5	0.01
Galactopentaose				23.4 ± 1.1	0.001
Galactohexaose				18.5 ± 0.7	0.001

Table 4.2 BT_4668 activity on galactan and galactooligosaccharides

Known concentrations of galactooligosaccharides with a DP of 3-6, were incubated at 37 °C with an appropriate concentration of BT_4668 and the rate of substrate depletion was determined by HPAEC. Table 4.2 shows catalytic efficiency ($k_{\text{cat}}/K_{\text{M}}$) of oligosaccharide degradation by Bt_4668, derived from the exponential loss of substrate over time. Substrate concentration was likely below the K_M of the enzyme. This was tested by halving the substrate concentration to ensure the rate remains constant (Appendix A.5) indicating the reaction is within the linear phase of the reaction (Figure 4.5d), allowing catalytic efficiency to be determined. Unfortunately, full kinetic analysis of

BT_4668 enzyme activity on the available galactooligosaccharides was not possible due to limited supply of these glycans as the nature of stopped-assay requires significant quantities of each oligosaccharide. The galactanase displays activity on all of the oligosaccharides tested, although it shows preference for oligosaccharides with a DP of ≥ 5 (Table 4.2). Maximum activity was observed using galactopentaose and galactohexaose as substrate. This indicates the enzyme possesses five subsites in the substrate binding cleft. Every subsite is occupied during catalysis of galactopentaose or galactohexaose (Table 4.2). Due to relative difficulty of generating galactooligosaccharides in quantities required for kinetic analysis there are currently no other reports of GH53 activity on galactooligosaccharides. As such the relatively low activity of BT_4668 on the galactooligosaccharides may be a characteristic of this specific enzyme or a shared property of GH53 galactanases, which, without further investigation this remains unclear.

a

```
MutantBT4668 AAATCAGTTCTTACCGCGTTAAAAGCGGTTGGGGTTACACCGGAATGGGTACAGGTAGGT 345
BT4668 AAATCAGTTCTTACCGCGTTAAAAGCGGTTGGGGTTACACCGGAATGGGTACAGGTAGGT 480

MutantBT4668 AATGCAACTCCCAGCATGATGCTTCTGTGGGGTCTGTTGATAATCCCGAACAACTG 405
BT4668 AATGAAACAACCTCCCAGCATGATGCTTCTGTGGGGTCTGTTGATAATCCCGAACAACTG 540

MutantBT4668 ACTGCACTCAACAATGCGGGATATGATGCCGTGAAAGCTATATGTCCCGACGCCAAAGTC 465
BT4668 ACTGCACTCAACAATGCGGGATATGATGCCGTGAAAGCTATATGTCCCGACGCCAAAGTC 600

MutantBT4668 ATTGTCCACCTGGACCGGGAAACGACCAATGGGTATAACAACCGCATGTTTCGACATTCTT 525
BT4668 ATTGTCCACCTGGACCGGGAAACGACCAATGGGTATAACAACCGCATGTTTCGACATTCTT 660

MutantBT4668 CAAGCCAATGGAGGTAATAACGACATGATAGGCATGTCGCTCTATCCATATTGGGCAGAG 585
BT4668 CAAGCCAATGGAGGTAATAACGACATGATAGGCATGTCGCTCTATCCATATTGGGCAGAG 720

MutantBT4668 CAGGAAGGCAAAACCGGTGGTTGGCTGAAAGTAGCCGATGACTGTATTGCCAATATCAAG 645
BT4668 CAGGAAGGCAAAACCGGTGGTTGGCTGAAAGTAGCCGATGACTGTATTGCCAATATCAAG 780

MutantBT4668 CATGTCAAGCAGAAATACAACAAACCTGTCAATGATTTGCGCATAGGTATGCCTTATGAT 705
BT4668 CATGTCAAGCAGAAATACAACAAACCTGTCAATGATTTGCGCATAGGTATGCCTTATGAT 840
```

b



Figure 4.6 Generation of an inactive catalytic residue mutant of BT_4668 by site directed mutagenesis. Nucleotide sequence alignment of wild type and mutated BT_4668. Plasmids containing potential mutations in *bt_4668* were harvested and sent to MWG for sequencing (a). TLC plate showing activity of the wild type BT_4668 and inactive E180A/E292A mutant of BT_4668 alongside known galactooligosaccharide standards, S, Gal1-6, E180A/E292A mutant, M, BT_4668 enzyme, E (b).

During catalysis of the defined oligosaccharides product profile data were also collected over the course of the reaction, and used to determine preferential binding modes for each substrate (Figure 4.7). For accurate assessment of product profiles of BT_4668 acting on each defined oligosaccharide only the initial reaction products were analysed, as later in the reaction the enzyme is likely to

degrade the larger oligosaccharide products generated, making data interpretation difficult. Understanding quantities of the oligosaccharides produced by the enzyme during a reaction can, along with end point assays, provide insights into possible productive binding modes and enable an estimation of binding energy at some of the subsites. Product peaks on the HPAEC chromatogram were compared to known concentrations of each oligosaccharide along with an internal standard to normalise for variations between HPAEC runs to give product concentration. BT_4668 was unable to hydrolyse galactobiose, which is left as an end product along with some Gal3 (Figure 4.5b). Digestion of Gal3 yields equal concentration of galactose and Gal2, although activity is very low (Figure 4.7a). It is well established that the -2 subsite plays a critical role in the activity of endo-acting enzymes employing the retaining mechanism (Davies 1995). For activity on Gal3 there is thus only one possible productive binding mode from -2 to +1 subsites (Figure 4.7d). The observed low activity of the enzyme on Gal3 could be, in part, due to the substrate occupying the enzyme active site in non-productive binding modes where the substrate does not occupy the -1 and +1 subsites simultaneously. However, the most likely explanation for low activity on Gal3 is that the substrate occupies only three of the possible five subsites resulting in reduced binding energy (Figure 4.7d). Possible products of Gal4 hydrolysis are two Gal2 or galactose and Gal3. The ratio of (2 x Gal2; binding -2 to +2):(Gal3 + Gal1; binding -3 to +1) was 5:3 ratio (Figure 4.7b), indicating a slightly higher binding energy at the +2 over the -3 subsite. As galactose, Gal2 and Gal3 are produced from Gal4 a -3 subsite must exist rather than a +3 subsite as, again, occupation at -2 is critical preventing a -1 to +3 binding mode. The product profile of Gal5 digestion is less clear, it shows galactose, Gal2, Gal3 and Gal4 are produced at similar concentrations (Figure 4.7c). These data indicate three possible productive binding modes; +1 to -4, -3 to +2 and -2 to +3 (Figure 4.7g). As the enzyme displays comparable activity against Gal6 and Gal5, the galactanase should contain only five subsites. The presence of Gal4 in the products generated from Gal5, however, points to a functional -4 subsite, while the generation of significant quantities of Gal2 from Gal4 indicates a productive +2 subsite. Thus, the product profiles suggest that the enzyme contains six subsites, while the similar activity of

the galactanase against Gal5 and Gal6 points to only five subsites. The products observed during digestion of Gal6 show a preference for the occupation of all six subsites generating Gal2 and Gal4 in higher quantities than any other oligosaccharide in the mixture of products (Figure 4.7d and 4.7h).

It is difficult to reconcile the binding mode of Gal6 and the biochemical data pointing to only five galactose binding subsites. Galactan displays a helical structure with a periodicity of six sugars (Cid *et al.* 2010), and thus likely adopts a very different conformation to Gal5. It is possible, therefore that while both Gal5 and Gal6 bind to subsites -2 to +1, they may occupy different distal subsites reflecting their conformational variation. For example Gal5 may bind preferentially from -3 to +2, but Gal6, because of its tight helical structure cannot access the -3 subsite but in this conformation it can bind to an additional negative subsite (-4 subsite), which is not accessible to Gal5. This proposal would explain why Gal6 is preferentially cleaved into Gal2 and Gal4. In Figure 4.7 it is proposed that the enzyme contains a preponderance of negative subsites, based primarily on the generation of galactose and Gal3 from galactotetraose. The opposite, however, may also be feasible (subsites from -2 to +4). To explore this issue further requires labelling the products released from the negative subsites after k_2 (the glycosylation step) using O^{18} water and mass spectrometry, which would distinguish, for example, between Gal6 binding -4 to +2 or +2 to -4 (Mcgregor *et al.*, 2016).

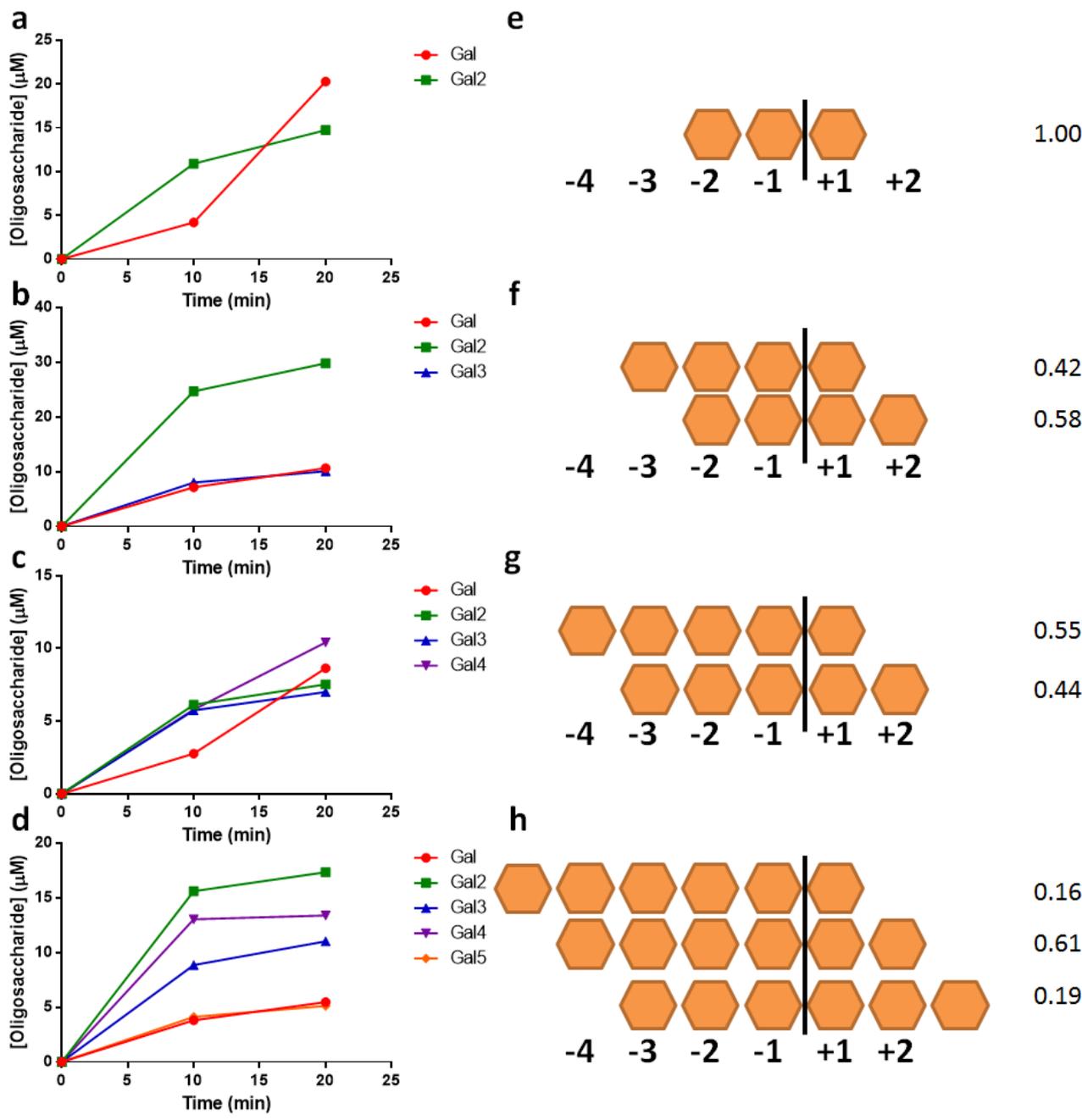


Figure 4.7 Products of oligosaccharide hydrolysis and potential binding modes of Bt4668. Product profiles as determined by HPLC analysis of assay samples at 0, 10 and 20 min of incubation at 37 °C of Gal3 (a), Gal4 (b), Gal5 (c) and Gal6 (d) digestion with BT_4668. And possible productive binding modes of Gal3 (e), Gal4 (f) Gal5 (g) and Gal6 (h) deduced from product profiles. Subsite number is given as +/- #, oligosaccharides represented as blue hexagons and yellow symbol denotes the bond which is hydrolysed in the reaction.

4.2.2.4 Location of BT_4668

LipoP analysis of the BT_4668 sequence indicated that the protein possess a type II signal peptide, allowing for trafficking to the outer membrane where it can be anchored to the outer face of the lipid membrane via a cysteine at position 20. To confirm this experimentally, a *B. thetaiotaomicron* mutant expressing inactive BT_4668 (E180A/E292A mutant lacking both catalytic residues) was generated. Inactive BT_4668 was cloned into a suicide vector and conjugated into *B. thetaiotaomicron tdk-* (see Chapter 2.9.4 for detailed methods). Through homologous recombination the inactive *bt_4668* sequence integrated into the genome replacing the wild type copy of the gene. Replacement of enzymes with inactive variants is preferable to entire gene deletions as PUL encoded enzymes and glycan binding proteins are thought to interact at the cell surface (Reeves *et al.*, 1997; Shipman *et al.*, 2000); an inactive enzyme would allow these interactions to remain intact.

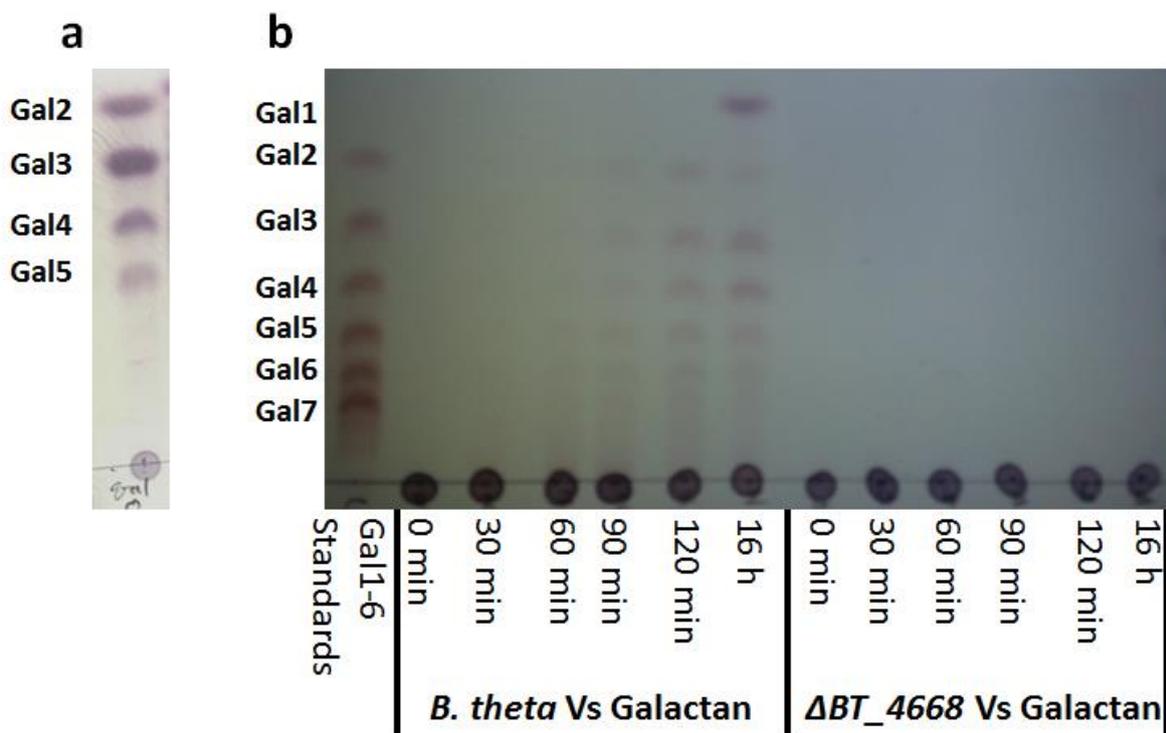


Figure 4.8 Whole cell assays using wild type *B. thetaiotaomicron* and the Δbt_{4668} mutant. Galactooligosaccharides used as growth substrate for wild type and Δbt_{4668} *B. thetaiotaomicron* (a). Surface enzyme assays of whole cells grown on oligosaccharides. Wild type and Δbt_{4668} *B. thetaiotaomicron* were grown to mid log phase before being harvested, concentrated and incubated in PBS with 1 mg/ml galactan at 37 °C over 16 h (b). *B. theta* refers to wild type *B. thetaiotaomicron*.

Whole cell assays were performed using *B. thetaiotaomicron* cells with galactan in aerobic conditions. As strict anaerobes, *Bacteroides* are unable to generate energy for the active transport required to import oligosaccharides generated by surface enzymes (Jordan *et al.*, 2013), leaving all oligosaccharides generated in solution. To upregulate PUL-Gal in both wild type and the mutant the cells were grown to exponential phase on minimal media (MM) + 0.5 % galactooligosaccharides from 3 h digestion of 1 % galactan with 1 μ M BT_4668 (Figure 4.8a). Products generated from incubation of galactan with wild type *B. thetaiotaomicron* (Figure 4.8b) were similar to that observed for recombinant BT_4668 acting on the polysaccharide (Figure 4.5b). No oligosaccharide products were generated by Δ bt_4668 *B. thetaiotaomicron*, indicating that BT_4668 is indeed present on the cell surface and responsible for extracellular galactan degradation.

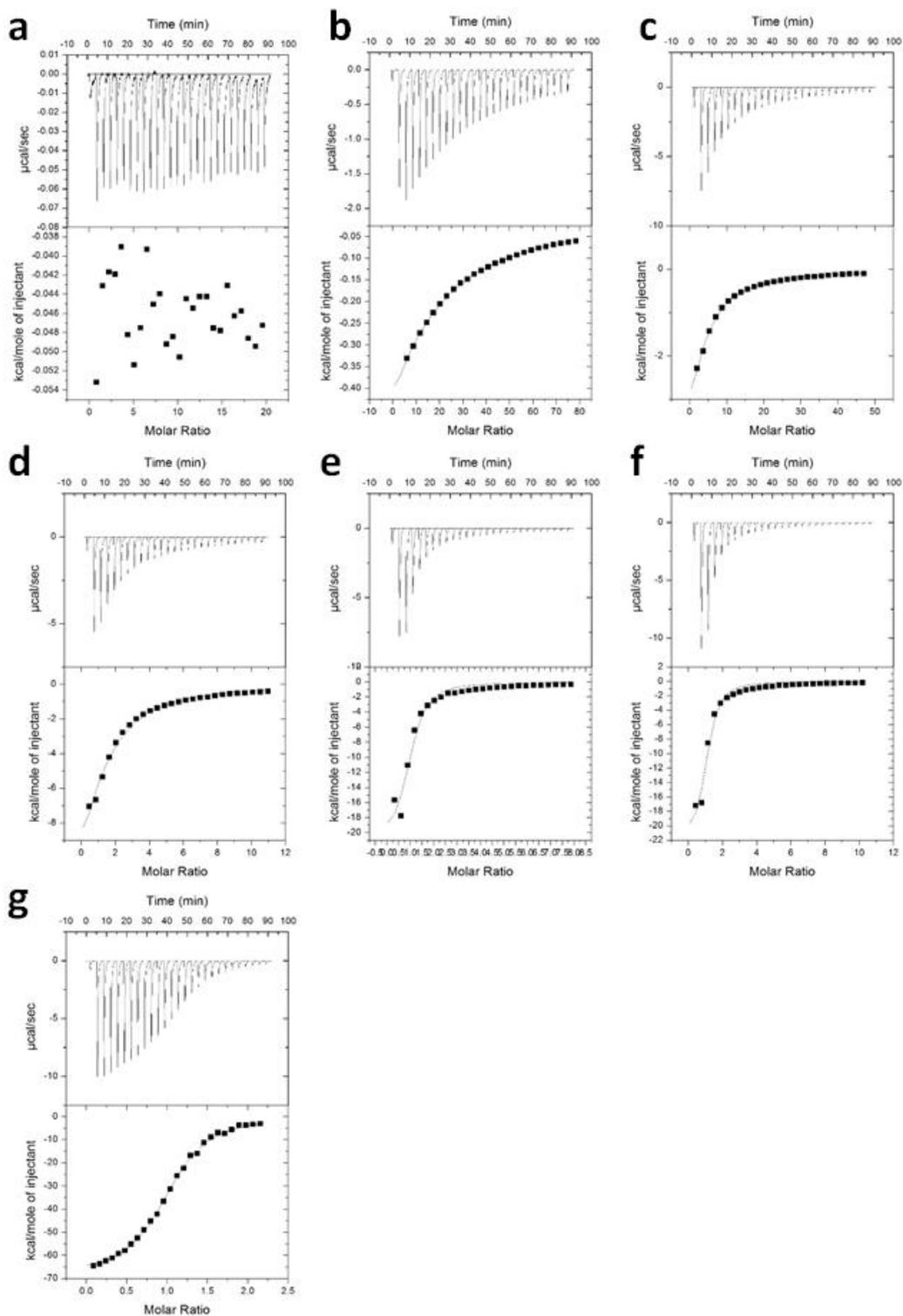


Figure 4.9 BT_4669 binding to ligands measured by isothermal titration calorimetry (ITC). The syringe contains 5 mM Gal2 (a); 5 mM Gal3 (b); 5 mM Gal4 (c); 5 mM Gal5 (d); 5 mM Gal6 (e); 5 mM Gal7 (f); 15 mg/ml galactan (g) titrated against 50 μ M BT_4669. The top half of each panel shows the raw ITC heats; the bottom half, the integrated peak areas fit using a one single binding model by MicroCal Origin 7 software. ITC was carried out in 50 mM Na/HEPES, pH 7.5 at 25 $^{\circ}$ C. Binding kinetic data are presented in Table 4.3.

4.2.2.5 Galactan PUL SusD-homologue, BT_4670, and Surface Glycan Binding Protein, BT_4669

Bioinformatic analysis of BT_4670 revealed significant homology to SusD, a known glycan binding protein from the *sus* locus (starch utilization system) of *B. thetaiotaomicron*. BT_4669 lacked homology to any known GH family; this, combined with the location of *bt_4669* adjacent to the *susD* homologue, indicates that the protein is most likely a surface glycan binding protein (SGBP). Based on these predicted activities, BT_4669 and BT_4670 are likely expressed as lipoproteins on the external face of the *B. thetaiotaomicron* outer membrane. This theory is supported by the presence of type II signal sequences present on both proteins, detected by LipoP. Isothermal titration calorimetry (ITC), a technique which measures directly the enthalpy and affinity of a protein-ligand interaction (enabling the change in entropy and Gibbs free energy of binding to be calculated), was used to assess binding to potato galactan and galactan derived oligosaccharides, Gal2 to Gal7 (Table 4.3). ITC data revealed BT_4669 and BT_4670 bound tightly to galactan (Figure 4.9g & 4.10f) and displayed a preference for galactooligosaccharides with a DP of 6/7, with only a small increase in affinity moving from Gal7 to the polysaccharide (Table 4.3). BT_4669 bound more tightly to its target ligands (Gal7/Galactan) than BT_4670 (Table 4.3 and Figure 4.10). BT_4669 showed relatively low affinity for Gal3 and no measurable binding to Gal2 (Figure 4.9a, b), while binding very tightly to galactan (Figure 4.9g). This binding profile indicates a binding site capable of accommodating oligosaccharides with a DP of 7, possibly slightly longer, due to the relatively small increase from Gal7 to Galactan. BT_4670 bound oligosaccharides with a DP >3 (Table 4.3, Figure 4.10). It appears both BT_4669 and BT_4670 preferentially bind galactan (Table 4.3).

Protein	Ligand	$K_a \times 10^3$ (M^{-1})	$G\Delta \times 10^3$ (Kcal mol $^{-1}$)	$\Delta H \times 10^3$ (Kcal mol $^{-1}$)	$t\Delta S \times 10^3$ (Kcal mol $^{-1}$)	N
BT_4669	Gal2					
	Gal3	0.38 ± 0.02	-3.12	-21.7 ± 4.6	-18.2	0.995
	Gal4	2.12 ± 0.18	-4.55	-28.1 ± 1.9	-23.6	1.06
	Gal5	10.4 ± 1.5	-5.48	-24.4 ± 0.65	-18.9	1.02
	Gal6	124 ± 42	-6.95	-21.9 ± 0.23	-14.9	0.952
	Gal7	138 ± 40	-7.02	-22.6 ± 0.19	-15.6	1.00
	Galactan	304 ± 14	-7.56	-68.4 ± 0.052	-60.8	1.03
BT_4670	Gal2					
	Gal3					
	Gal4	8.12 ± 1.1	-5.33	-8.15 ± 1.1	-2.82	1.05
	Gal5	11.1 ± 1.2	-5.52	-7.62 ± 1.67	-2.11	1.07
	Gal6	16.8 ± 1.4	-5.76	-7.73 ± 0.80	-1.96	1.04
	Gal7	22.0 ± 2.5	-5.92	-7.18 ± 0.79	-1.26	1.09
	Galactan	84.8 ± 2.2	-5.70	-6.21 ± 0.60	-0.51	1.02
BT_4673	Gal2	38.0 ± 2.4	-6.21	-17.9 ± 0.89	-11.7	1.05
	Gal3	48.5 ± 9.2	-6.39	-16.1 ± 2.3	-9.72	1.01
	Gal4	98.5 ± 9.2	-6.80	-12.9 ± 0.42	-6.17	1.04
	Gal5	86.9 ± 1.4	-6.74	-6.06 ± 0.33	-0.67	1.08

Table 4.3 Thermodynamic parameters of the binding of BT_4669, BT_4670 and BT_4673 to their ligands determined by ITC.

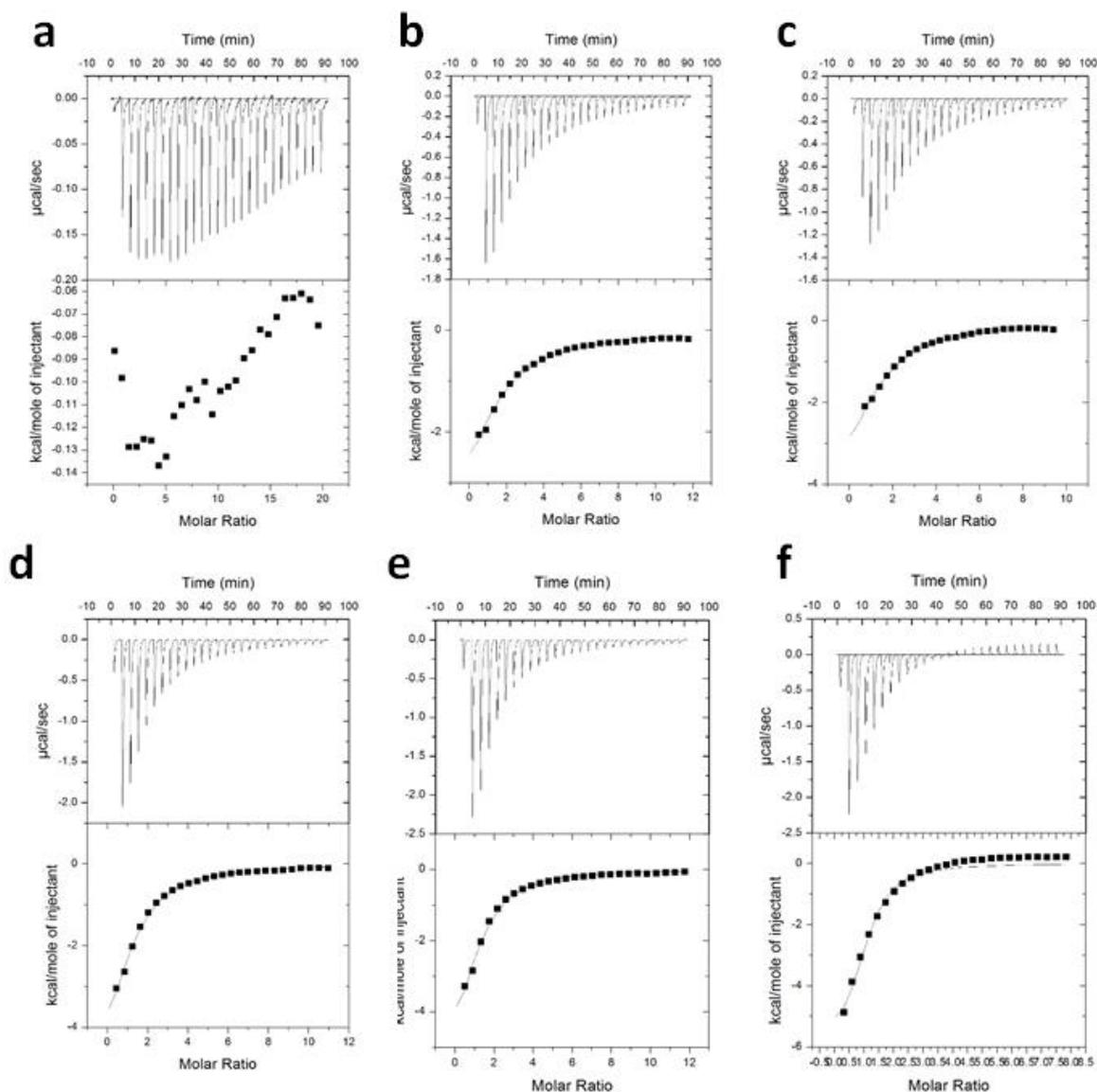


Figure 4.10 BT_4670 binding to ligands measured by ITC. 5 mM Gal3 (a), 5 mM Gal4 (b), 5 mM Gal5 (c), 5 mM Gal6 (d), 5 mM Gal7 (e), 15 mg/ml galactan (f) titrated against 50 μ M BT_4670. The top half of each panel shows the raw ITC heats; the bottom half, the integrated peak areas fit using a one single binding model by MicroCal Origin 7 software. ITC was carried out in 50 mM HEPES 150 mM NaCl, pH 7.5 at 25 $^{\circ}$ C.

4.2.2.6 Galactooligosaccharide degradation by BT_4667

Analysis of the amino acid sequence of BT_4667 revealed significant sequence identity to GH2 enzymes with an N-terminal type I signal peptide and thus is likely located in the periplasm. Activity of the recombinant protein was assessed by incubation with a range of different galactosides (Figure 4.11a). Of the glycans tested BT_4667 demonstrated hydrolytic activity against Gal- β 1,4-GlcNAc, Gal-

β 1,4-Gal (Gal2) and Gal- β 1,4-Glc (Lactose), while Gal- β 1,3-GlcNAc and Gal- β 1,3-Gal remained intact after 16 h incubation (Figure 4.11a) with 1 μ M BT_4667 at 37 °C. This result indicates specificity for Gal- β 1,4-linked substrates and thus the enzyme is a β 1,4-galactosidase.

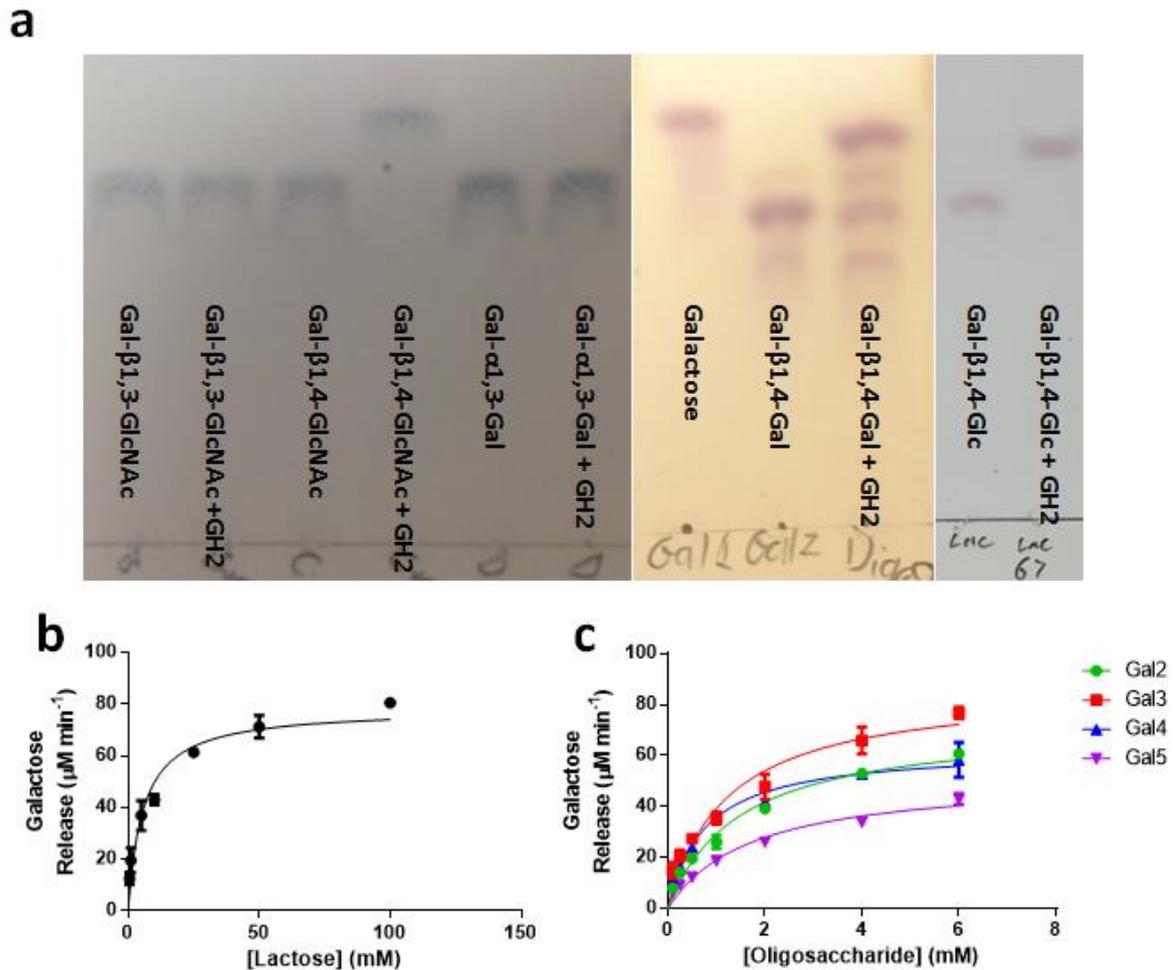


Figure 4.11 BT_4667 activity on galactosides. TLC of BT_4667 (defined as GH2) activity screened against a range of galactosides (a). Michaelis-Menten kinetics of BT_4667 activity against lactose (b) and galactooligosaccharides Gal2-5 (c) in 20 mM phosphate buffer 150 mM NaCl pH 7.5 at 37 °C.

The kinetic parameters of BT_4667 activity on a range of galactooligosaccharides and lactose, a galactoside commonly found in the human diet, were assayed (Figure 4.11b, c). As an exo-acting enzyme BT_4667 hydrolyses linkages in galactosides to release a terminal galactose, the accumulation of which was detected by coupling galactose release with conversion of NAD^+ to NADH

by galactose dehydrogenase activity. NADH accumulation was measured by absorbance at 340nm (Chapter 2.8.1.3). Enzyme activity for each substrate was recorded in Table 4.4. Although BT_4667 demonstrated relatively high and similar activity on the galactooligosaccharides tested compared to lactose, which was hydrolysed slowly (Table 4.4, Figure 4.11c). The moderate reduction in activity on longer galactooligosaccharides may be a result of steric clashes due to the helical structure of the oligosaccharides, which becomes more pronounced as the chain length increases (Cid *et al.*, 2010). In any event, the data indicated that BT_4667 contains a -1 and +1 subsite, which both target galactose.

Substrate	V _{max} ($\mu\text{M min}^{-1}$)	K _M (mM)	k _{cat} (min ⁻¹)	k _{cat} /K _M (min ⁻¹ mM ⁻¹)
Lactose	78.2 ± 4.48	5.75 ± 1.28	15.65 ± 0.62	2.72 ± 0.27
Gal2	72.7 ± 4.33	1.48 ± 0.23	1454 ± 150	980 ± 103
Gal3	86.8 ± 6.36	1.21 ± 0.26	1735 ± 119	1432 ± 147
Gal4	63 ± 2.65	0.76 ± 0.11	1260 ± 105	1652 ± 241
Gal5	51.9 ± 3.37	1.72 ± 0.29	1037 ± 50.9	604 ± 9.36
	V _{max} ($\mu\text{M min}^{-1}$)	K _M (mg ml ⁻¹)	k _{cat} (min ⁻¹)	k _{cat} /K _M (min ⁻¹ ml mg ⁻¹)
RGI	70.35 ± 0.35	2.9 ± 0.21	70.35 ± 0.28	24.25 ± 0.45

Table 4.4 BT_4667 activity on galactooligosaccharides, RGI and lactose. Each reaction was performed in triplicate.

4.2.2.7 Binding of BT_4673 Sensor domain to activating ligand

Bioinformatic data indicate BT_4673 as the HTCS that regulated the galactan PUL in response to appropriate ligands presented to the periplasm. The full length regulator is predicted to contain a histidine kinase A domain (cd00082), HTH_AraC domain (pfam00165), Y_Y_Y domain (pfam07495), periplasmic ligand binding domain (COG3292) and a signal transduction histidine kinase domain (cd00075), using Pfam 27.0 (<http://pfam.sanger.ac.uk/>). The native form of BT_4673 is predicted to span the inner membrane of *B. thetaiotaomicron* binding oligosaccharides in the periplasm via a

sensor domain. DNA encoding the sensor (ligand binding domain) of BT_4673 was cloned into pET28b. The ligand specificity of the resultant recombinant protein, purified by IMAC, was assessed by ITC (Figure 4.12, Table 4.3). The protein displayed highest affinity for Gal4, although affinity was also elevated for Gal5 (Table 4.3). Preference for mid-length oligosaccharides such as Gal4 allows regulation specificity as galactose and, to a lesser extent, Gal2 are found as components of other polysaccharides. This contrasts with the previously characterised ligand binding domain of the HTCS responsible for regulation of the levan PUL, which preferentially binds the monosaccharide fructose (Sonnenburg *et al.*, 2010). Previously described HTCS ligand binding domains are highly specific for oligosaccharide breakdown products of the target glycan of the PUL (Martens *et al.*, 2011; Lowe *et al.*, 2012). The activating ligands typically comprise undecorated oligosaccharides derived from the target polysaccharide backbone, even when the glycan possesses side chains. There has been no cross specificity found for HTCS binding to glycans. Recognition of oligosaccharides allows for differentiation based, not only on sugar content, but also on linkage, explaining why the activating ligands are highly PUL specific (Martens *et al.*, 2011).

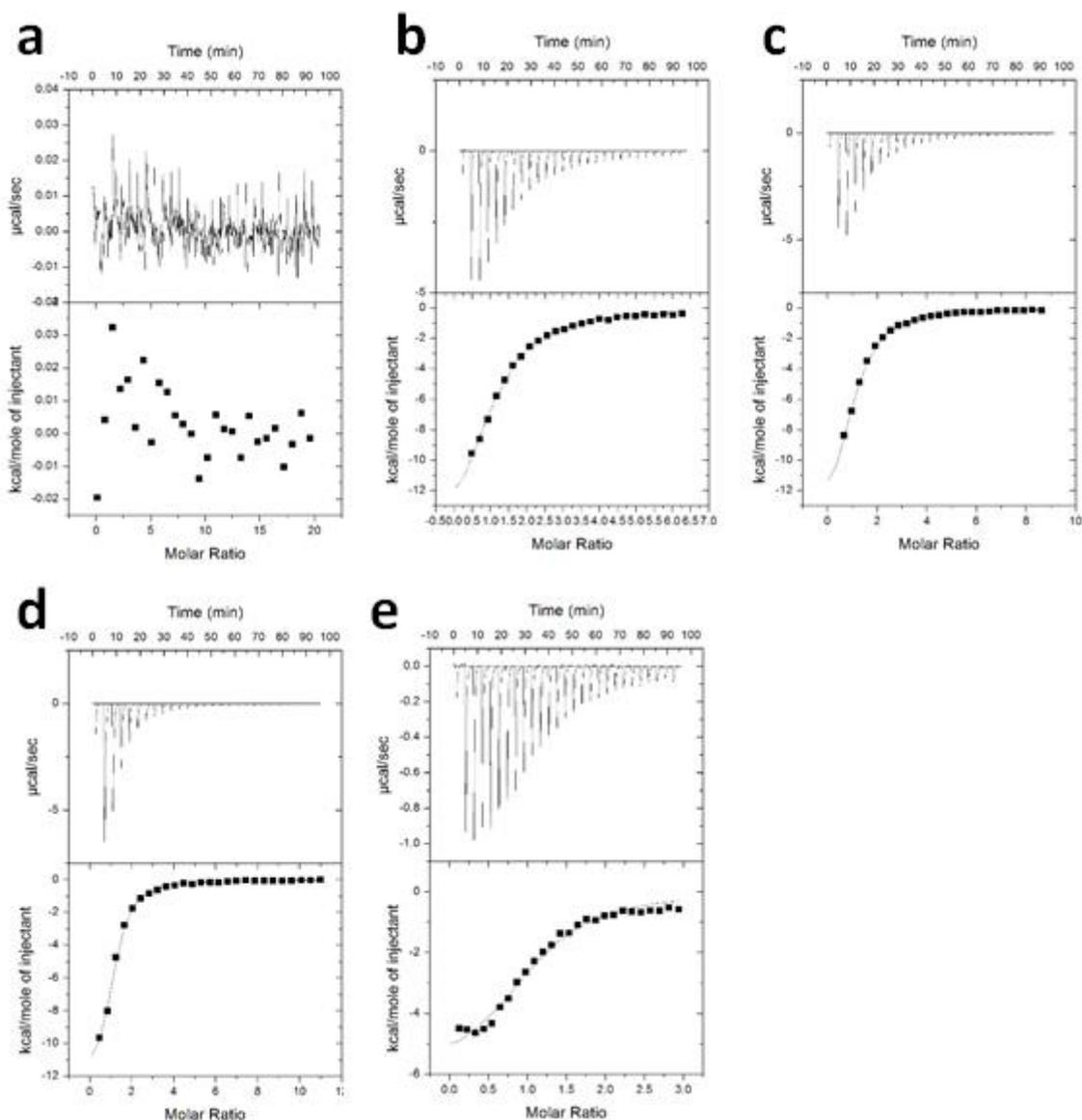


Figure 4.12 BT_4670 binding to ligands measured by ITC. Syringe loaded with galactose (a), Gal2 (b), Gal3 (c), Gal4 (d), Gal5 (e) all at 5 mM, titrated against 50 μ M BT_4673. The top half of each panel shows the raw ITC heats; the bottom half, the integrated peak areas fitted using a one single binding model by MicroCal Origin 7 software. ITC was carried out in 50 mM HEPES 150 mM NaCl, pH 7.5 at 25 $^{\circ}$ C.

The crystal structure of BT_4673 ligand binding domain was previously solved in apo form (Zhang *et al.*, 2014b), although no structure has been reported for the ligand bound conformation. Screening crystallisation conditions with the aim of growing protein crystals of BT_4673 was attempted.

Crystals formed in a condition (100 mM HEPES pH 7.0, 10 % Polyvynal Pyridone K15) of the Midas

solution set. The crystals harvested were unsuitable for diffraction requiring optimisation of the original condition. After optimisation crystals harvested were found to diffract at 3.9 Å.

Unfortunately resolution was too low to solve the structure.

Substrate	V _{max} ($\mu\text{M min}^{-1}$)	K _M (mM)	k _{cat} (min ⁻¹)	k _{cat} /K _M (min ⁻¹ mM ⁻¹)
Gal2	No activity			
Gal3	165 ± 8.38	0.75 ± 0.11	165 ± 47.1	231 ± 46.9
Gal4	208 ± 12.0	0.52 ± 0.09	208 ± 36.1	399 ± 61.5
Gal5	274 ± 8.75	0.73 ± 0.07	274 ± 25.8	374 ± 21.1
Gal6	275 ± 18.7	0.18 ± 0.05	275 ± 25.3	1512 ± 5.77
	V _{max} ($\mu\text{M min}^{-1}$)	K _M (mg ml ⁻¹)	k _{cat} (min ⁻¹)	k _{cat} /K _M (min ⁻¹ ml mg ⁻¹)
Galactan	306 ± 10.2	0.29 ± 0.03	1224 ± 31.7	4154 ± 294
RGI	150 ± 15.3	8.28 ± 1.78	74.75 ± 6.12	9.03 ± 0.92

Table 4.5 BACOVA_05493 activity on galactooligosaccharides, galactan and RGI. Reactions performed in triplicate.

4.2.2.8 BACOVA_05493, an unassigned glycoside hydrolase in the *B. ovatus* galactan PUL

Sequence alignment of galactan PULs from *Bacteroides* uncovered an open reading frame (ORF) annotated as belonging to an unspecified glycoside hydrolase family on the CAZy database. This ORF is highly conserved in *B. ovatus* ATCC 8433, *B. cellulosityicus* DSM 14838, *B. caccae* ATCC 43185 and *B. finefoldii* DSM 17565. The ORF is, however, absent from the *B. thetaiotaomicron* galactan PUL (Figure 4.13a). The ORF from *B. ovatus* encoding BACOVA_05493 was cloned into the *E. coli* pET28 expression vector and the encoded protein expressed in TUNER *E. coli* cells. BACOVA_05493 shares a high degree of identity with BACCAC_02093 (82%), BACCELL_05455 (63%) and BACFIN_00640 (80%) (Figure 4.13b). Each of the encoding genes are located between the SusC-homologue and the HTCS of each *Bacteroides* galactan PUL.

Recombinant BACOVA_05493 was screened for activity against galactosides (Figure 4.14a). The only product observed when the enzyme was active on a substrate was galactose and undigested Gal2 (Figure 4.14a), indicating BACOVA_05493 is an exo-acting β 1,4-galactosidase that was unable to hydrolyse Gal2. The enzyme was active on RGI side chains and galactan. BACOVA_05493 demonstrated similar activity on Gal3 to Gal5, with a significant increase against Gal6 (Table 4.5). As k_{cat} was similar on all of the oligosaccharides including Gal6; the higher activity against the hexasaccharide reflected a decrease in K_M (Table 4.5) indicating BACOVA_05493 displays a much higher binding affinity for the longer oligosaccharide substrates.

Lack of activity on Gal2 indicates that substrate must bind at least to three subsites to form a productive complex with the enzyme. The preference for Gal6 may suggest a +5 subsite, but this seems unlikely as the enzyme lacks functional +3 and +4 subsite as the enzyme displays similar activity from Gal3 to Gal5. More likely galactooligosaccharides adopt a six-fold screw axis helical structure which is in its most stable conformation when the DP is ≥ 6 (Cid *et al.*, 2010). Thus, the substrate binding site of the enzyme, extending from -1 to +2, is optimized to bind to galactohexaose in its six-fold helical conformation.

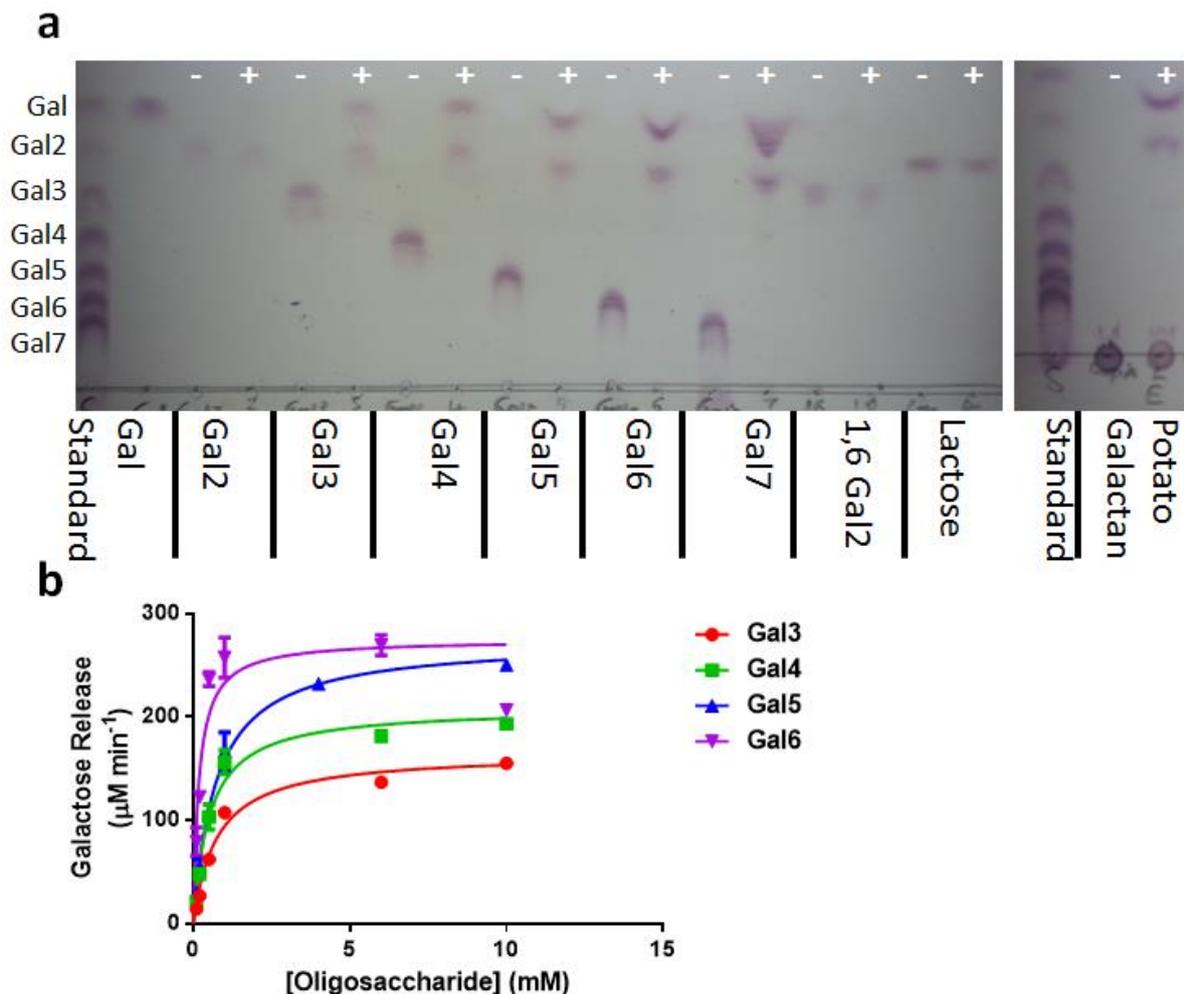


Figure 4.14 BACOVA_05493 activity on galactooligosaccharides. TLC of BACOVA_05493 activity screened against galactooligosaccharides, 1 mM Gal2-7, 1 mM lactose and 1 mg/ml galactan over 16 h in 20 mM sodium phosphate 150 mM NaCl pH 7.5 at 37 °C (a). Minus symbol indicates sample without enzyme and plus symbol indicates where enzyme is included. Michaelis-Menten plots of BACOVA_05493 activity on galactooligosaccharides, Gal3-6, in 20 mM sodium phosphate 150 mM NaCl pH 7.5 at 37 °C (b).

4.2.2.9 Galactan PUL Mutants

Genomic mutations were introduced into each of the genes that comprise the *B. thetaiotaomicron* and *B. ovatus* PUL-Gal to either inactivate the encoded proteins, by substitution of catalytic residues, or entire gene deletions when critical residues were unclear (Chapter 2.9.4). The mutants generated were sequenced to ensure the intended genomic mutation was achieved. Primers used for sequencing anneal to the gene in the case of gene inactivation, while the flanking sequence used to replace the target gene in gene deletions were sequenced. In both strategies the regions around the

mutations were also sequenced to ensure insertion occurred in the right place in the genome.

Mutants are named using the locus tag of the inactivated/deleted gene, $\Delta btXXXX$ or $\Delta bacovaXXXXX$.

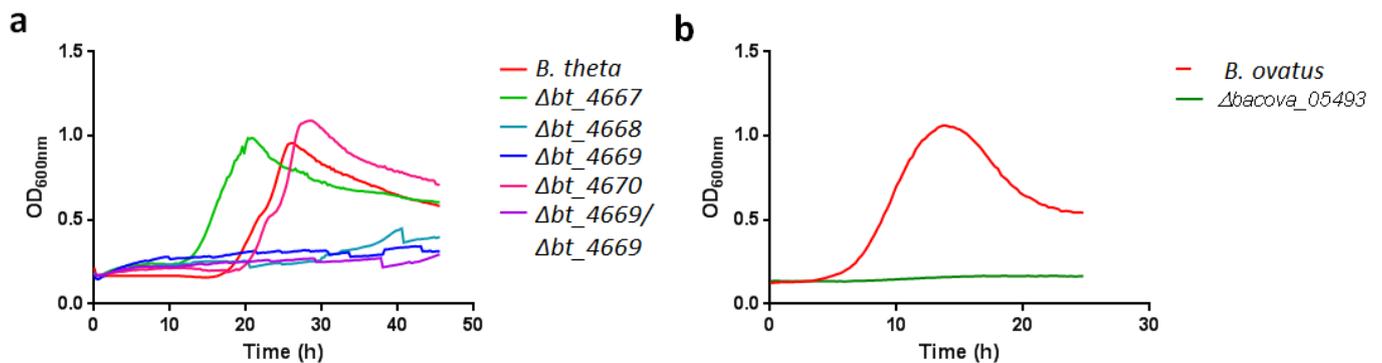


Figure 4.15 Growth of galactan PUL mutants on galactan. Mutants of the components of the *B. theta* galactan PUL grown in MM+ 0.5 % galactan (a). Comparison of growth of *B. ovatus* and $\Delta bacova_{05493}$ grown in MM+ 0.5% galactan (b). Growths were performed using 96-well format with 200 μ l (minimal medium) MM + 0.5 galactan at 37 °C using an automatic plate reader in anaerobic conditions. *B. theta* refers to wild type *B. theta* and *B. ovatus* refers to wild type *B. ovatus*. Each growth was performed in triplicate, errors not shown due to the number of data points.

The growth of each mutant was assessed on galactan minimal medium (MM) as described in Chapter 2.9.1. Previously published data suggest the SusD-homologue in a glycan utilisation system is vital for utilisation and uptake of oligosaccharides generated at the cell surface (Shipman *et al.*, 2000; Koropatkin *et al.*, 2008). Here, Δbt_{4670} (encodes SGBP) demonstrated only a slight phenotype when grown on galactan, while Δbt_{4669} (encodes SusD-homologue) shows no growth (Figure 4.15a). Currently unpublished data regarding glycan utilisation systems that target simpler glycans, dextran and inulin, show similar phenotypes when the associated SusD-homologue is deleted, suggesting this may be a common trait (Bolam unpublished data). No phenotype observed when Δbt_{4667} (encodes the GH2 β -galactosidase) was grown on galactan (Figure 4.15a). Lack of a phenotype may be the result of redundancy of GH2 enzymes in the *B. theta* genome, as there are 32 genes encoding GH2 enzymes in the *B. theta* genome (Xu *et al.*, 2003). Despite no upregulation of these genes in response to galactan (Martens *et al.*, 2011), basal expression may be sufficient to degrade galactooligosaccharides in the periplasm. Inactivation of Δbt_{4668} resulted in a complete lack of growth on galactan (Figure 4.15a), indicating that the endo-

activity of the surface GH53 galactanase is vital for galactan utilisation. This reliance on BT_4668 suggests undigested galactan is too large for transport through the galactan SusC-homologue BT_4671.

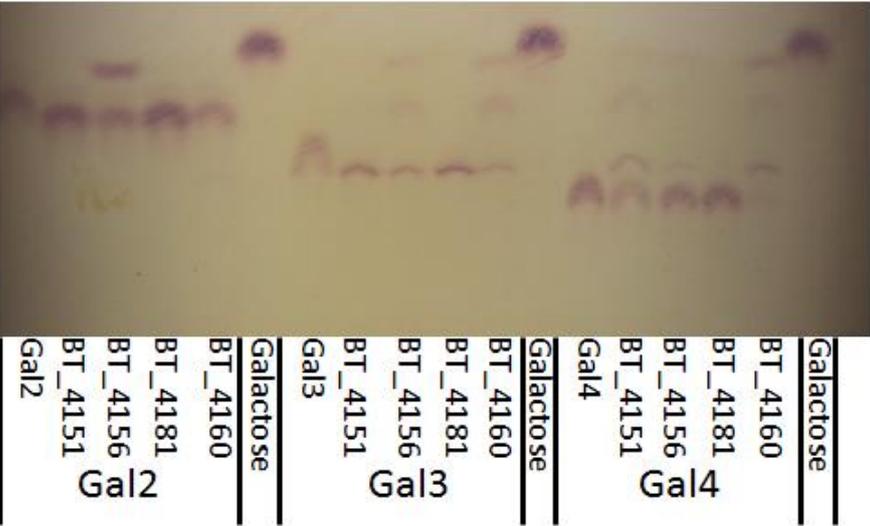
The extra glycoside hydrolase present in the *B. ovatus* was inactivated in the genome to create the *Δbacova_05493* mutant. When grown on galactan the mutant was unable to utilise the polysaccharide (Figure 4.16b). This result was interesting as the other components of the *B. ovatus* galactan PUL share a very high degree of identity to those of the corresponding *B. thetaiotaomicron* locus. This suggests they provide the same function in the *B. ovatus* PUL and thus it is unclear why BACOVA_05493 should be an essential component of the galactan degrading apparatus.

4.2.3 RGI Galactosidases

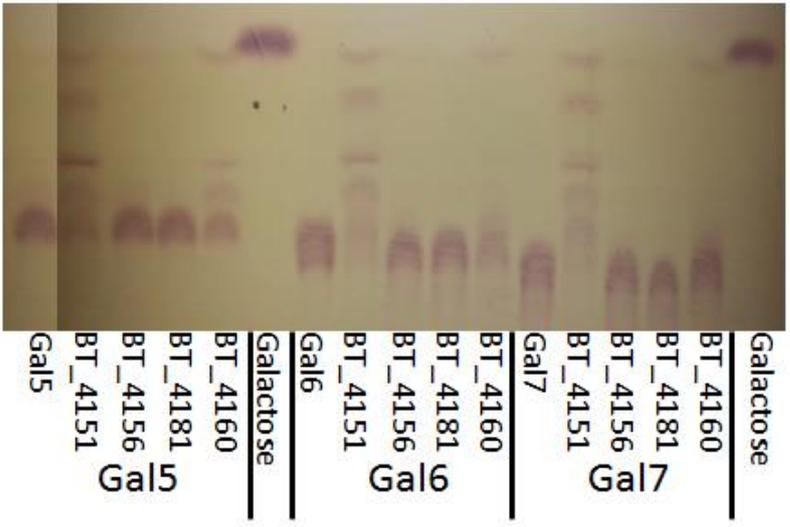
Once the majority of galactan is liberated from RGI, short oligosaccharides remain attached to the RGI backbone. The *B. thetaiotaomicron* RGI PUL, transcribed in response to growth on the polysaccharide encodes three GH2 proteins (BT_4151, BT_4156 and BT_4181), while BT_4160 is located in GH35. All these proteins are predicted to display β 1,4-galactosidase activity. Sequence analysis of each protein using LipoP predict all are present in the periplasm of *B. thetaiotaomicron*.

TLC was used to screen the activity of BT_4151, BT_4156, BT_4160 and BT_4181 on a range of galactooligosaccharides (Figure 4.16a, b), galactan (Figure 4.16c) and RGI (Figure 4.16d). BT_4151 partially hydrolysed Gal4-7 and galactan, indicating a clear specificity for galactan and galactooligosaccharides with a high DP (Figure 4.16).

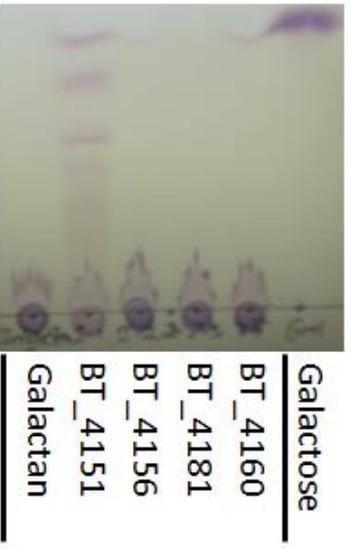
a



b



c



d

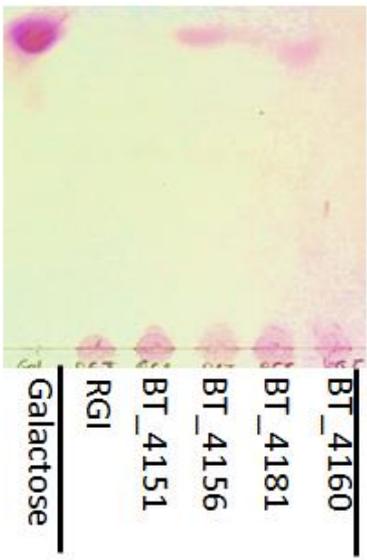


Figure 4.16
BT_4151,
BT_4156, BT_4160
and BT_4181
activity screening
TLCs. Enzymes were incubated for 20 h with 1 mM Gal2-4 (a), Gal5-7 (b), 1 mg/ml galactan (c) and 1 mg/ml RGI (d) in 20 mM sodium phosphate 150 mM NaCl pH 7.5 at 37 °C. Stopped reactions were run alongside galactooligosaccharide standards, Gal- Gal7.

BT_4156 was the only enzyme active on Gal2 (Figure 4.16a). The galactosidase, however, was active on RGI (Figure 4.16d), releasing galactose, suggesting this enzyme is specifically targeting the shortest galactooligosaccharides attached to the RGI backbone. The lower activity of BT_4156 than BT_4151 on RGI (Table 4.6 and Figure 4.17), reflected by an elevated K_M , may result from less available substrate for BT_4156 within RGI. Previous work showed BT_4156 removes the last galactose in the galactan side chain (Zhang 2014). Deletion of BT_4156 caused a slight growth phenotype when the Δbt_{4156} was grown in MM+0.5% RGI (Figure 4.18b). Although the Δbt_{4156} did achieve a cell density similar to that of the wild type *B. thetaiotaomicron*, the mutant demonstrated a longer lag phase than the wild type bacterium (Figure 4.18b).

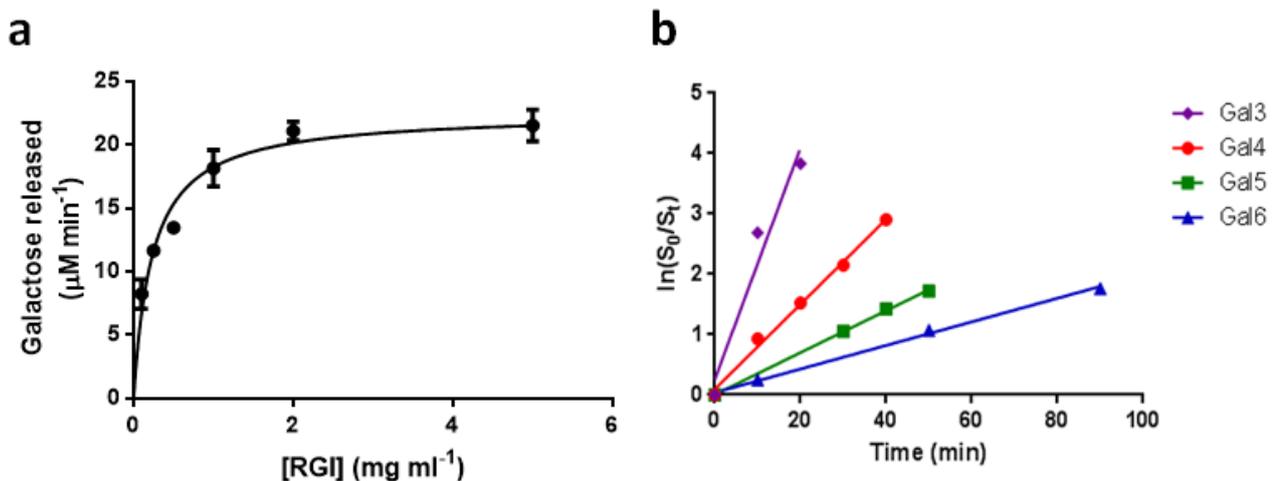


Figure 4.17 BT_4151 and BT_4160 activity. Michaelis-Menten plots of BT_4151 activity on RGI (a). Plots of the inverse rate of substrate depletion of defined galactooligosaccharides when incubated with BT_4160 (b) in 20 mM sodium phosphate 150 mM NaCl pH 7.5 at 37 °C to give catalytic efficiency (k_{cat}/K_M).

BT_4181 was inactive on all substrates tested (Figure 4.16), a faint streak is present but this was overspill from the adjacent reaction carried over during staining (Figure 4.16d). On further investigation the catalytic nucleophile of BT_4181 was missing in the primary sequence of the protein (Figure 4.19a). This enzyme may have lost catalytic function but retained binding to act as a glycan binding protein to aid degradation. The protein, however, showed no binding to RGI or

galactan as ligand (Figure 4.19c,d), and exhibited no synergistic effects when co-incubated with BT_4151, BT_4156 or BT_4161 (Figure 4.19b).

In the activity screen BT_4161 completely degraded all galactooligosaccharides with the exception of Gal2 which remained as a faint band after the digest (Figure 4.16a, b). The enzyme demonstrated very low activity on galactan (Figure 4.16c) but generated more galactose from RGI (Figure 4.16d). Activity of BT_4160 on galactooligosaccharides was tested using HPAEC and substrate concentrations sufficiently below the K_M to allow k_{cat}/K_M to be determined from the inverse of the rate of substrate depletion (Table 4.6, Figure 4.17). BT_4160 displayed maximum activity against Gal3, although the decline in catalytic efficiency against longer oligosaccharides was modest. This indicated that BT_4160 contains three subsites involved in productive substrate binding. The minimal activity against galactan (inferred from TLC, Figure 4.16c and Table 4.6) likely reflects the very low concentration of available substrate (galactose at the non-reducing end of the polysaccharide). The native substrate for this enzyme would most likely be free galactooligosaccharides and those which decorate fragments of RGI in the periplasm, and hence would not require an extended binding site.

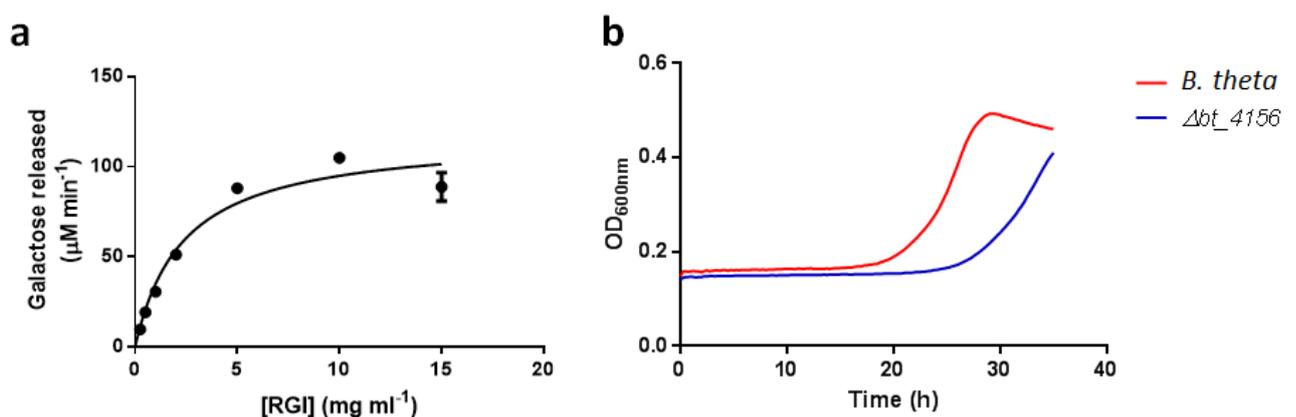


Figure 4.18 Characterisation of BT_4156. Michaelis-Menten plot of BT_4156 activity on RGI (a). Growths of the *B. theta* Δbt_{4156} mutant and wild type *B. theta* in MM + 0.5% RGI (b). Growths were performed in 200 µl media in 96-well format under anaerobic conditions at 37 °C. Growths were recorded by absorbance (A_{600nm}) using an automatic plate reader.

Substrate	V _{max} ($\mu\text{M min}^{-1}$)	K _M (mg ml^{-1})	k _{cat} (min^{-1})	k _{cat} /K _M ($\text{min}^{-1} \text{ml mg}^{-1}$)
BT_4151				
RGI	22.5 ± 0.84	0.24 ± 0.24	22.5 ± 0.59	95.7 ± 1.80
Galactan	3.57 ± 0.33	0.20 ± 0.08	3.57 ± 0.26	17.5 ± 8.54
BT_4156				
RGI	117 ± 7.50	2.36 ± 0.48	117 ± 9.19	49.68 ± 5.07
BT_4160				
RGI				2.07 ± 0.28
				k_{cat}/K_M ($\text{min}^{-1} \text{mM}^{-1}$)
Gal3				198
Gal4				70.4
Gal5				34.7
Gal6				19.5

Table 4.6 BT_4151, BT_4156 and BT_4160 activity. Reactions were performed in triplicate with the exception of galactooligosaccharides, where substrate was limited.

Figure 4.19 Characterisation of BT_4181. Sequence alignment of BT_4181 with typical GH2 family enzymes, LacZ from *E. coli*, BT_4156 and BT_4151 from *B. thetaiotaomicron*. Catalytic residues identified by, blue General acid/base, green, catalytic nucleophile, red, missing acid/base in BT_4181 (a). Relative activity of 1 μ M BT_4151, BT_4156 and BT_4160 alone and in combination with 1 μ M BT_4181 on 1 mg/ml RGI (b) in 20 mM sodium phosphate 150 mM NaCl pH 7.5 at 37 °C. Relative activity given as a percentage of galactose released of the total galactose available. ITC titration of 10 mg/ml galactan (c) and RGI (b) against 50 M BT_4181 in 20 mM HEPES 150 mM pH 7.5.

4.2.3.1 BT_4158

Bioinformatic analysis of BT_4158 showed that the enzyme belongs to a new carbohydrate esterase (CE) family. The esterase demonstrates activity on 4-nitrophenyl-acetate, monitored by accumulation of the chromogenic product 4-nitrophenolate at $A_{405\text{nm}}$ (Figure 4.20a), and acetylated birchwood xylan (Figure 4.20b). BT_4158 demonstrated no activity on RGI or RGI pre-treated with the β -galactosidases encoded by RGI PUL (removes the galactan side chains; Figure 4.20b). On further investigation, the technique used to purify RGI from plant cell walls was found to include an alkali treatment step. As acetyl groups are alkali labile they were most likely removed during purification of the polysaccharide. Previous work showed BT_4158 activity on RGI (Zhang 2014), however, this was an older batch of the polysaccharide purified by a slightly different method to that used for the RGI batch used here. The difference is most likely the severity of the alkali treatment used by Megazyme in the polysaccharide purification process, explaining the difference in observed activity on RGI.

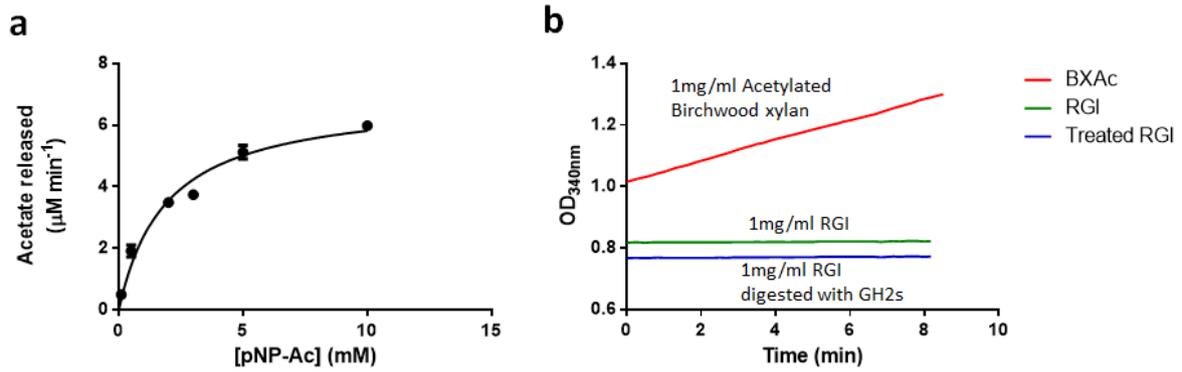


Figure 4.20 BT_4158 activity on acetylated substrates. Michaelis-Menten plot of the acetyl esterase activity of BT_4158 activity on 4-nitrophenyl-acetate (**a**), on 1 mg/ml acetylated Birchwood xylan (BxAc *red*), RGI (*green*) and RGI pre-treated with BT_4151, BT_4156 and BT_4181 (Treated RGI *blue*) (**b**). BT_4158 was at 1 μM when evaluated against the polysaccharides the synthetic substrate.

4.2.4 Arabinan PUL Characterisation

The arabinan PUL, extending from *bt_0360* to *bt_0369* and *bt_0348*, is activated by and is thus likely to orchestrate the degradation of the polysaccharide. The PUL encodes the surface GH43 enzymes, BT_0360 and BT_0367, which were characterised by Dr Lauren Mackie (Cartmell *et al.*, 2011), while the HTCS ligand binding domain derived from this locus was investigated by Dr Lis Lowe (Martens *et al.*, 2011).

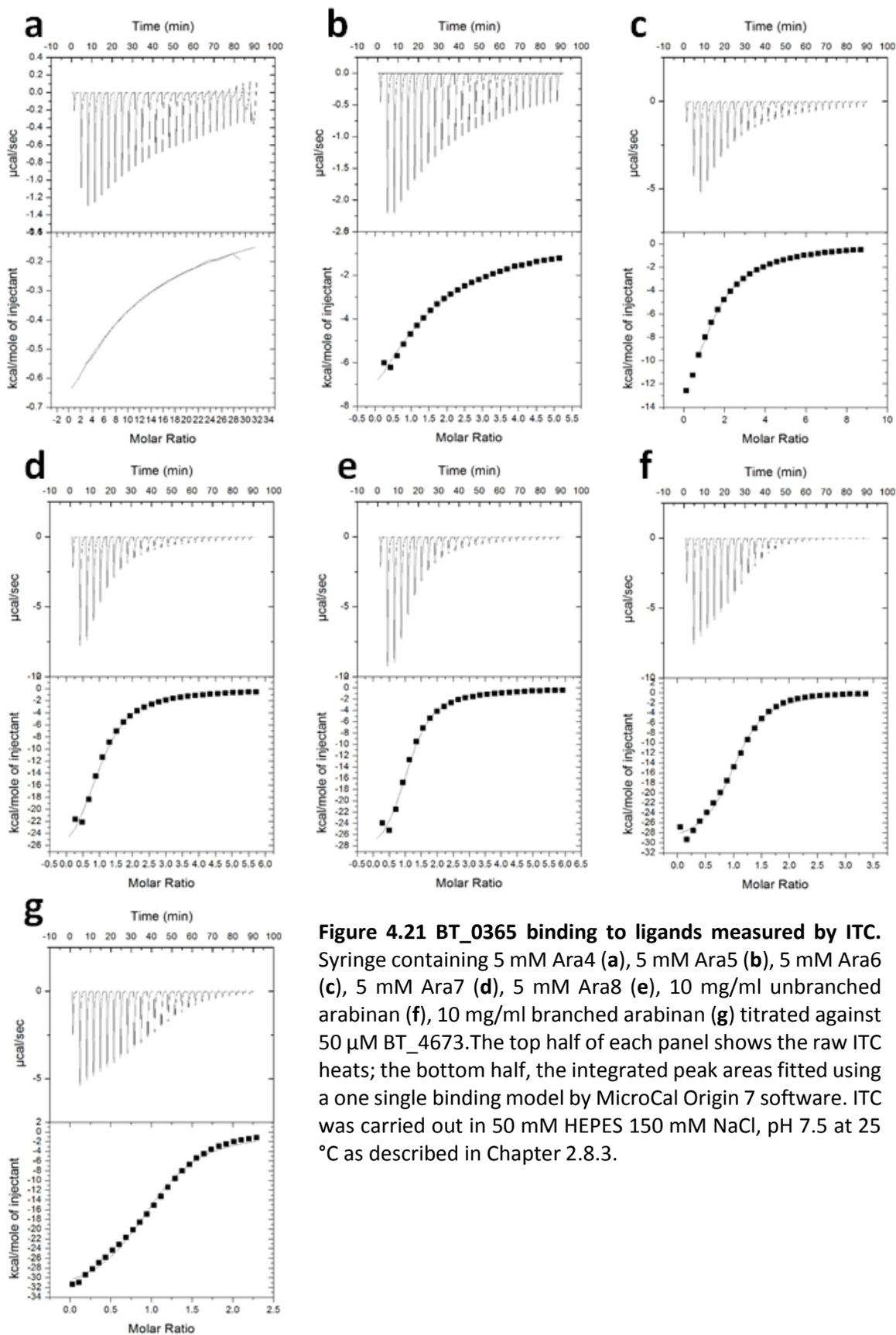


Figure 4.21 BT_0365 binding to ligands measured by ITC. Syringe containing 5 mM Ara4 (a), 5 mM Ara5 (b), 5 mM Ara6 (c), 5 mM Ara7 (d), 5 mM Ara8 (e), 10 mg/ml unbranched arabinan (f), 10 mg/ml branched arabinan (g) titrated against 50 μM BT_4673. The top half of each panel shows the raw ITC heats; the bottom half, the integrated peak areas fitted using a one single binding model by MicroCal Origin 7 software. ITC was carried out in 50 mM HEPES 150 mM NaCl, pH 7.5 at 25 $^{\circ}\text{C}$ as described in Chapter 2.8.3.

4.2.4.1 Surface Binding Proteins

In addition to encoding enzymes, the arabinan PUL contains three putative surface binding proteins; BT_0361 and BT_0363, two SusD-homologues each paired with a specific SusC-homologue, and BT_0365, a SGBP. Binding of each recombinant protein to arabinan and arabinooligosaccharide ligands were measured by ITC. The SGBP bound to α 1,5-linked arabinooligosacchrides with a DP between 4 and 8 (defined as AraX where X is the DP of the molecule), both branched and unbranched arabinan polysaccharide (Figure 4.21). Affinity for the polysaccharides was similar to that of Ara8 (Table 4.7), indicating the ligand binding site is able to accommodate around 8 arabinose residues. Interestingly BT_0365 was able to bind branched arabinan with affinity only 1.4-fold lower than the unbranched polysaccharide (Table 4.7). These data suggest the ligand binding site of BT_0365 is able to tolerate decorations on the arabinan backbone. BT_0361 and BT_0363 were unable to bind any arabinan or arabinan oligosaccharide mixture tested (Figure 4.22). The lack of binding observed here may be due to the requirement of the associated SusC-homologue for binding to occur. Recent structural data suggests the association the SusCD pair is far tighter than previously believed (Glenwright unpublished data 2016). The SusC may have a role in stabilising the associated SusD or may be directly involved in ligand binding.

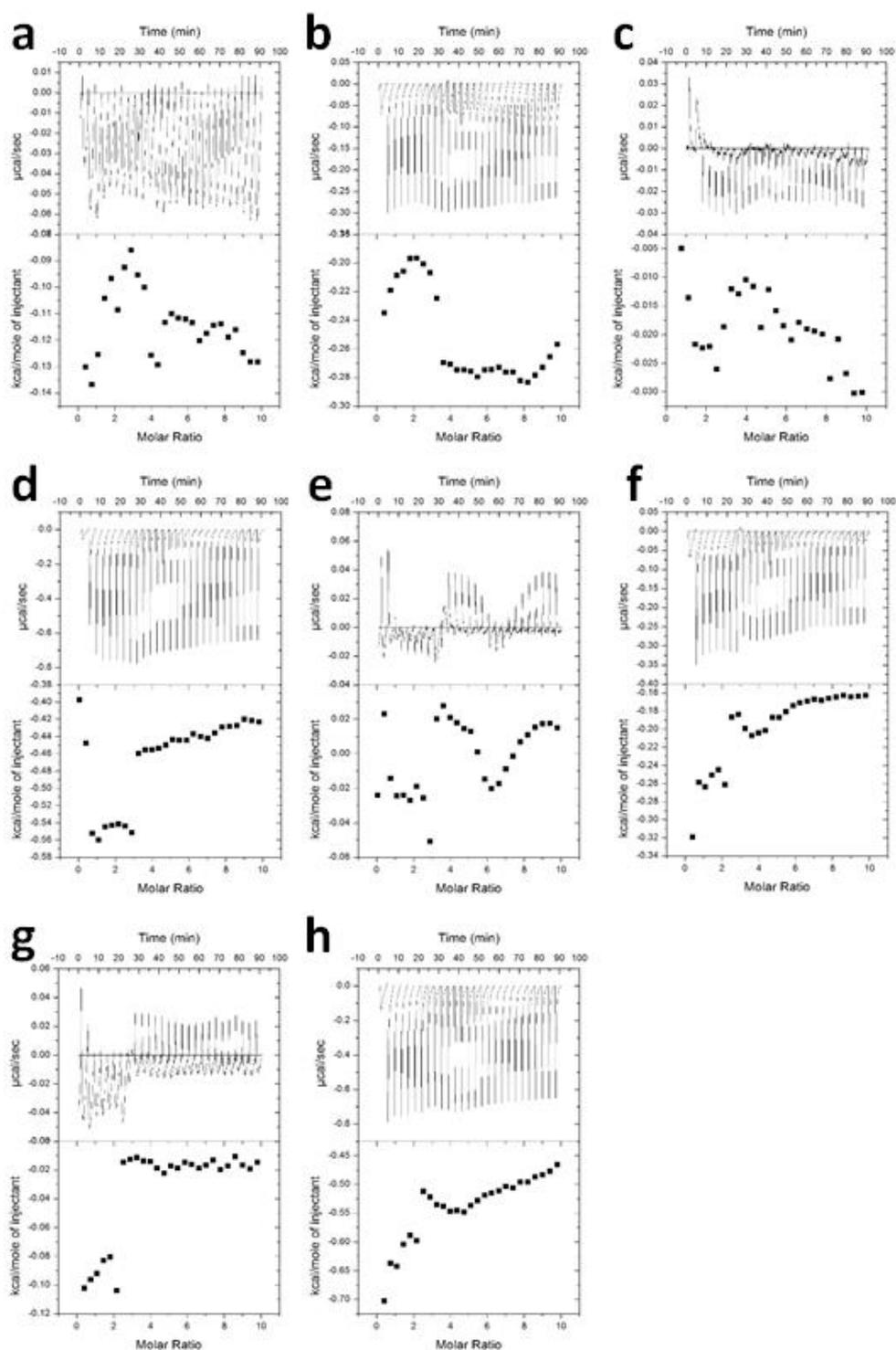


Figure 4.22 BT_0361 and BT_0363 binding to ligands measured by ITC. Binding titration of branched arabinan (a), unbranched arabinan (b), digested branched arabinan (c), digested unbranched arabinan (d) against BT_0361 and branched arabinan (e), unbranched arabinan (f), digested branched arabinan (g), digested unbranched arabinan (h) against BT_0363. The top half of each panel shows the raw ITC heats; the bottom half, the integrated peak areas fitted using a one single binding model by MicroCal Origin 7 software. ITC was carried out in 50 mM HEPES 150 mM NaCl, pH 7.5 at 25 °C.

Ligand	N	$\Delta G \times 10^3$ (kcal mol ⁻¹)	$T\Delta S \times 10^3$ (kcal mol ⁻¹)	$\Delta H \times 10^3$ (kcal mol ⁻¹)	$K_a \times 10^3$ (M ⁻¹)
Ara4	0.99	-3.75	-19.6	-23.4 ± 1.2	0.57 ± 0.12
Ara5	0.98	-5.26	-21.0	-26.2 ± 1.2	7.26 ± 1.2
Ara6	1.02	-4.32	-20.7	-25.5 ± 7.2	14.4 ± 0.72
Ara7	1.01	-6.59	-25.5	-32.1 ± 1.7	68.4 ± 8.2
Ara8	1.08	-6.84	-24.8	-31.6 ± 1.3	112 ± 15
Branched Arabinan	1.02	-7.15	-26.6	-33.8 ± 0.64	170 ± 15
Unbranched Arabinan	1.02	-7.32	-23.0	-30.3 ± 0.53	246 ± 25

Table 4.7 BT_0365 binding to arabinooligosaccharides and arabinan

4.2.4.2 Periplasmic GH51s

Expressed in the periplasm are two GH51 enzymes, BT_0348 and BT_0368 (predicted from signal sequence), and BT_0369, a GH43 arabinofuranosidase that targets α -1,2-linked arabinose units appended to the backbone of arabinan (Cartmell *et al.*, 2011).

The activities of both GH51 enzymes were determined in a reaction coupled with galactose dehydrogenase, an enzyme that recognises galactan and arabinan as substrates generating NADH in a 1:1 ratio with galactose/arabinose (see Chapter 2.8.1.3 for methods). BT_0348 demonstrated low activity on all α -1,5-arabinooligosaccharides tested and linear arabinan (Figure 4.23b, Table 4.8). Catalytic efficiency against branched arabinan was 18-fold greater than that of unbranched arabinan (Table 4.8), indicating the preferred substrate of BT_0348 is present in branched but not linear arabinan. These data suggest BT_0348 is an exo-acting arabinofuranosidase that acts on O2 and/or O3 substitutions found in branched arabinan.

BT_0368 can be considered a more typical α 1,5-arabinofuranosidase highly active on the arabinooligosaccharides tested (Table 4.8, Figure 4.23a), displaying similar activity on each substrate. The enzyme shows low activity on both branched and unbranched arabinan (Table 4.8), particularly on the decorated polysaccharide (Table 4.8). These data suggest BT_0368 shows specificity for

unbranched arabinooligosaccharides, possibly due to a restricted substrate binding site that makes steric clashes with arabinan decorations. The poor activity against linear arabinan reflects the low concentration of available substrate (non-reducing terminal arabinose) (Figure 4.23c, d).

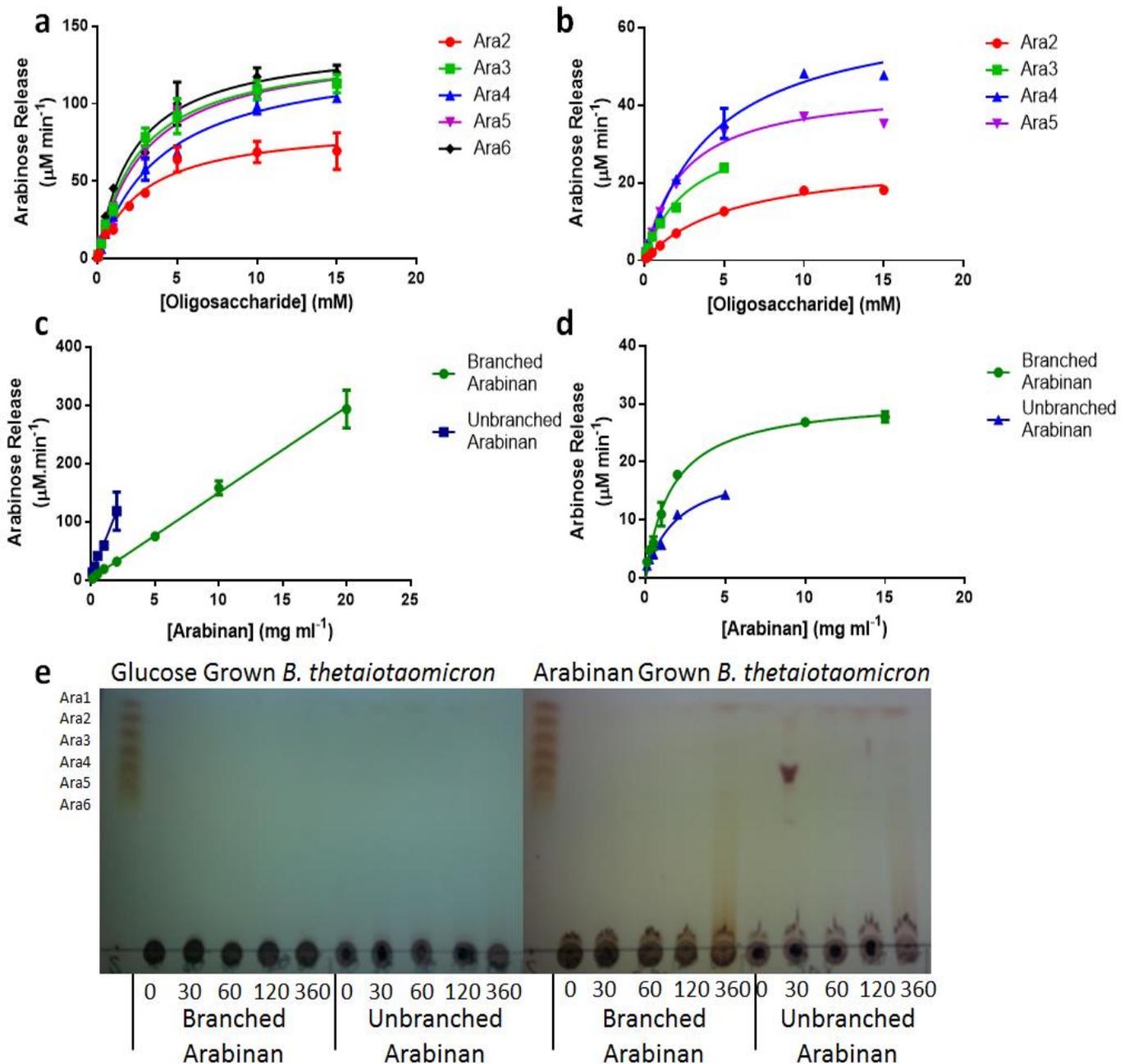


Figure 4.23 Activity of GH51 enzymes BT_0348 and BT_0368. Michaelis-Menten plots of, BT_0368 (a), and BT_0348 (b) activity on arabinooligosaccharides. Activity of BT_0368 (c) and BT_0348 (d) on branched and unbranched arabinan. All reactions were performed in 20 mM sodium phosphate 150 mM NaCl pH 7.5 at 37 °C. Whole cell assays of *B. thetaiotaomicron* grown on glucose and arabinan were concentrated and incubated with branched or unbranched arabinan under aerobic conditions for the times indicated. The supernatants were subjected to TLC, numbers under each lane indicate minutes incubated (e).

Substrate	V _{max} ($\mu\text{M min}^{-1}$)	K _M (mM)	k _{cat} (min ⁻¹)	k _{cat} /K _M (min ⁻¹ mM ⁻¹)
BT_0348				
Ara2	26.6 ± 1.14	5.6 ± 0.53	26.6 ± 1.95	4.79 ± 0.36
Ara3	36.7 ± 2.45	1.55 ± 0.34	36.7 ± 1.8	23.7 ± 3.33
Ara4	65.6 ± 2.81	4.26 ± 0.45	65.6 ± 4.11	15.4 ± 0.42
Ara5	45.1 ± 1.69	2.40 ± 0.26	45.1 ± 3.22	18.8 ± 2.17
	V _{max} ($\mu\text{M min}^{-1}$)	K _M (mg ml ⁻¹)	k _{cat} (min ⁻¹)	k _{cat} /K _M (min ⁻¹ ml mg ⁻¹)
Branched Arabinan	31.21 ± 0.85	1.689 ± 0.15	312.1 ± 2.5	184.8 ± 15.4
Unbranched Arabinan	19.79 ± 1.76	1.904 ± 0.38	19.79 ± 0.19	10.39 ± 0.80
BT_0368				
	V _{max} ($\mu\text{M min}^{-1}$)	K _M (mM)	k _{cat} (min ⁻¹)	k _{cat} /K _M (min ⁻¹ mM ⁻¹)
Ara2	87.18 ± 4.8	2.83 ± 0.44	17436 ± 1854	6165 ± 573
Ara3	136.7 ± 4.7	2.63 ± 0.29	27340 ± 571	10415 ± 1147
Ara4	133.4 ± 5.5	2.95 ± 0.40	26680 ± 3798	9032 ± 993
Ara5	138.8 ± 6.5	2.96 ± 0.34	27760 ± 1566	9366 ± 2299
Ara6	140.7 ± 4.9	2.33 ± 0.25	28140 ± 1808	12087 ± 1162
				k _{cat} /K _M (min ⁻¹ ml mg ⁻¹)
Branched Arabinan				14.70 ± 0.38
Debranched Arabinan				53.93 ± 5.6

Table 4.8 GH51 activity on arabinooligosaccharides and arabinan. Reactions were performed in triplicate.

4.2.4.3 Arabinan PUL Mutants

Genomic mutations were made in components of the arabinan utilisation system through inactivation or deletion of genes. Each mutant was grown in both MM + 0.5% branched arabinan and MM + 0.5% debranched arabinan in the 96-well format in anaerobic conditions at 37 °C. Mutation in the surface GH43 enzymes led to different phenotypes despite each enzyme displaying only modest

differences in activity and specificity *in vitro* (Figure 4.24a, b). The Δbt_0360 *B. thetaiotaomicron* mutant showed only a modest growth defect when cultured on unbranched arabinan (Figure 4.24b), while on branched arabinan growth was indistinguishable from wild type *B. thetaiotaomicron* (Figure 4.24a). The Δbt_0367 *B. thetaiotaomicron* mutant, however, demonstrated little to no growth on either of the arabinan polysaccharides (Figure 4.24). The double mutant, $\Delta bt_0360\Delta bt_0367$ *B. thetaiotaomicron* was unable to grow on either of the arabinan polysaccharides (Figure 4.24). Although initial growth rate was unaffected by loss of the SGBP, the Δbt_0365 *B. thetaiotaomicron* mutant was unable to match the cell density of the wild type *bacterium* (Figure 4.24). Phenotype of the arabinan PUL SusC-homologue deletions were analogous to those observed for the pair of surface GH43 enzymes. Deletion of *bt_0364* caused a dramatic phenotype with no significant growth on either arabinan (Figure 4.24). The *bt_0362* deletion strain grew slightly slower than wild type on branched arabinan (Figure 4.24a) but the mutation did not influence the capacity of the bacterium to grow on linear arabinan (Figure 4.24b).

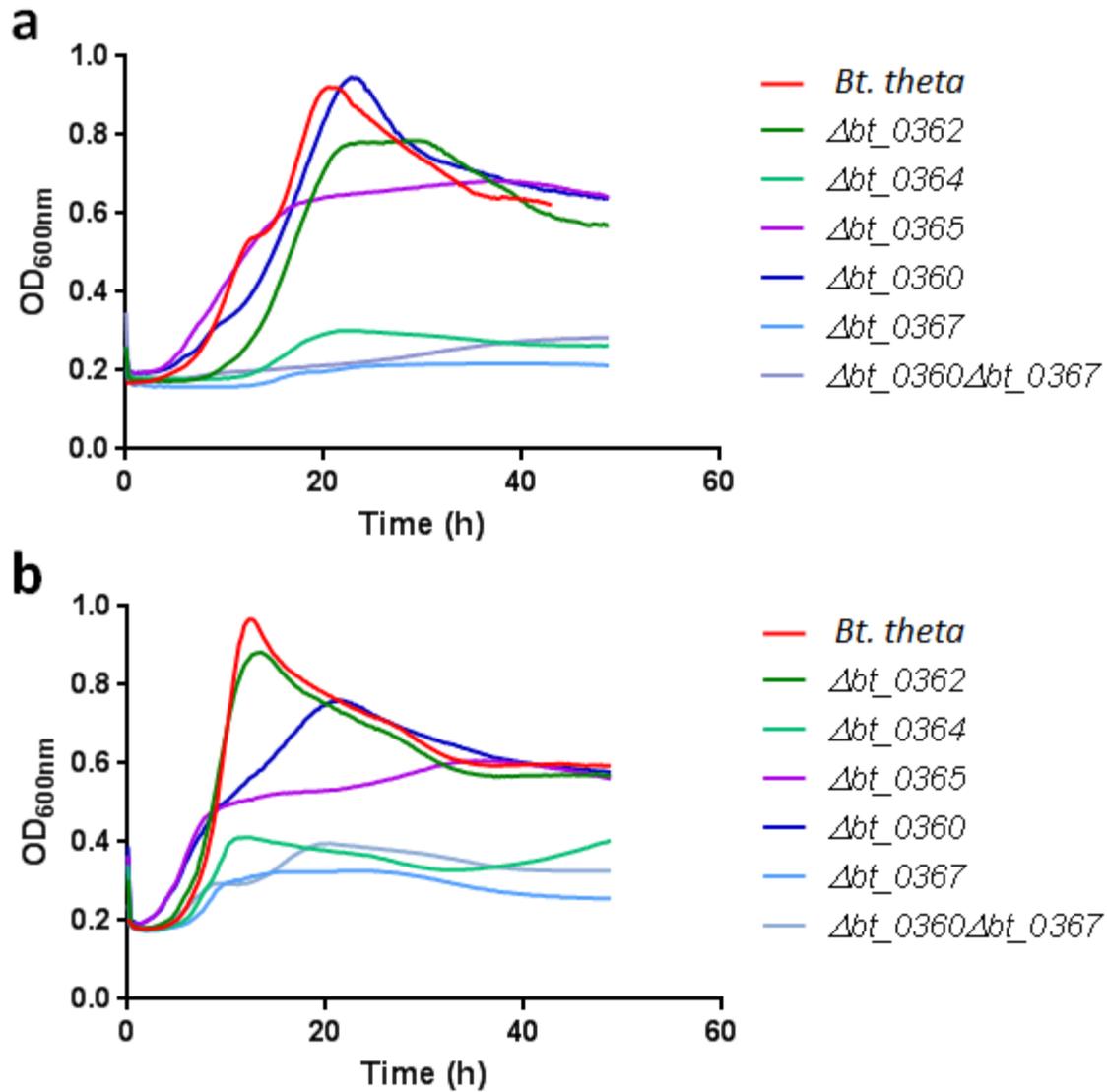


Figure 4.24 Growth of *Bacteroides* with mutations in the arabinan PUL. Growth of the *B. theta* *taotaomicron* arabinan PUL mutants and wild type *B. theta* *taotaomicron* (*Bt. theta*) on in MM+0.5% branched arabinan (a) and MM+0.5% unbranched arabinan (b). Growths were performed in 200 μ l media in 96-well format under anaerobic conditions at 37 °C as described in Chapter 2.9.1. Growth was recorded by absorbance (A_{600nm}) using an automatic plate reader.

4.2.5 Pectin Cross-feeding

Reports of cross-feeding between members of the HGM have shown potential for a currently poorly defined interaction in the human gut (Rakoff-Nahoum *et al.*, 2014; 2016) .

4.2.5.1 Supernatant oligosaccharides

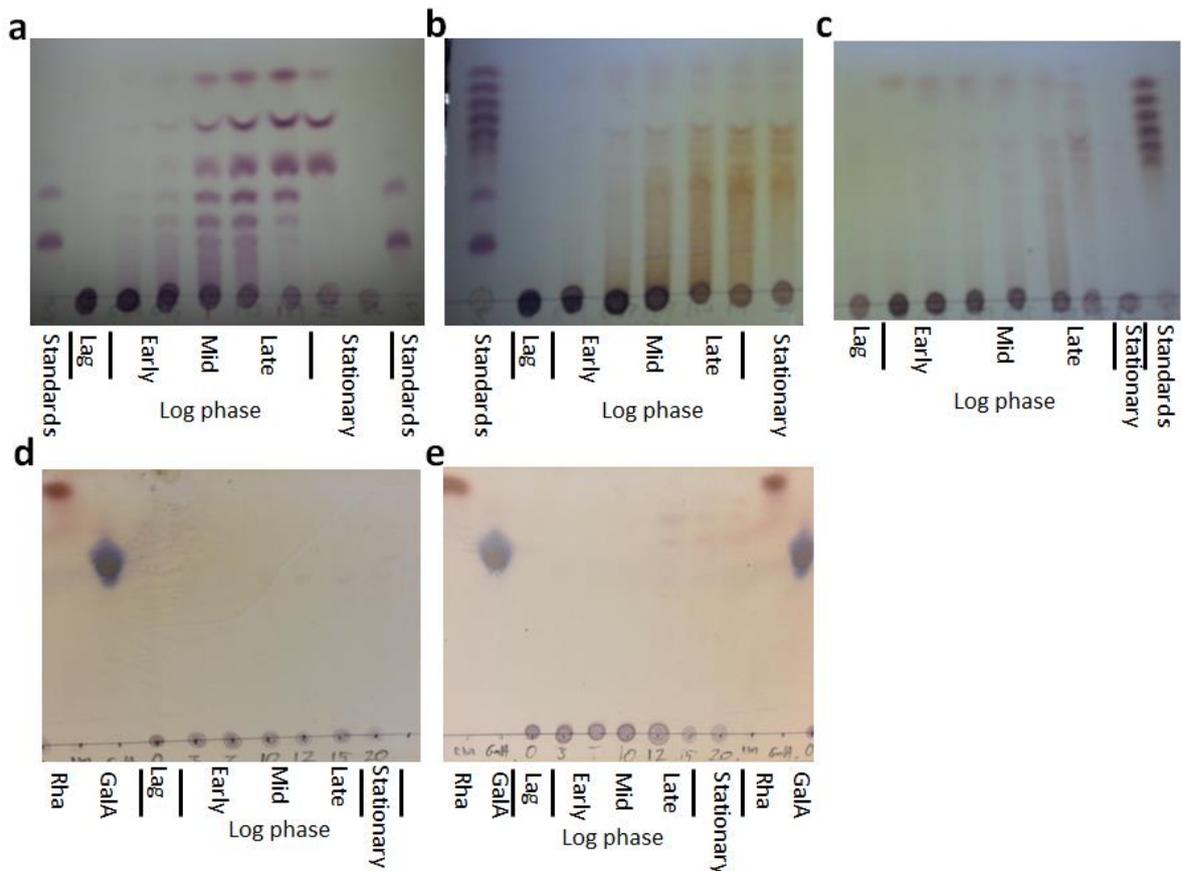


Figure 4.25 TLC of oligosaccharides released by *B. thetaiotaomicron* during growth on pectins. Samples taken over the course of a growth curve were spotted on TLC and run alongside known oligosaccharide and monosaccharide standards. *B. thetaiotaomicron* was grown at 37 °C under anaerobic conditions in MM + 0.5 % galactan (a), branched arabinan (b), unbranched arabinan (c) mucilage corresponds to undecorated RGI backbone (d) RGII (e).

To assess the potential for nutrient release and cross-feeding during *B. thetaiotaomicron* utilisation of pectic polysaccharides, wild type *B. thetaiotaomicron* was grown in MM+ 0.5% galactan, arabinan, Mucilage (Mucilage is undecorated RGI from Arabidopsis), and RGII, in separate cultures. At time points that represented different phases of growth, as determined by OD of the growing culture,

samples of the culture supernatant were taken and evaluated for glycan content by TLC (Figure 4.25). Common to growth on all glycans tested, only the polysaccharide was present in lag phase (Figure 4.25), *B. thetaiotaomicron* is not at a sufficient density to begin significant degradation of the glycans. At early log phase faint smears of long oligosaccharide products were present in the culture supernatant of the galactan and branched arabinan growths (Figure 4.25a, b), while the unbranched arabinan growth supernatant appeared to be devoid of oligosaccharides (Figure 4.25c). During mid log phase discreet oligosaccharide bands were present in galactan and branched arabinan supernatant (Figure 4.25a, b), while the smear of long undefined oligosaccharides grew more intense in the unbranched arabinan culture (Figure 4.25b). In the galactan growth longer oligosaccharides were beginning to disappear at late log/early stationary phase (Figure 4.25a), likely being imported by *B. thetaiotaomicron* or hydrolysed by the surface GH53 endo-galactanase, generating shorter oligosaccharides, which continues into early stationary phase where only galactose, Gal2 and Gal3 were present (Figure 4.25a). The disappearance of longer oligosaccharides was not observed until stationary phase for either arabinan culture (Figure 4.25b, c).

During growth in MM + 0.5% Mucilage the origin spot of polysaccharide became fainter over the course of the growth, however, there were no intense bands to suggest release of oligosaccharides during glycan utilisation (Figure 4.25d). There was a very faint band present which could not be identified (Figure 4.25d), which migrated slower than GalA on TLC suggesting the oligosaccharide present is larger than the monosaccharide. The unidentified oligosaccharide did not grow fainter in the later stages of the growth suggesting it was not used by *B. thetaiotaomicron*. Similar observations were made for the growth on MM+ 0.5% RGII in which presumed monosaccharides or disaccharides accumulated towards the end of the growth. In this regard the RGII-derived monosaccharides apiose, Kdo, 2-O-Me-fucose and aceric acid, and the disaccharide 2-O-Me-xylose- α 1,3-fucose, were not used by *B. thetaiotaomicron* (Ndeh unpublished data), and thus would accumulate in the culture supernatant. Cell supernatant was not run on HPAEC-PAD as cell debris can cause damage to the highly sensitive detection electrode of the HPAEC machine.

4.2.5.2 Mutant growth on oligosaccharides and polysaccharides

Mutants of surface endo-active enzymes encoded by each of the PULs orchestrating the degradation of the pectic polysaccharides galactan, arabinan, RGI and RGII were generated by either the introduction of mutations that inactivated the cognate enzyme or deletion of the entire gene were made of BT_4668 (β 1,4-galactanase) in the galactan PUL (Δbt_4668), BT_0360 and BT_0367 (α 1,5-arabanases, (Cartmell *et al.*, 2011) in the arabinan PUL ($\Delta bt_0360\Delta bt_0367$), BT_4170 (pectic lyase, Luis unpublished data) in the RGI PUL (Δbt_4170 , generated by Ana Luis), BT_1023 (pectic lyase, Didier Ndeh personal communication) in the RGII PUL (Δbt_1023 , generated by Didier Ndeh) and deletion of the entire RGII PUL ($\Delta RGIIPUL$, generated by Wade Abbott). These mutations were then tested for growth during culturing on MM + 0.5% of the relevant polysaccharide and MM+0.5 % of oligosaccharides generated from the relevant substrate by digestion with the mutated enzyme(s). While Δbt_4668 , $\Delta bt_0360/\Delta bt_0367$ and, Δbt_1023 and $\Delta RGIIPUL$ demonstrated no growth on galactan, arabinan and RGII (Figure 4.26), respectively. The mutant Δbt_4170 showed good growth on potato RGI despite lacking BT_4170 activity (Figure 4.26). The potato RGI used here has been shown to contain galactan side chains of sufficient length to be cleaved by BT_4668 and transcription data showed the galactan PUL is activated by the polysaccharide (Martens *et al.* 2011). A double mutant was made of the entire galactan PUL and BT_4170, $\Delta bt_4170\Delta GALPUL$. This double mutant was unable to utilise galactan, however still grew on RGI. Complete acid hydrolysis of RGI was performed and run on HPLC to determine if there were contaminants in the RGI which may explain the growth observed by $\Delta bt_4170\Delta GALPUL$. As RGI could not be easily purified from the contaminants another source of RGI, Mucilage, was tested for growth (Figure 4.26c). As stated above Mucilage is the RGI backbone without any decorations. When Δbt_4170 was cultured on mucilage, despite high background absorbance from the polysaccharide, the mutant showed no growth (Figure 4.26c), although the variant was able to utilise Mucilage derived oligosaccharides (Figure 4.26c) generated by incubation of the linear RGI (10 mg/ml) with 1 μ M BT_4170 for 3 h.

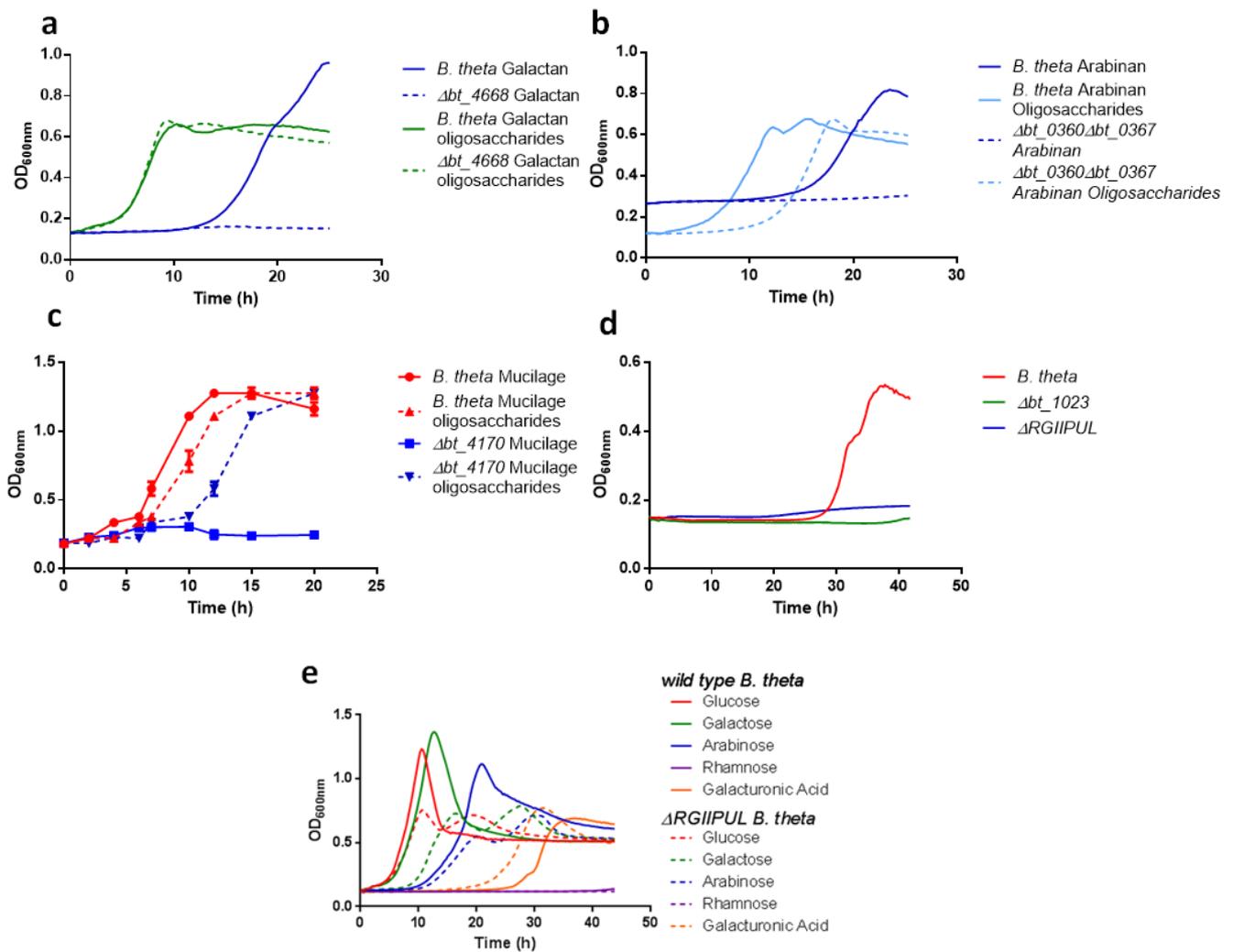


Figure 4.26 *B. theta* mutants lacking surface endo-acting enzymes encoded by pectin PULs grown on oligosaccharides and polysaccharides. Δbt_{4668} grown in MM+0.5 % galactan and MM+0.5 % galactooligosaccharide (a). $\Delta bt_{0360}\Delta bt_{0367}$ grown in MM+ 0.5 % arabinan and MM+0.5 % arabinooligosaccharide (b). Δbt_{4170} grown in MM+ 0.5 % mucilage and MM+0.5 % mucilage-oligosaccharide in 5 ml tube cultures (c). Comparison of Δbt_{1023} $\Delta RGII PUL$ and wild type *B. theta* grown in MM + 0.5 % RGII (d). $\Delta RGII PUL$ and wild type *B. theta* grown in MM + 0.5 % galactose, arabinose, rhamnose and galacturonic acid (e). Growths a, b, d, e were monitored by automatic plate reader at A_{600nm}, whereas c was performed in 5 ml tube cultures with manual absorbance readings at A_{600nm}, all in anaerobic conditions at 37 °C. Growths were performed in triplicate.

In co-culture, to differentiate between wild type *B. theta* and the mutant, a unique nucleotide sequence was introduced into each variant and the wild type at the *att1* site in the genome. The two signature-tag sequences, referred to as tag11 and tag1, are on the pNBU2-tetQb vector developed by the Martens/Gordon labs (Martens *et al.*, 2008). NBU2 is a 2.5 kbp mobilisation

region which integrates into one of two *att* sites in the *B. thetaiotaomicron* genome carrying the signature-tags. The *att* integration sites are 13 bp sequences at the 3' end of two Ser-tRNA genes, Ser-tRNA_{UGA} and Ser-tRNA_{UGA2} (Wang *et al.*, 2000). Wild type *B. thetaiotaomicron* harboured the *tag1* sequence while each mutant contained a copy of the *tag11* sequence. The proportion of wild type to mutant *B. thetaiotaomicron* was determined by performing qPCR using primers specific to each tag on gDNA preparations of culture samples taken at intervals during growth on the target glycan. The ratio was used, in conjunction with viable counts taken at the same time points, to give an accurate cell density of both wild type and mutant *B. thetaiotaomicron* cells in the co-culture. Each mutant was tested for presence of tag1, tag11 and an intact *att1* site. The observed lack of tag1 in the mutants ensure that any tag1 products detected are from the wild type *B. thetaiotaomicron*, while tag11 products (200 bp) confirm successful insertion of the qPCR tag. Presence of PCR product (2500 bp) using the ATT-diagnostic primers (Chapter 2.9.4, Appendix A.3) demonstrates the tag inserted in *att2* site as *att1* remains intact (Figure 4.27a).

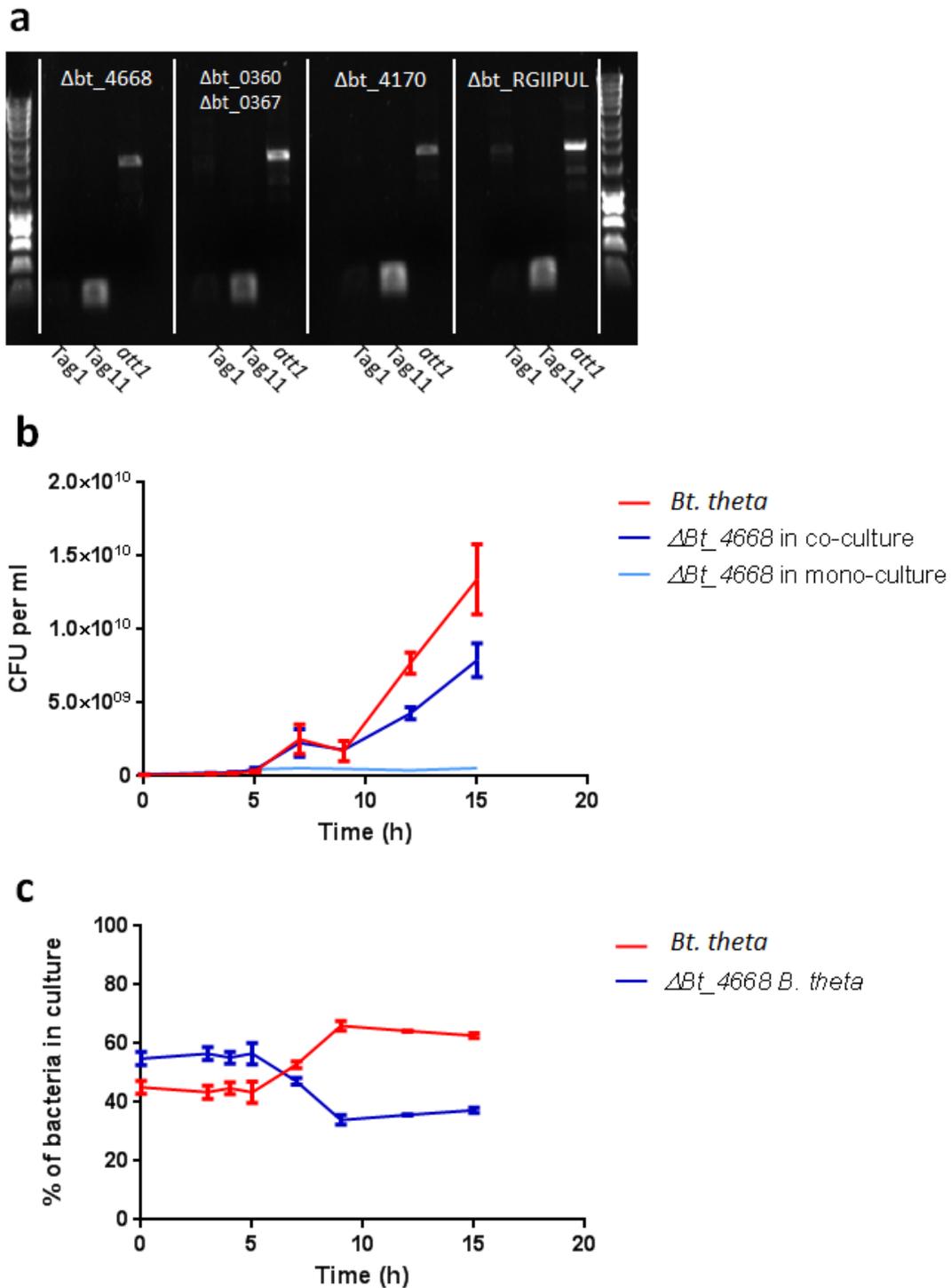


Figure 4.27 Galactan cross-feeding experiment. Agarose gel electrophoresis (0.8% gel) of qPCR tags from each of the mutants used in the cross-feeding experiments, PCR products of Tag1, Tag11 and *att1* integration site with known standards (a). Wild type *B. theta* (red) and Δbt_{4668} (blue) co-cultured in MM+0.5% galactan at 37 °C under anaerobic conditions. CFU/ml of wild type *B. theta* in co-culture, Δbt_{4668} in co-culture and in mono-culture (b). Ratio of wild type and Δbt_{4668} *B. theta* in the co-culture on galactan (c). *B. theta* refers to wild type *B. theta*.

4.2.5.3 Galactan Cross-Feeding

The *B. thetaiotaomicron* GH53 mutant, Δbt_4668 , was mixed in a 1:1 ratio with wild type *B. thetaiotaomicron* and used to inoculate MM+ 0.5% galactan. Initially the mutant and wild type *B. thetaiotaomicron* maintained the inoculation ratio of 55:45, favouring the mutant (Figure 4.27b). The ratio shifted at 7 h when the culture was at the beginning of log phase, here the wild type *B. thetaiotaomicron* cell numbers rose to become the majority of the culture at 10 h, which was maintained until the end of the growth (Figure 4.27b). The mutant Δbt_4668 was able to maintain a 30-40 % share of the culture (Figure 4.27b), most likely due to the high concentration of galactooligosaccharides released by wild type *B. thetaiotaomicron* during galactan utilisation. Final cell density of the mutant in co-culture with the wild type was 24-fold greater than that of the mutant alone on galactan (Figure 4.27a).

4.2.5.4 Arabinan Cross-feeding

A 1:1 ratio of wild type and $\Delta bt_0360\Delta bt_0367$ *B. thetaiotaomicron* was attempted by OD of initial starter cultures prior to mixing and inoculation. However, often small variation in OD translate to large differences in cell numbers, as demonstrated by the inoculation ratio observed for the arabinan co-cultures (Figure 4.28b, d). During co-culture on branched and unbranched arabinan $\Delta bt_0360\Delta bt_0367$ remains in the minority, dropping to below 2% and 10%, respectively (Figure 4.28b, d). Despite this, in co-culture the mutant achieves a cell density 4-fold and 8-fold greater than when grown alone on branched and unbranched arabinan, respectively (Figure 4.28a, c). Despite high concentrations of oligosaccharides produced by *B. thetaiotaomicron* during growth $\Delta bt_0360\Delta bt_0367$ cannot grow to the same level observed on other glycans during co-culture. This may be due to restrictions on size of oligosaccharide the SusCD-pairs are capable of transporting causing a build up of long oligosaccharides which require degradation by the slow acting surface GH43 enzymes at the surface of the wild type *B. thetaiotaomicron*.

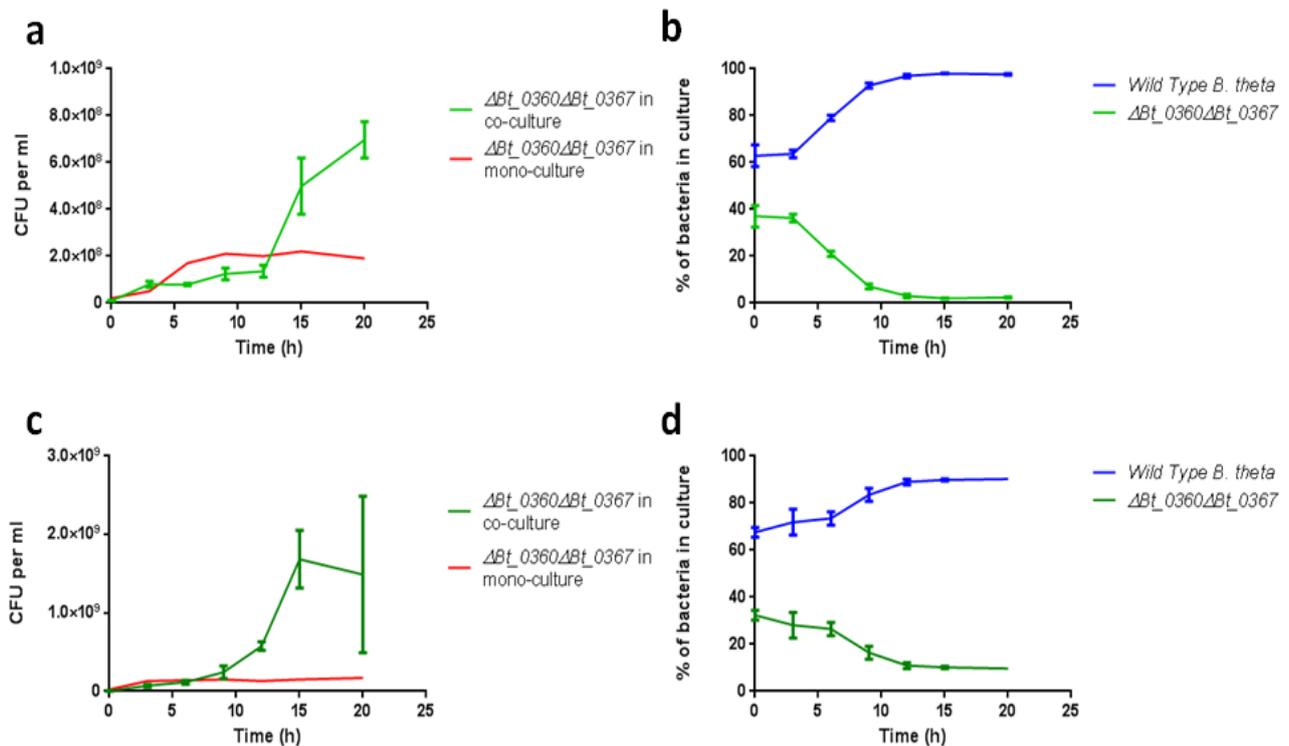


Figure 4.28 Arabinan cross-feeding experiments. Wildtype *B. theta* and $\Delta BT_0360\Delta BT_0367$ co-cultured in MM + 0.5 % arabinan at 37 °C under anaerobic conditions. CFU/ml of wildtype *B. theta* in co-culture, $\Delta bt_0360\Delta bt_0367$ in co-culture on branched (a) and unbranched (b) arabinan. Ratio of wildtype and $\Delta bt_0360\Delta bt_0367$ *B. theta* in the co-culture on branched (c) and unbranched (d) arabinan.

4.2.4.5 Mucilage Cross-feeding

Counterintuitively, growth of wild type *B. theta* on mucilage supports good growth of Δbt_4170 , much more than analysis of the supernatant oligosaccharide content would suggest (Figure 4.25d). The mutant, Δbt_4170 achieved a 40% final proportion of the culture, which equates to less than 75-fold greater viable count than the Δbt_4170 mono-culture on mucilage (Figure 4.29a, b). This may reflect the capacity of the outer membrane RGI porin to transport larger oligosaccharides, and such molecules are not visible by TLC.

4.2.4.6 RGII Cross-feeding

Two mutants were used as recipients in co-growths on RGII, the Δbt_1023 mutant (encodes the PL1 pectate lyase) and $\Delta RGIIPUL$, which has the entire RGII PUL deleted from the genome. In co-culture

with wild type *B. thetaiotaomicron* on RGII, Δ RGII*PUL* was unable to grow showing no significant increase when the mutant was grown in mono-culture on the pectic polysaccharide. The mutant lacking only the surface PL1 pectate lyase fared much better on RGII in co-culture with wild type *B. thetaiotaomicron* (Figure 4.29). The Δ *bt_1023* strain was able to maintain 30-40 % of the co-culture despite being unable to generate small oligosaccharides from RGII. This translates to a 55-fold greater number of cells of Δ *bt_1023* in the co-culture over the mono-culture of the same mutant on MM+0.5% RGII (Figure 4.29c-f). Cross-feeding occurs between wildtype and mutant *B. thetaiotaomicron* on RGII occurs when the binding and transport apparatus is intact as in the Δ *bt_1023* mutant.

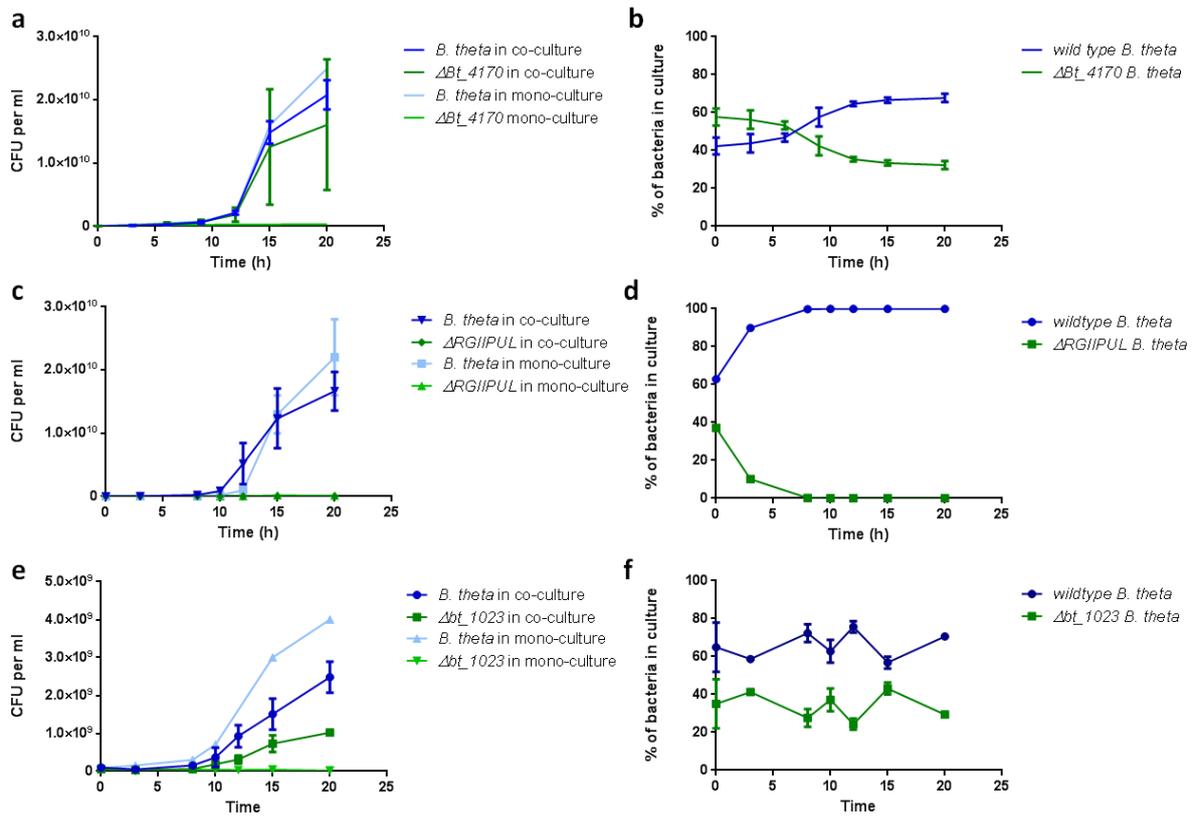


Figure 4.29 Mucilage and RGII cross-feeding experiments. Wild type *B. thetaiotaomicron* and Δ *bt_4170* co-cultured in MM+0.5% mucilage (a), Wildtype *B. thetaiotaomicron* and Δ RGII*PUL* (c), Wildtype *B. thetaiotaomicron* and Δ *bt_1023* (e) in MM+0.5% RGII at 37 °C under anaerobic conditions. Panels a, c and e include monoculture and co-culture of the wild type *B. thetaiotaomicron* and the relevant mutant. Ratio of wild type and Δ *bt_4170* *B. thetaiotaomicron* in the co-culture on mucilage (b). Ratio of wild type and Δ RGII*PUL* *B. thetaiotaomicron* in the co-culture on RGII (d), and ratio of wild type and Δ *bt_1023* *B. thetaiotaomicron* in the co-culture on mucilage (f). *B. theta* refers to wild type *B. thetaiotaomicron*.

4.3 Discussion

4.3.1 Utilisation of galactan

In the context of the gut, the HGM is flooded with a mixture of polysaccharides from plant cell walls, which are major components of the human diet (Backhed *et al.*, 2005). Here, RGI consists of a rhamnogalaturonic acid backbone decorated with long chains of galactan, and branched/unbranched arabinan (Mohnen, 2008). When *B. thetaiotaomicron* comes into contact with RGI it can liberate the arabinan and galactan side chains by the action of surface endo-acting enzymes, BT_4668 (galactan), BT_0360 (arabinan) and BT_0367 (arabinan). The data presented above describe each component of the galactan utilisation system in isolation. These data in conjunction with other publications can be used to create a model for the degradation and import of galactan by *B. thetaiotaomicron* (summarised in Figure 4.30). As with all PULs, the galactan PUL is constitutively expressed to survey the environment for potential polysaccharide nutrients (Martens *et al.*, 2011). *B. thetaiotaomicron* encounters galactan at the external face of the outer membrane where it is bound by BT_4669 a lipoprotein SGBP. BT_4669 shows preference for the polysaccharide over oligosaccharides (Table 4.3), suggesting the presence of an extended binding site. Binding of galactan to the cell surface increases local concentrations of the glycan, enhances the proximity of the endo-galactanase (BT_4668) to its substrate leading to increased catalytic efficiency. Deletion of BT_4669 demonstrates the importance of the SGBP in galactan utilisation. Binding of galactan may open up the structure, uncoiling the helical conformation and allowing BT_4668 to cleave the glycan backbone. It should be emphasised, however, that the endo-galactanase (BT_4668) displays typical activity for GH53 enzymes and thus appears to be functional in the absence of accessory carbohydrate binding proteins. Another possibility is BT_4669 fulfils an as yet unclear structural role in stabilising the surface apparatus involved in binding and deconstructing galactan, without which surface activity does not occur.

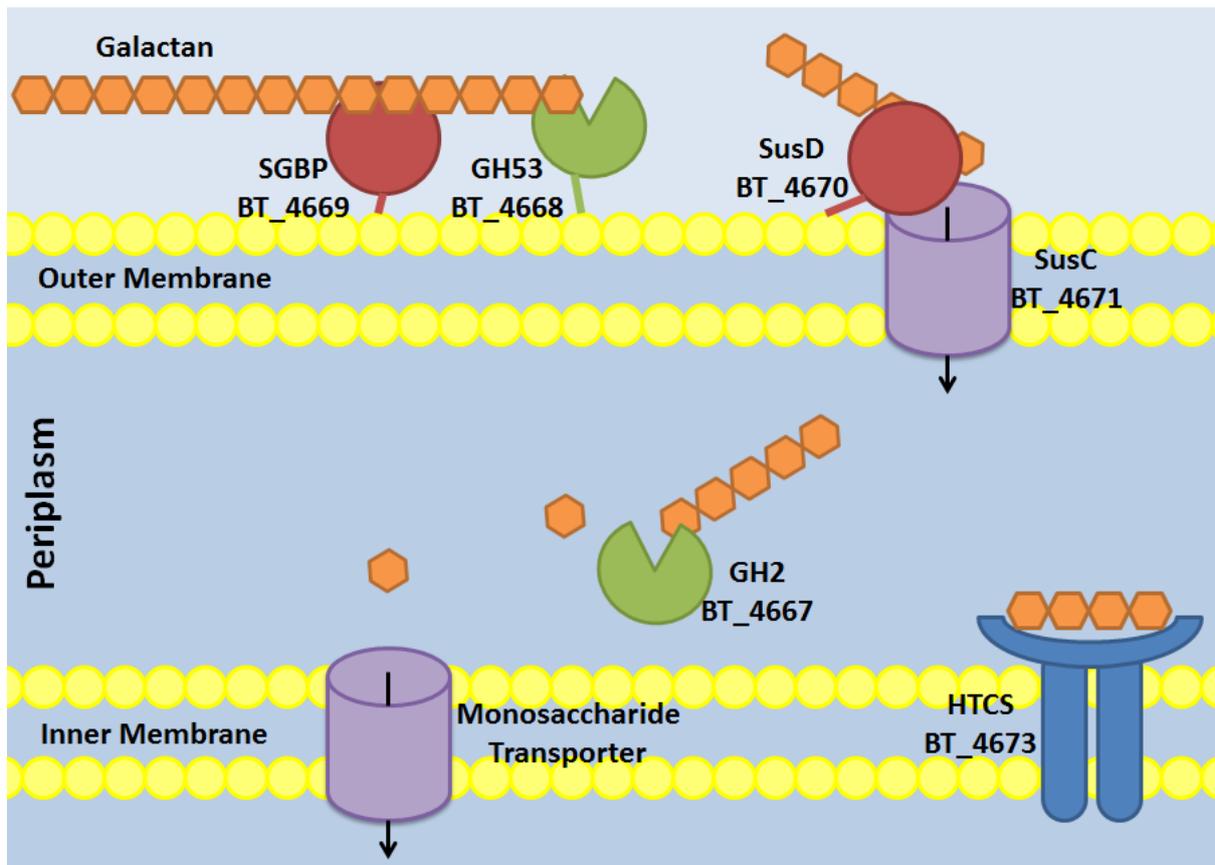


Figure 4.30 *B. thetaiotaomicron* galactan utilisation. Diagram summarising the breakdown and utilisation of galactan at the cell surface and in the periplasm of *B. thetaiotaomicron*.

The surface lipoprotein BT_4668, is an endo- β 1,4-galactanase generating long galactooligosaccharides from galactan polysaccharide (Figure 4.5b). The oligosaccharide products from BT_4668 digestion of galactan at the cell surface are bound by BT_4670 (Figure 4.10, Table 4.3), the galactan PUL SusD-homologue and channelled through BT_4671, the galactan PUL SusC-homologue, into the periplasm. The small difference in affinity for Gal6 and galactan suggests the ligand binding site is able to accommodate around six galactosyl residues (Table 4.3). Galactan forms a helical structure with each turn approximately six residues long (Cid *et al.*, 2010). Ligand specificity of BT_4670 may be based on recognition of the structure and chemical composition, similar to the starch utilisation system where SusD binds more tightly to cyclic than linear oligosaccharides of the same length (Koropatkin *et al.*, 2008). This process is not efficient, leading to oligosaccharide release

at the cell surface (Figure 4.25a). The inefficiency derives from an imbalance between the rate at which oligosaccharides are generated by BT_4668 (Table 4.2), and the rate at which the SusCD pair transport these saccharides. Saturation of the transport system does not appear to cause any detrimental effects on *B. thetaiotaomicron* growth on galactan *in vitro*, perhaps transporter expression is limited to allow for maximal growth without flooding the periplasm with undigested oligosaccharides which may cause unwanted interactions at high concentrations in the compartment. The action of BT_4668 is vital to galactan utilisation as very large oligosaccharides cannot be transported into the periplasm by the SusCD-homologue pair. Once in the periplasm oligosaccharides are sequestered away from any potential competitors, here the oligosaccharides are degraded to their monosaccharide constituents by BT_4667 (Figure 4.11), an exo-acting β 1,4-galactosidase. Interestingly BT_4667 is much more active against Gal2 than lactose (Table 4.4), indicating a preference for galactose over glucose at the +1 subsite; lactose comprises Gal- β 1,4-Glc. This preference for homogenous galactooligosaccharide substrates indicates a specialisation consistent with its location in the galactan degrading apparatus of *B. thetaiotaomicron*; the enzyme encounters galactooligosaccharides rather than lactose in the periplasm. The intermediate products of galactan utilisation can then be bound by the ligand binding domain of the HTCS sensor, BT_4673 (Figure 4.12, Table 4.3). Binding affinity data of BT_4673 shows a slight preference for Gal4 over Gal5 (Table 4.3). Thus, mid-range oligosaccharides are the activating ligand for the galactan PUL. This specificity protects *B. thetaiotaomicron* from inappropriate PUL upregulation from ligands which may contain galactose but are not derived from the polysaccharide. The galactooligosaccharides are completely degraded to mono- or di- saccharides in the periplasm and transported into the cytoplasm for entry into fermentation and energy generation pathways.

4.3.1.1 BACOVA_05493

BACOVA_05493, the glycoside hydrolase found in the *B. ovatus* galactan PUL, but not in the corresponding *B. thetaiotaomicron* locus, appears to be a β 1,4-galactosidase and shows significant

homology to GH2 family enzymes (Figure 4.13). BACOVA_05493 most likely belongs to a GH2 subfamily. BACOVA_05493 displays slight preference for the longer oligosaccharides tested and requires a minimum substrate length of 3 DP for activity (Figure 4.14a, Table 4.5). A preference for the longer galactooligosaccharides may indicate a divergence in the galactan utilisation systems of *B. thetaiotaomicron* and *Bacteroides* spp. possessing a BACOVA_05493 homologue. Those with a BACOVA_05493 homologue may import longer galactooligosaccharides than *B. thetaiotaomicron* requiring another β 1,4-galactosidase to efficiently degrade these extended substrates in the periplasm. Interestingly, when BACOVA_05493 is inactivated the Δ *bacova_05493* *B. ovatus* mutant is unable to utilise galactan (Figure 4.15b), despite possessing the same machinery encoded by the *B. thetaiotaomicron* galactan PUL. A possible explanation is that the HTCS, which activates the galactan PUL in *B. ovatus*, requires products generated by BACOVA_05493 to upregulate the PUL. An alternative explanation is that the GH53 and/or GH2 present in the *B. ovatus* galactan PULs have divergent activities from their counterparts in the *B. thetaiotaomicron* galactan PUL, hence making the presence of BACOVA_05493 a requirement for utilisation. However this seems unlikely as the *B. ovatus* homologues to BT_4667 and BT_4668 are practically identical to the corresponding enzymes in the *B. thetaiotaomicron* galactan PUL based on sequence alignment (data not shown). It is possible that the GH2 and/or GH53 enzyme is poorly expressed in species containing the additional galactosidase, or the outer membrane importer is specific to long galactooligosaccharides.

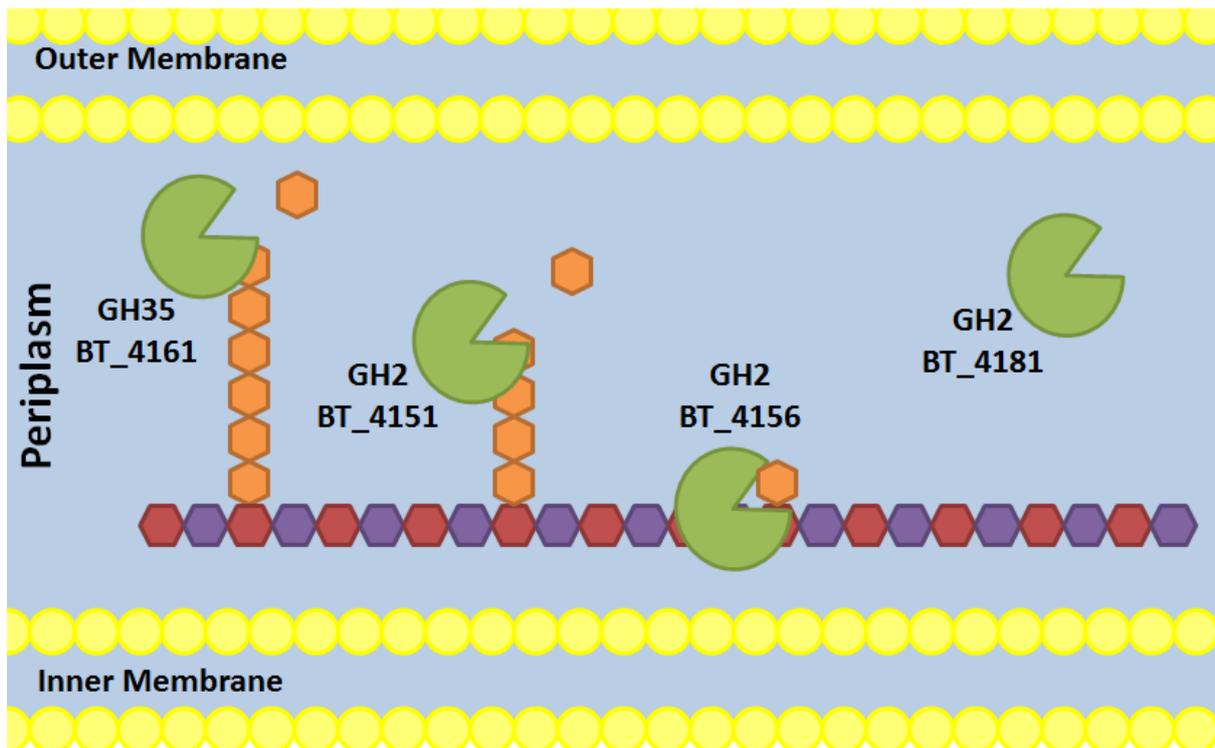


Figure 4.31 RGI galactooligosaccharide sidechain breakdown. Diagram summarising deconstruction of galactan sidechains of RGI by BT_4151, BT_4156, BT_4181 and BT_4160 in the periplasm of *B. thetaiotaomicron*.

4.3.2 Degradation of Galactooligosaccharide side chains on RGI

When galactan and arabinan are removed from the RGI backbone short chains of galactosyl residues remain. The surface polysaccharide lyase family 9 (PL9) enzyme, BT4170, cleaves the RGI backbone into large oligosaccharides that contain the galactose side chains (Ana Luis personal communication). These branched oligosaccharides are then imported into the periplasm to complete degradation without further loss of products. In the periplasm four β 1,4-galactosidases, one belonging to GH35 (BT4160) and three to GH2 (BT4151, BT4156 and BT4181), hydrolyse the galactosyl side chains that decorate the RGI oligosaccharides, releasing galactose as the product. This model is based on the assumption that removal of the galactan side chains prior to the action of BT_4170 is inefficient as this would negate the requirement for these enzymes derived from the RGI PUL. It is also possible that a proportion of the RGI backbone imported by *B. thetaiotaomicron* is derived from the pectin degrading systems of other bacteria in the human gut. These molecules

contain remnants of the galactan backbone that are too short to be released by the surface endo-galactanase BT_4668, thus the galactan apparatus would not be activated. BT_4160 demonstrates good activity on all free galactooligosaccharides tested, galactan and some activity on galactooligosaccharide side chains on RGI (Figure 4.16, Table 4.6). The GH35 enzyme does not require the galactooligosaccharide to be in the context of RGI for activity. Low activity on RGI suggest BT_4160 is unable to trim the shorter oligosaccharides still appended to the RGI backbone after removal of the galactan by BT_4668 (Figure 4.17c). The Presence of the backbone may cause steric hindrance when the oligosaccharide side chains are relatively short, reducing BT_4160 activity on RGI. Of the four β 1,4-galactosidases BT_4151 showed greatest activity on RGI targeted galactooligosaccharides with a DP of 4 to 7 and galactan (Figure 4.16). BT_4151 shows low activity on shorter oligosaccharides leaving a ladder of products similar to that of endo-acting enzymes (Figure 4.16). Accumulation of galactose, however, confirms the exo-activity of the enzyme. BT_4156 is inactive on free galactooligosaccharide with the exception of Gal2, but releases galactose from RGI (Figure 4.16), implying activity on only the shortest oligosaccharides and tolerance, if not preference, for the RGI backbone bound in the positive subsites of the enzyme. This confirms the preferred substrate of BT_4156 to be the Rha-Gal bound at the base of the galactooligosaccharide side chain once it is degraded to a single galactosyl decoration by the action of BT_4151 and BT_4160. Of the three active β 1,4-galactosidases, BT_4156 has the most impact on RGI utilisation, without the ability to remove the galactose substitution on the rhamnose residues of the RGI backbone the action of the exo-acting GH28 and GH106 enzymes, which successively remove rhamnose and GalA from the backbone, are blocked hindering utilisation of large portions of the RGI backbone. The importance of BT_4156 is reflected in the mutant Δbt_4156 , which has a growth curve with a much longer lag phase than wild type *B. thetaiotaomicron*, although there is no change to the maximum cell density in MM + 0.5% RGI (Figure 4.18b). The final GH2 present in the RGI PUL, BT_4181 lacks the appropriate catalytic residues for activity (Figure 4.19a) and thus, not surprisingly, was inactive on all substrates tested (Figure 4.16). This inactive enzyme, despite harbouring no obvious mutations to

the substrate binding site, was unable to bind any RGI derived glycan, and so its role in RGI utilisation is unclear (Figure 4.19b). This data had been compiled to create Figure 4.31, in which the roles of each of the four enzymes is described along with its position in the breakdown of RGI galactooligosaccharide sidechains.

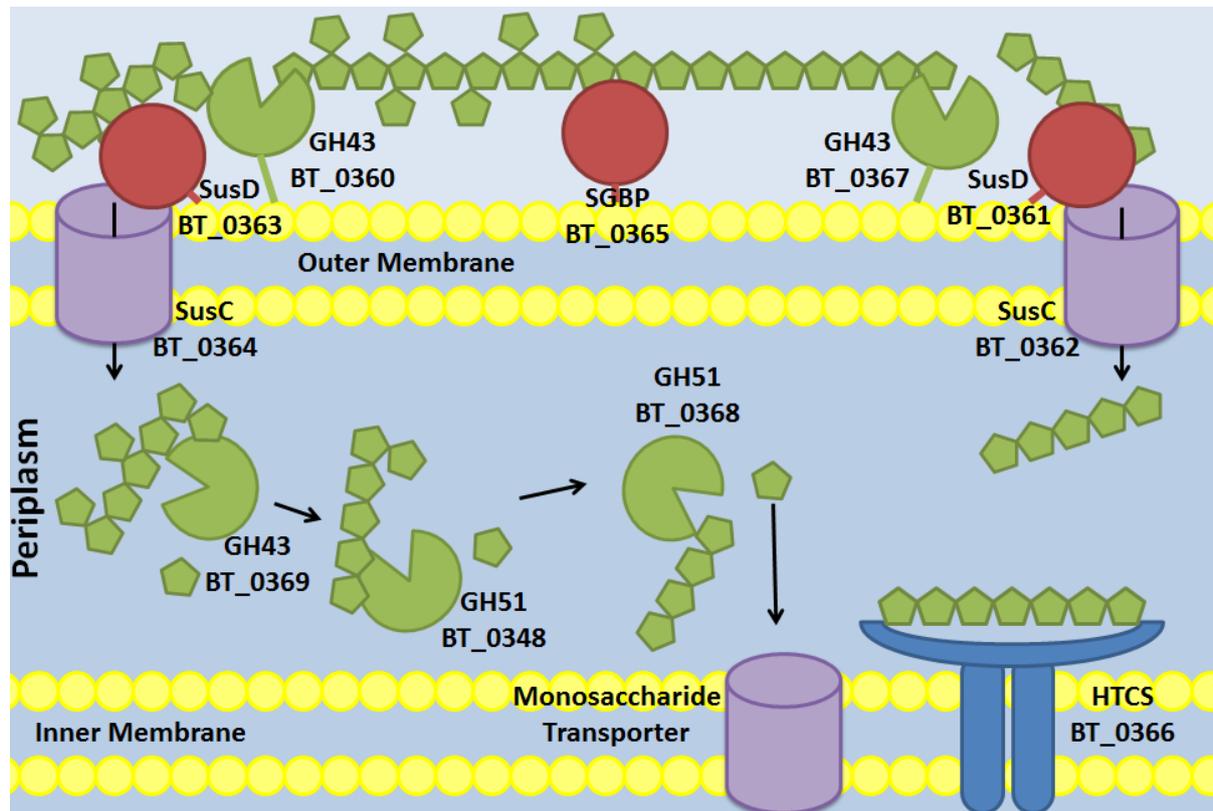


Figure 4.32 Arabinan utilisation. Diagram summarising the breakdown and utilisation of branched and unbranched arabinan at the cell surface and in the periplasm of *B. thetaiotaomicron*.

4.3.3 Arabinan Utilisation

Arabinan degradation and utilisation differs from the simple galactan utilisation system. Arabinan is present with or without arabinose decorations to the backbone (Mohnen, 2008), unlike linear galactan, hence requires a greater number of degradative enzymes to complete hydrolysis of each unique linkage present. Once liberated from RGI arabinan, regardless of the presence of arabinose decorations, can be bound by the SGBP, BT_0365 (Table 4.7). Bound linear and branched arabinan is degraded by the GH43 endo- α 1,5-arabinanases BT0360 and BT0367 (Cartmell *et al.*, 2011). The

activity of these previously characterised enzymes, however, was low (Cartmell *et al.*, 2011). Whole cell assays and oligosaccharides present in the growth supernatant is consistent with the activity found with recombinant surface enzymes, demonstrating slow oligosaccharide release and low activity at the cell surface (Figure 4.23d, 4.25b, c). Double mutation of the genes encoding the two arabinanases, $\Delta bt_0360/\Delta bt_0367$, confirmed surface degradation of arabinan is required for transport of the target glycan into the periplasm (Figure 4.24). Single mutations of each surface GH43 enzyme showed that BT_0367 has a greater contribution to the utilisation of arabinan (Figure 4.24). These data are surprising as sugar beat arabinan is heavily branched and is thus a better substrate for BT_0360 which showed preference for branched arabinan.

B. thetaiotaomicron expresses two SusCD-homologue pairs for utilisation of arabinan (Figure 4.2b). Unfortunately the true ligand of the SusD-homologues, BT_0361 and BT_0363, could not be defined (Figure 4.22). Deletion of each SusC-homologue showed BT_0364 had a greater contribution to arabinan utilisation than BT_0362. The Δbt_0364 mutation had a similar effect on linear and branched arabinan utilisation, suggesting the transporter does not distinguish between the two forms of the polysaccharide (Figure 4.24). It is entirely possible that BT_0361, and by extension BT_0362, has a greater contribution to transport when *B. thetaiotaomicron* encounters different arabinan structures that have not been explored here. Once in the periplasm, the oligosaccharides enter one of two pathways depending on the presence of decorations. Rare arabinose side chains are removed by BT_0369, a α 1,2-arabinofuranosidase that targets the O2 arabinosyl linkage in the context of single or double substitutions (Cartmell *et al.*, 2011). With the release of all O2 linked arabinose, BT_0348, an α 1,3-arabinofuranosidase hydrolyses the O3 linkages of the remaining arabinose substitutions (Figure 4.23.b, Table 4.8). Activity on linear oligosaccharides suggests BT_0348 has side activity on α 1,5-arabinofuranosidase. Once the substitutions are removed BT_0368, an exo- α 1,5-arabinofuranosidase, degrades the linear or unbranched oligosaccharides releasing arabinose which (Figure 4.23.a, Table 4.8), like galactose, is transported into the cytoplasm where it enters fermentation pathways resulting in energy generation (Turner and Roberton, 1979).

BT_0366, the HTCS, binds Ara7 in the periplasm to upregulate the arabinan PUL (Martens *et al.*, 2011). A long oligosaccharide activating ligand suggests large oligosaccharides must be imported by the two SusCD-homologue pairs as the activating ligand of previously characterised PULs tend to be an intermediate product of periplasmic glycan degradation. In this respect it is interesting that the α 1,3-arabinofuranosidase is substantially less active than the α 1,5-arabinofuranosidase. It is possible that the slow removal of the side chains blocks the action of the α 1,5-arabinofuranosidase enabling the large arabinooligosaccharides to be present throughout the growth cycle ensuring that the arabinan PUL is not prematurely switched off.

4.3.4 Cross-feeding with pectic polysaccharides

The cross-feeding experiments performed here involve the use of mutants lacking a surface endo-acting enzyme, which renders the mutant reliant on exogenously generated oligosaccharides. Using these mutants creates an artificial situation. However, these experiments were designed to evaluate the capacity for oligosaccharide cross-feeding using a strategy that only evaluates whether oligosaccharides are released into the environment and are available to other organisms. It was for this reason we deployed a mutant of *B. thetaiotaomicron* lacking the critical surface enzyme but retaining the rest of the degrading apparatus. Thus the mutants retained their oligosaccharide import and intracellular (periplasm and cytoplasm) utilization systems.

4.3.4.1 Galactooligosaccharide Cross-Feeding

Galactan, the least complex of the polysaccharides studied here, was digested at the cell surface of *B. thetaiotaomicron* releasing high concentrations of oligosaccharides (Figure 4.25a) that, in the natural environment of the bacterium, would be highly accessible to competitors within the same niche. Galactooligosaccharide utilisation is well represented in the human gut microbiota (Scott *et al.*, 2014). Maybe reflecting the simplicity of the degradative system, only requiring a β 1,4-galactosidase and transporter/binding system. Galactooligosaccharides are given as prebiotic treatment to boost *Bifidobacterium* and other species which have been found to promote good gut

health (Weaver *et al.*, 2011; Whisner *et al.*, 2013; Furuse *et al.*, 2014). The concentration of galactooligosaccharides released by *B. thetaiotaomicron* during growth on galactan suggests this is a low priority growth substrate for the bacterium, given that galactan can be utilised by many species in the gut and *B. thetaiotaomicron* is able to utilise a wide range of alternative glycans (Martens *et al.*, 2011; Larsbrink *et al.*, 2014; Cuskin *et al.*, 2015). The oligosaccharides are produced at the cell surface by the β 1,4-galactanase, BT_4668 (Figure 4.8). If a more selfish approach to galactan utilisation was required *B. thetaiotaomicron* may have, at some point in its evolutionary history, selected for a slower surface enzyme, similar to the *B. ovatus* xylan surface GH10 (Zhang *et al.*, 2014a). This has not occurred suggesting there is some benefit to releasing galactooligosaccharides. A prebiotic has been previously explored through the digestion of galactan with GH53 enzymes from *Emericella nidulans* to administer pre-digested galactooligosaccharides (Michalak *et al.*, 2012). Supplementing infant formula with various length galactooligosaccharides have been explored, demonstrating enrichment for gut microbes similar to the natural breast-fed microbiota (Veereman-Wauters *et al.*, 2011). In adults, galactooligosaccharide supplements have been linked to a reduction in travellers' diarrhoea and alleviation of irritable bowel syndrome (Drakoularakou *et al.*, 2009; Silk *et al.*, 2009). Perhaps, similar to administering galactooligosaccharides in the diet, *B. thetaiotaomicron* releases galactooligosaccharides promoting gut health, which benefits *B. thetaiotaomicron* indirectly. This treatment strategy may allow slow release of galactooligosaccharides rather than flooding the gut with simple oligosaccharides, hence prolonging the positive effects attributed to the treatment.

4.3.4.2 Arabinan Cross-feeding

Unlike the galactan cross-feeding experiment, there was only a modest difference between the mutant lacking the surface GH43 enzymes, $\Delta bt_0360/\Delta bt_0367$, in mono-culture and in co-culture with wild type *B. thetaiotaomicron* on MM + 0.5% branched and unbranched arabinan (Figure 4.28). This may stem from a combination of the low activity of each GH43 enzyme on their preferred

substrate (Cartmell *et al.*, 2011) and the presence of two SusCD-homologue pairs in the arabinan PUL. Reduced surface activity would cause fewer arabinooligosaccharides to be generated at the cell surface reducing the concentration of arabinooligosaccharides released into the growth supernatant. This reduced activity at the cell surface is responsible for arabinooligosaccharides being present in greater concentrations at mid-late stage of exponential growth phase in the cell supernatant (Figure 4.25.b, c).

Despite the relatively small effect of cross-feeding on the surface arabinanases mutant, the arabinooligosaccharides released may have a greater effect on certain *Bifidobacterium* species which have been shown to thrive on a mixture of low molecular weight arabinooligosaccharides (Al-Tamimi *et al.*, 2006). *Bifidobacterium* express oligosaccharide specific utilisations systems which may make them much more efficient than *Bacteroides* at oligosaccharide utilisation (Ejby *et al.*, 2013; Shigehisa *et al.*, 2015). If this assumption is correct, *B. thetaiotaomicron* could function as a probiotic species in the human gut, generating oligosaccharides which would act as *Bifidobacteria* bespoke prebiotics. This would avoid the costly alternative of including arabinooligosaccharides in the human diet to specifically boost oligosaccharide users in the gut (Al-Tamimi *et al.*, 2006).

4.3.4.3 RGI backbone (Mucilage) Cross-feeding

The undecorated Rha-GalA backbone of RGI was used as the growth substrate for the cross-feeding experiment as Δbt_4170 was unable to utilise the RGI backbone (Figure 4.26). When grown alongside wild type *B. thetaiotaomicron*, the mutant was able to grow (Figure 4.29a). This indicates that, despite the lack of detection of mucilage-derived oligosaccharides in the *B. thetaiotaomicron* growth supernatant, cross-feeding was evident (Figure 4.25d). It could be argued that oligosaccharide transport is extremely quick and thus these molecules are rapidly imported by the wild type and mutant bacteria, and thus do not accumulate in the medium. An alternative explanation for the data are that only very high molecular weight oligosaccharides are generated, similar to the GH98 endoxylanase in the arabinoxylan degrading system of *B. ovatus* (Rogowski *et al.*, 2015), which are

not visible by TLC. However, BT4170 is a very active enzyme that rapidly generates visible low molecular weight RGI oligosaccharides, which are evident in whole cell assays (Ana Luis personal communication). Thus the lack of visible oligosaccharides is unlikely to be the result of the production of high molecular weight molecules.

4.3.4.4 RGII Cross-feeding

Despite lack of detectable levels of oligosaccharides in the growth supernatant of *B. thetaiotaomicron* cultured in MM + 0.5% RGII (Figure 4.25e), the wild type bacterium was able to support the growth of the Δbt_1023 mutant (Figure 4.29e) lacking the PL1 pectate lyase required to cleave the backbone of RGII for oligosaccharide import. Interestingly, the mutant lacking the entire RGII PUL, $\Delta RGIIPUL$, was unable to grow in co-culture with wild type *B. thetaiotaomicron* (Figure 4.29c), indicating no usable monosaccharides are released during growth on RGII; oligosaccharides however, are present. Lack of detection of the oligosaccharides by TLC suggests they are either at a very low concentration or large enough not to migrate from the origin point of the TLC. Large oligosaccharide production is supported by lack of access to the backbone of RGII due to the presence of side chains blocking the activity of the surface PL1 lyase, forcing the enzyme to mediate infrequent cleavage of the backbone. This strategy has the added advantage of reducing energy output required by *B. thetaiotaomicron* to utilise a substrate, by importing larger oligosaccharides through the TonB-dependent SusC-homologues directed against RGII (Jordan *et al.*, 2013). Less ATP would be required to import a restricted number of large molecules than import multiple smaller substrates, offsetting the energy investment required to express the apparatus required to utilise a complex substrate.

4.4 Conclusion

Utilisation of pectic polysaccharides by *B. thetaiotaomicron* is a complex and costly process for the bacterium, requiring many enzymes, binding proteins and transporters to achieve complete utilisation of each fraction of these GalA-containing polysaccharides, including the highly complex

RGII. Pectin is often added as a gelling agent to processed foods, which are a large part of the western diet (Munarin *et al.*, 2012). Pectins are naturally occurring in plant material and hence a significant part of the human diet, even discounting the use of pectin in processed foods (Mohnen, 2008). As such, the HGM has evolved with the human diet as a major selective pressure, which has led to the development of comprehensive pectin degradation and utilisation systems exemplified by the apparatus synthesised by *B. thetaiotaomicron* (Martens *et al.*, 2011). Indeed, this *Bacteroides* species is capable of utilising every fraction of pectin from simple galactan to RGII, the most complex plant glycan known. Galactooligosaccharides are used in food supplements generated from galactan digestion or by transglycosylation of lactose (Macfarlane *et al.*, 2008). Use of these oligosaccharides have been shown to have positive effects on the gut and are bifidogenic (Macfarlane *et al.*, 2008). *B. thetaiotaomicron* is capable of releasing high concentrations of galactooligosaccharide during growth on galactan, which may prolong the positive effects record for diet supplementation with oligosaccharides. While similar effects have yet to be observed for other oligosaccharides derived from other pectic glycans, data here shows oligosaccharide release from *B. thetaiotaomicron* growth on these glycans are sufficient to support a second bacterium, unable to use the polysaccharide.

4.5 Future Work

The potential effects of cross-feeding and growth of *B. thetaiotaomicron* among other members of the gut microbiota on pectin polysaccharides should be monitored along with any potential recipients of the oligosaccharides generated at the *B. thetaiotaomicron* cell surface. This could be performed in a humanised murine gut microbiota to demonstrate the effect observed in the cross-feeding experiments here in a more biologically relevant setting. Galactan cross-feeding is likely to reveal recipient bacterial species, however, the oligosaccharides generated from other pectic glycans enriching members of the gut microbiota has not currently been shown.

The interactions between surface enzymes and binding proteins are currently poorly understood. Introducing mutations into coding regions most likely to be involved with these predicted protein-

protein interactions could show which residues are responsible for the interactions and the effect on the system from loss of these interactions.

Chapter 5: Cross-feeding during *Bacteroides* utilisation of Fructans

5.1 Introduction

5.1.1 Background

Current prebiotic research has a large focus on fructooligosaccharides (FOS), inulin-type fructans and their positive effects for the human gut and human gut microbiota (HGM) (Riviere *et al.*, 2016). Inulin-type fructans and inulin derived FOS (FOS-I) consist of β 2,1-linked fructose with each chain terminating in an α 1,2-linked glucose residue. The length of inulin chains can vary but are usually 2 to 60 fructose residues long, with FOS-I considered to be under 10 monosaccharides in length. Inulin is predominantly purified from chicory root, although can be purified from garlic and onion (Mensink *et al.*, 2015). Many studies have shown Inulin and FOS-I enrich *Bifidobacterium spp* and other beneficial bacterial species in the HGM, making these fructans important therapeutic tools in treating gut disease like inflammatory disorders and even shown to inhibit gut related cancers by promoting butyrate production (Hoepli *et al.* 2015). Inulin treatment has also been shown to have direct effects on the host by increasing intestinal calcium absorption (Abrams *et al.*, 2005). Studies investigating inulin and FOS-I utilisation by *Bifidobacterium* have shown preference for FOS over inulin, with some *Bifidobacterium spp* being unable to utilise the polysaccharide (Watson *et al.*, 2013; Selak *et al.*, 2016).

Levan-type fructan differs from inulin type fructan, it consists of β 2,6-linked fructose and lacks the glucose cap found in inulin chains. Levan usually exists in much longer chains than inulin. Low molecular weight levan is derived from plant material (Vijn and Smeekens, 1999), while high molecular weight forms of the glycan are often branched and found in the extra cellular polysaccharide matrix surrounding biofilms or the bacterial capsule. Levan produced by *Erwinia herbicola* and *Bacillus subtilis* is highly branched and assists in biofilm cohesion (Blake *et al.*, 1982; Benigar *et al.*, 2014).

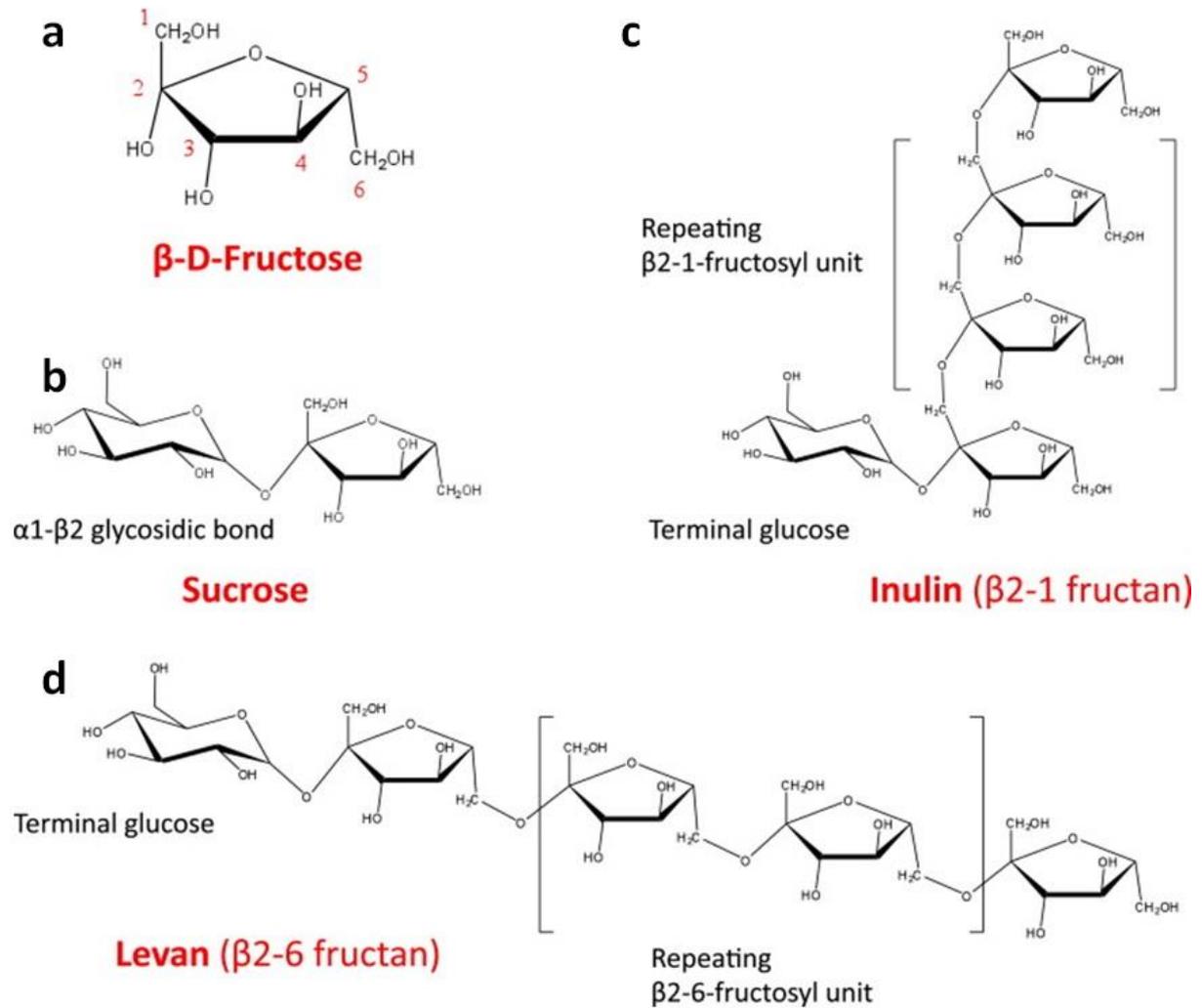


Figure 5.1 Schematic diagram of fructan structures. Structures of β -D-Fructose (a), Sucrose (b), β 2-1 linked fructan, Inulin (c) and β 2-1 linked fructan, Levan (d). Adapted from Sonnenburg *et al.* 2010.

Evidence shows humans have consumed fructan since the prehistoric era, estimating 135 g were consumed per day (Leach and Sobolik, 2010). Current fructan/FOS consumption has dropped to only 3-11 g per day (van Loo *et al.*, 1995), indicating drastic changes to the human diet which evolution of the host may not yet have compensated for, hence and a potential requirement for dietary fructan supplementation.

HGM are flooded with dietary complex carbohydrates and host glycans produced by intestinal epithelial cells as part of the protective mucous layer in the gut. As such members of the HGM have developed systems to effectively utilise the glycans available in the human gut, the strategies for

which are influenced by the presence or absence of a periplasm between the inner and outer membrane. The two contrasting strategies discussed here are that of Gram positive *Bifidobacterium* and Gram negative *Bacteroides*, each presenting specific obstacles to overcome in utilisation of glycans (Van der Meulen *et al.*, 2006). *Bifidobacterium* possesses a thick peptidoglycan cell wall attached to a cell membrane. An extracellular solute binding protein (ESBP) is expressed which extends through the peptidoglycan cell wall to capture glycans the bacterium encounters. Once bound by the ESBP the glycan is delivered to an associated ABC-transporter in the cell membrane where it is imported into the cytoplasm. Once sequestered in the cytoplasm glycoside hydrolases expressed by *Bifidobacterium* digest the target glycan into fermentable substrates, mono- or disaccharides (Ejby *et al.*, 2013).

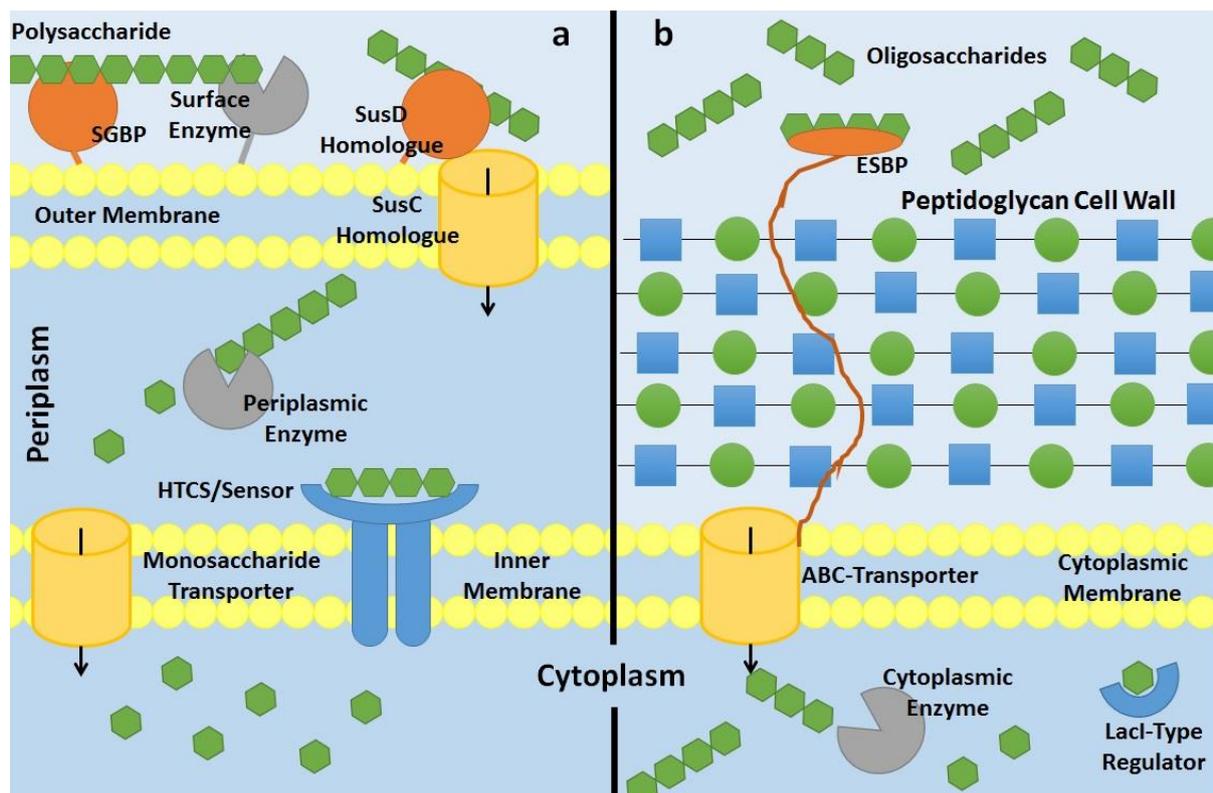


Figure 5.2 Simplified diagram of glycan utilisation. Diagram of *Bacteroides* glycan utilisation system (a) and *Bifidobacterium* glycan utilisation system (b). Binding proteins are shown in orange, Transport proteins in yellow, enzymes in grey and sensor/regulators in blue. Glycan structure is represented by green hexagons and the phospholipid bilayer of the inner/outer membranes as yellow circles.

Bacteroides possess both an inner and outer membrane, requiring digestion of glycans by surface enzymes to generate oligosaccharides which are small enough to be transport into the periplasm through the transporter system comprising of complexes of SusC/SusD homologues present in the outer membrane. The SusD homologue binds the oligosaccharides produced at the cell surface while the SusC homologue is predicted to be a TonB-dependent transporter which allows passage of oligosaccharides through the outer-membrane. In the periplasm, primarily exo-acting glycoside hydrolases further deconstruct the oligosaccharides into fermentable substrates which are then imported into the cytoplasm (for further discussion of *Bacteroides* glycan utilisation systems, including the SusC/SusD complex, see Chapter 4.3). The presence of the periplasm in *Bacteroides* cells allow for deconstruction of glycans and fermentation to take place in separate compartments (Martens *et al.*, 2009). Whereas *Bifidobacteria* tend to utilise oligosaccharides and some short polysaccharides thus, glycan utilisation is limited by size of the target polysaccharide as import requires shorter substrates (Ejby *et al.*, 2013). Once in the cytoplasm, the exo-acting glycoside hydrolases expressed by *Bifidobacterium* can degrade the glycan into fermentable monosaccharides.

Cross-feeding between gut microbes grown on polysaccharides is a relatively unexplored interaction between members of the gut microbiota. Studies have shown cross-feeding of oligosaccharides between *Bacteroides* species (Rakoff-Nahomen *et al.*, 2014) and between *B. thetaiotaomicron* and *Bifidobacterium* (Van der Meulen *et al.*, 2006), however the former focused on recipient growth on *Bacteroides* conditioned media over and extended growth period while the latter recorded production of metabolites as a measure of growth rather than cell density of each species. This chapter aims to demonstrate growth of both glycan donor and recipient in co-culture on fructans.

5.1.2 Aims

- Demonstrate growth of *Bifidobacterium adolescentis* (*Bi. adolescentis*) and *Bi. longum* growth on inulin and levan and FOS derived from each fructan.
- Explore genes/potential loci responsible for fructan utilisation in *Bi. adolescentis* and *Bi. longum*.
- Demonstrate effect of crossfeeding of both inulin and levan type fructans.

5.2 Results

5.2.1 Growth on Fructans

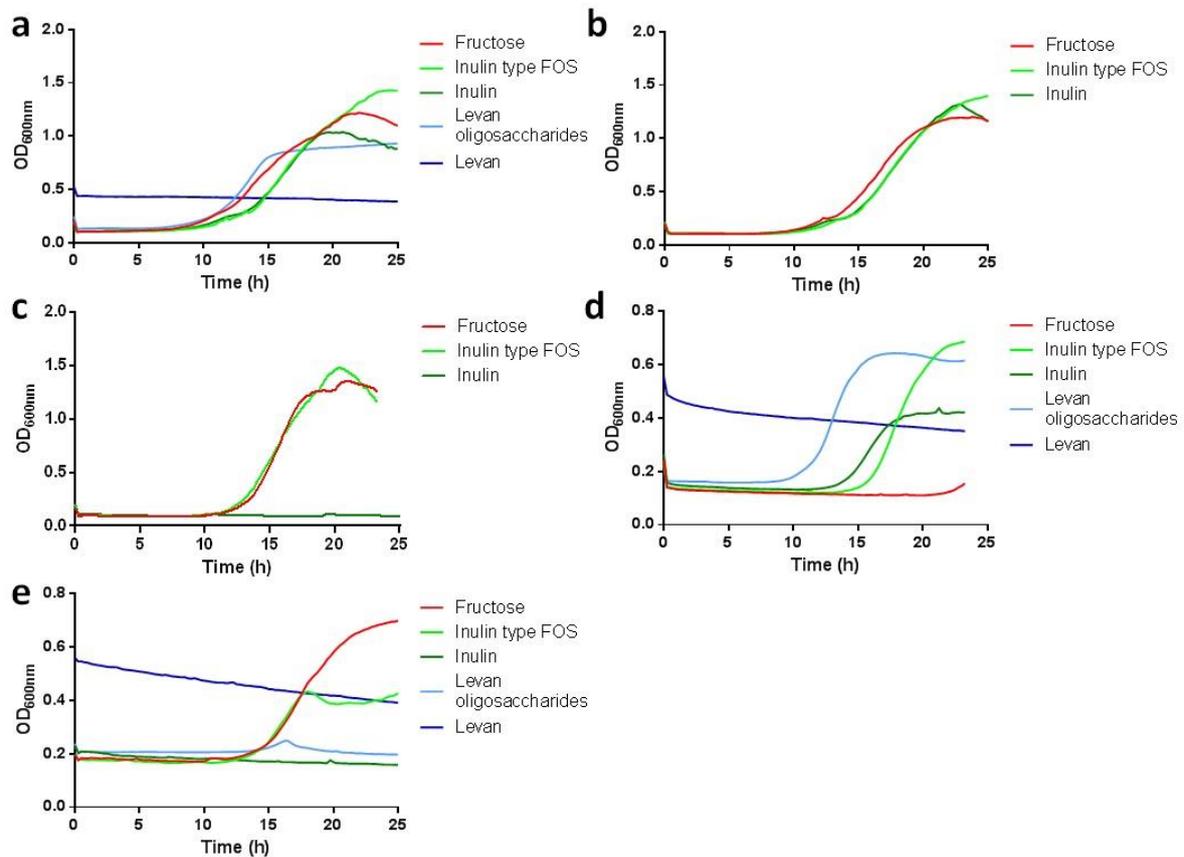


Figure 5.3 Growth of *B. ovatus*, *B. vulgatus*, *Bi. adolescentis* and *Bi. longum* on fructans, FOS-I, FOS-L and fructose. *B. ovatus* (a), *B. ovatus* Δ GH91 mutant (b), *B. vulgatus* (c), *Bi. adolescentis* (d) and *Bi. longum* (e) were cultured on fructose, FOS-I, FOS-L and fructans. Growths were performed in 2 ml plater reader cultures measured using an automatic plate reader in an anaerobic chamber at 37 °C. *Bacteroides* and *Bifidobacterium* were grown in minimal media (MM) and *Bifidobacterium* minimal media (BiMM), respectively, the composition of which can be found in Chapter 2.3.1.

5.2.1.1 *B. ovatus*, *B. vulgatus*, *Bi. adolescentis* and *Bi. longum* growth on fructans

Growth of *Bifidobacterium* spp. and *Bacteroides* spp. on fructans is well described in the literature (Kolida and Gibson, 2007; Sonnenburg *et al.*, 2010). To confirm the organisms used here displayed this phenotype the growth of the bacteria employed in this study was tested on fructans and FOS.

Bacteroides spp. were grown in minimal media (MM) while *Bifidobacterium spp.* were grown in *Bifidobacterium* minimal media (BiMM) (Van der Meulen *et al.*, 2006), which were supplemented with 0.5 % (w/v) sugar and haematin. Growths were performed in an automatic plate reader under anaerobic conditions at 37 °C. *Bi. adolescentis* and *Bi. longum* produce a thick capsule during growth. In the 200 µl wells with smaller diameter capsule production caused false growth readings due to occlusion of the plate reader light paths, using a larger diameter well led to more reliable data, hence 2 ml wells were used for *Bifidobacterium* in the automatic plate reader.

The data showed *B. ovatus* and *Bi. adolescentis* could utilise both inulin and FOS-I (Figure 5.3a,d), while *B. vulgatus* and *Bi. longum* grew on FOS-I but not the polysaccharide inulin (Figure 5.3c, e). *B. ovatus*, *B. vulgatus*, *Bi. adolescentis* and *Bi. longum* were unable to utilise levan (Figure 5.3a,c,d,e). However, *B. ovatus*, *Bi. adolescentis* and *Bi. longum* grew on FOS-L (DP 2-8), generated by partial digestion of 10 mg/ml levan with 0.5 µM BT_1760 for 30 min (Figure 5.5b). The three *Bacteroides* species and *Bi. longum* grew on fructose, while *Bi. adolescentis* was unable to use the monosaccharide. These data are in agreement with previous studies investigating *Bacteroides spp.* and *Bifidobacterium spp.* growth on fructans (Kolida and Gibson, 2007), indicating the bacterial species/strains display the expected growth phenotypes. These *Bacteroides* and *Bifidobacterium spp.* were selected due to their fructan utilisation profiles (Figure 5.3). The *Bacteroides spp.* were selected as they are able to utilise inulin or levan with the possibility of oligosaccharide release. Whereas *Bi. adolescentis* ability to utilise both FOS-I and inulin gave an opportunity to investigate competition with *B. ovatus*. *Bi. longum* and *B. vulgatus* were able to utilise FOS-I but not inulin allowing investigation of crossfeeding with *B. ovatus*.

The mutant of *B. ovatus* lacking the surface GH91 endo-inulinase, designated $\Delta GH91$, was generated by Sarah Shapiro (Shapiro 2015). This Mutant was capable of utilising inulin, FOS-I and fructose (Figure 5.3b).

5.2.1.2 Growth Supernatants

During growth on inulin- and levan-type fructans samples of culture supernatant were subjected to TLC to evaluate oligosaccharide content (Figure 5.4). *B. ovatus* culture supernatant grown on inulin showed a high concentration of a wide range of oligosaccharides over the entire growth (Figure 5.4a), including early stationary phase. During early growth there was very little fructose present in the supernatant, but low concentrations of the monosaccharide were evident in late log and stationary phase (Figure 5.4a). Longer oligosaccharides, apparent as a long smear from the origin spot, are used quickly and were almost completely absent by mid-late log phase (Figure 5.4a). A band present in all supernatant samples from log phase onwards was identified as difructose anhydride (dFA) by Sarah Shapiro; this molecule is implicated in increased calcium and iron absorption in the gut (Suzuki and Hara, 2004; Hara *et al.*, 2010) but cannot be metabolised by *B. ovatus* (Shapiro 2015).

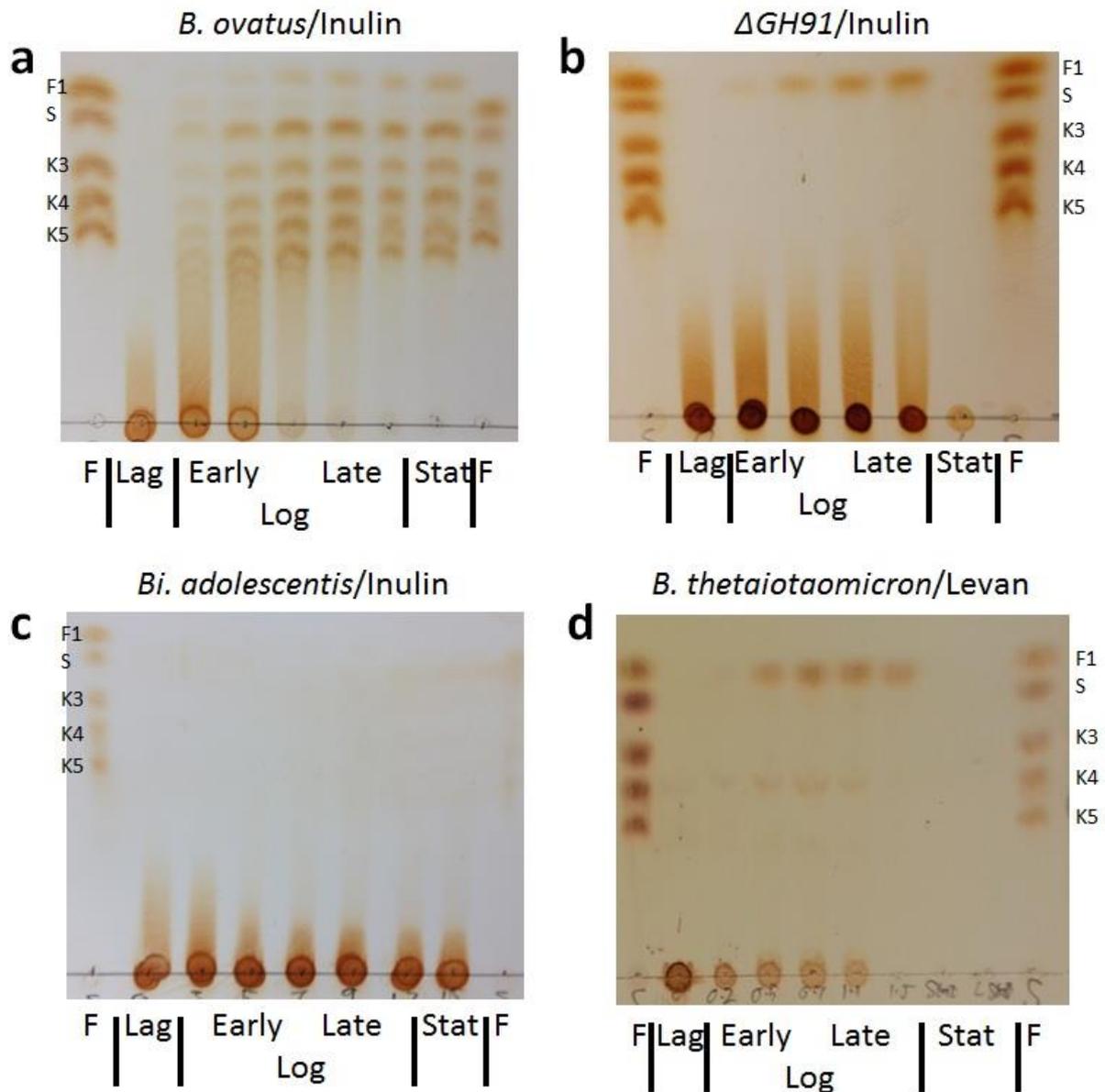


Figure 5.4 TLC images of fructan growth supernatants. Samples were taken from *B. ovatus* growth on inulin (a), $\Delta GH91$ mutant growth on inulin (b), *Bi. adolescentis* growth on inulin (c) and *B. thetaiotaomicron* growth on levan (d). Growth supernatant taken at each phase of growth and run alongside known fructooligosaccharide standards (F). Standards used (lane F) were fructose, F1, sucrose, S, and the FOS-Is kestotriose, K3, kestotetraose, K4 and kestopentaose, K5.

A mutant of *B. ovatus* created by Sarah Shapiro, $\Delta GH91$, was also grown on inulin and supernatant samples run on TLC (Figure 5.4b). Analysis of the supernatant of the mutant revealed no oligosaccharide release and much slower utilisation of inulin than wild type *B. ovatus*, as evidenced by the presence of a dark spot present at the origin until stationary phase (Figure 5.4b). This reflects

inulin degradation in the mutant has shifted to the periplasm where the transient oligosaccharides generated during the degradative process are not released into the environment (Figure 5.4b).

Similar to the *B. ovatus* $\Delta GH91$ mutant, *Bi. adolescentis* utilises inulin without surface enzyme activity, hence no oligosaccharides were present in the growth supernatant (Figure 5.4c).

Throughout the growth the long oligosaccharides present, shown as a smear in on the TLC plate, was shortened, indicating *Bi. adolescentis* was utilising the long oligosaccharides. Although *Bi. adolescentis* utilised inulin to provide good growth, the growth supernatant shows very little inulin was used (reflecting the retention of the glycan spot at the origin; Figure 5.4c) compared to wild type *B. ovatus* (Figure 5.4a).

When characterised the surface GH32 of the *B. thetaiotaomicron* levan PUL showed high activity on levan, producing a high concentration of FOS-L in vitro (Sonnenburg *et al.*, 2010). However, the culture supernatant of *B. thetaiotaomicron* grown on levan contained low levels of a single oligosaccharide that co-migrated with kestotetraose and fructose. The fructose band increased in intensity until late exponential, then was absent in stationary phase (Figure 5.4d). Low supernatant FOS-L concentration could reflect an efficient binding and transport system which stops accumulation of FOS-L DP 2-3, or perhaps the surface GH32 is not highly expressed in the utilisation system, resulting in lower surface activity. Although this is unlikely as *B. thetaiotaomicron* completely utilised the levan starting material before reaching stationary phase (Figure 5.4d), indicating a high activity on levan polysaccharide.

It should be noted here that TLC is only effective at separating oligosaccharides and monosaccharides in a fairly narrow range of DP. There are other techniques which are effective at separating longer glycans from a sample solution. HPAEC, for example can be used to separate glycans based on charge/size of the glycans in solution. Using different gradients of NaOH concentration to elute glycans could allow for separation of long FOS which appears as a smear or does not leave the origin spot on the TLC plate.

5.2.1.3 Fructooligosaccharide utilisation by *Bi. adolescentis* and *Bi. longum*

Bi. adolescentis and *Bi. longum* were grown to stationary phase on inulin, levan, FOS-I and FOS-L and products released in to the supernatant analysed by TLC. As no growth was observed for either bacterium on levan, or *Bi. longum* on inulin, there was no stationary phase sample. In the case of no growth a sample was taken at 20 h to assay for any partial breakdown of the polysaccharide (Figure 5.5), the same incubation time for *Bi. adolescentis* to reach late stationary phase on inulin. The stationary phase samples of the bacteria grown on inulin were identical to the starting material supernatant. Growth on FOS-I showed that *Bi. adolescentis* and *Bi. Longum* utilised longer and shorter oligosaccharides, respectively (Figure 5.5a). In agreement with growth data in the previous section (Figure 5.3), *Bi. adolescentis* did not utilise fructose (Figure 5.5a).

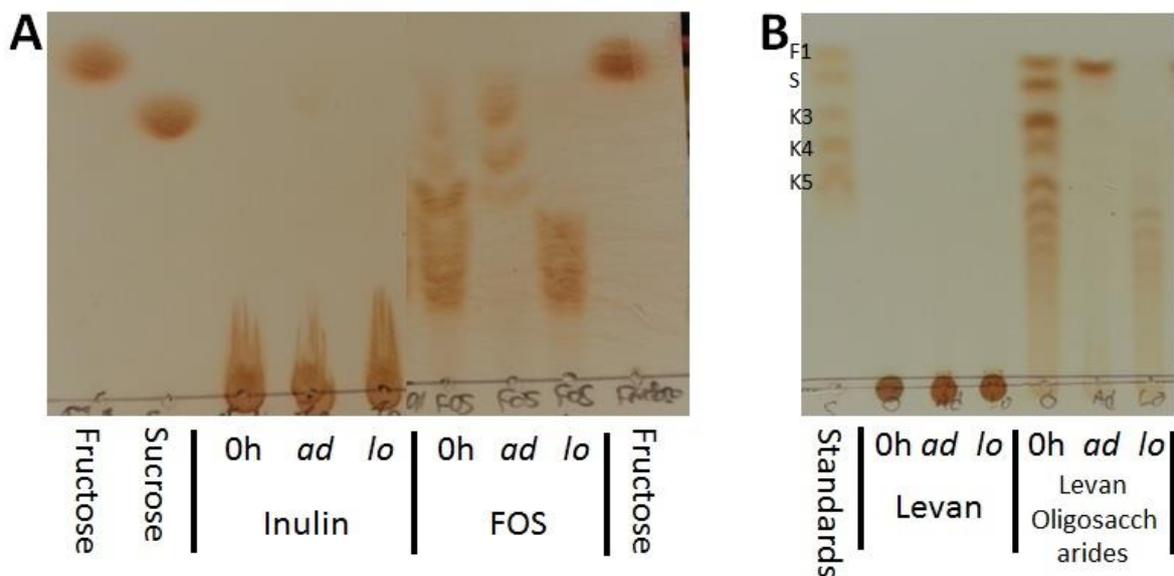


Figure 5.5 Fructooligosaccharide utilisation by *Bi. adolescentis* and *Bi. longum*. *Bi. adolescentis* (ad) and *Bi. longum* (lo) were grown to stationary phase on inulin, FOS-I (a) and FOS-L (b) and the glycan content of the culture supernatant was evaluated by TLC. Stationary phase supernatants (ad for *Bi. adolescentis* and lo for *Bi. longum*) were compared to the glycan content of media prior to addition of bacteria (0h) and run alongside known standards, fructose, F1, sucrose, S, kestotriose, K3, kestotetraose, K4 and kestopentaose, K5.

Bi. adolescentis completely utilised all oligosaccharides present, with the exception of fructose. A similar pattern of utilisation was observed for *Bi. longum* on FOS-L (Figure 5.5b) as the FOS-I (Figure 5.5a). *Bi. longum* was able to utilise fructose and oligosaccharides up to a DP of 5 fructose monosaccharides in length (Figure 3.5b). Bands corresponding to higher DP oligosaccharides, however, appeared to be less intense than in the starting material, indicating *Bi. longum* was capable of utilising oligosaccharides with a DP >5 but at a slow rate. These data show *Bi. adolescentis* and *Bi. longum* preferentially utilise different fractions of FOS regardless of linkage type.

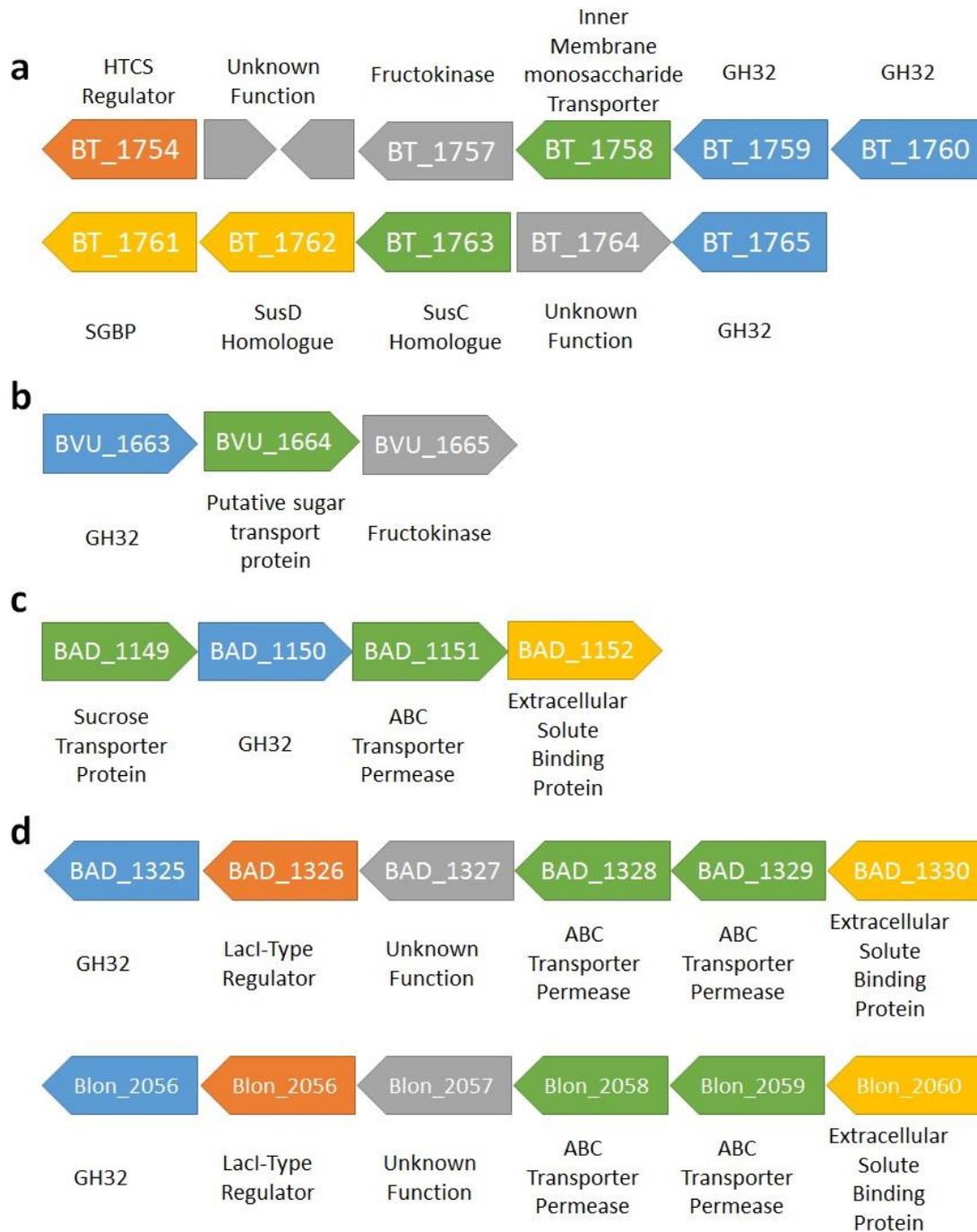


Figure 5.6 Schematic representation of predicted fructan utilisation loci of *Bi. adolescentis* and *Bi. longum*. Predicted glycan utilisation loci of *B. thetaiotaomicron* (a), *B. vulgatus* (b), *Bi. adolescentis* (c) and *Bi. longum* (d). The colour code for the predicted proteins are as follows: transporter proteins, green; glycoside hydrolases, blue; Surface glycan binding proteins, yellow; Regulator protein, orange; unknown function, grey. Symbol size is not representative of gene size.

5.2.2 Analysis of inulin and levan utilisation loci

The observed differences in fructan utilisation among the bacterial species tested here (Figure 5.3) may be explained by analysis of the loci most likely to be involved in the process. Unlike *B. thetaiotaomicron* which encodes multiple GH32 enzymes (Sonnenburg *et al.*, 2010)(Figure 5.6a), *B. vulgatus* expresses a single GH32 enzyme (BVU_1663, Figure 5.6b). Members of the GH32 family typically display activity on fructans or FOS, hence indicate an ability to utilise fructans. According to signal peptide analysis using LipoP server (LipoP site/reference), BVU_1663 is most likely expressed in the periplasm. BVU_1663 shows 64 % sequence identity to BT_1765, a non-specific exo-acting fructosidase (Sonnenburg *et al.*, 2010) and likely performs the same function. The presence of the GH32 enzyme in the periplasm suggests *B. vulgatus* requires digestion of fructans into FOS to be completed externally prior to import into the periplasm where BVU_1663 is able to complete digestion.

Bi. longum and *Bi. adolescentis* utilisation systems typically target oligosaccharides rather than polysaccharides, which is supported by growth data collected here (Figure 5.5). *Bi. adolescentis* encodes a pair of GH32 enzymes BAD_1150 and BAD_1325. *bad_1150* is flanked by genes predicted to encode components of an ABC transporter and an ESBP (Figure 5.6a). This gene organisational cluster is suggestive of an oligosaccharide utilisation system. *bad_1325*, in contrast, is immediately preceded by a LacI-type regulator gene, while ABC transporter components and an ESBP are encoded by genes upstream. These predicted loci were found by search sequence data for homologues of GH32 enzymes in the *Bi. adolescentis* genome for when they are found adjacent to ABC-pemases and ESBP homologues.

Bi. longum also encodes two GH32 enzymes, Blon_0787 and Blon_2056. *blon_2056* is surrounded by genes encoding a predicted transport, ESBP and LacI-type regulator (Figure 5.6b). This gene organisation is the same as the locus identified in *Bi. adolescentis*, *bad_1325* to *bad_1330*, implying similar function. The gene second GH32 enzyme of *Bi. longum*, Blon_0787, is not in any obvious loci,

although may be upregulated along with genes further away on the genome to form a functional oligosaccharide utilisation system.

5.2.3 Cross-feeding with fructans

Utilisation of inulin and levan by *B. ovatus* and *B. thetaiotaomicron*, respectively, leads to release of oligosaccharides into the culture supernatant (Figure 5.4). In co-culture with FOS users we can explore how species equipped with polysaccharide utilisation systems can create growth substrates for oligosaccharide utilising bacteria present in the human gut.

5.2.3.1 Inulin Cross-feeding

Fructans and FOS are currently used as dietary supplements to enrich for gut microbes associated with good gut health (Ouwehand *et al.*, 2002). Studies have shown supplementation of food with FOS-I and inulin enrich the HGM for *Bifidobacterium* increasing butyrate in the gut (Ouwehand *et al.*, 2002). Butyrate is utilised for energy by colonocytes, inhibits proliferation of cancerous cells and triggers maturation of naive T-cells into Treg cells (Hoepli *et al.* 2015).

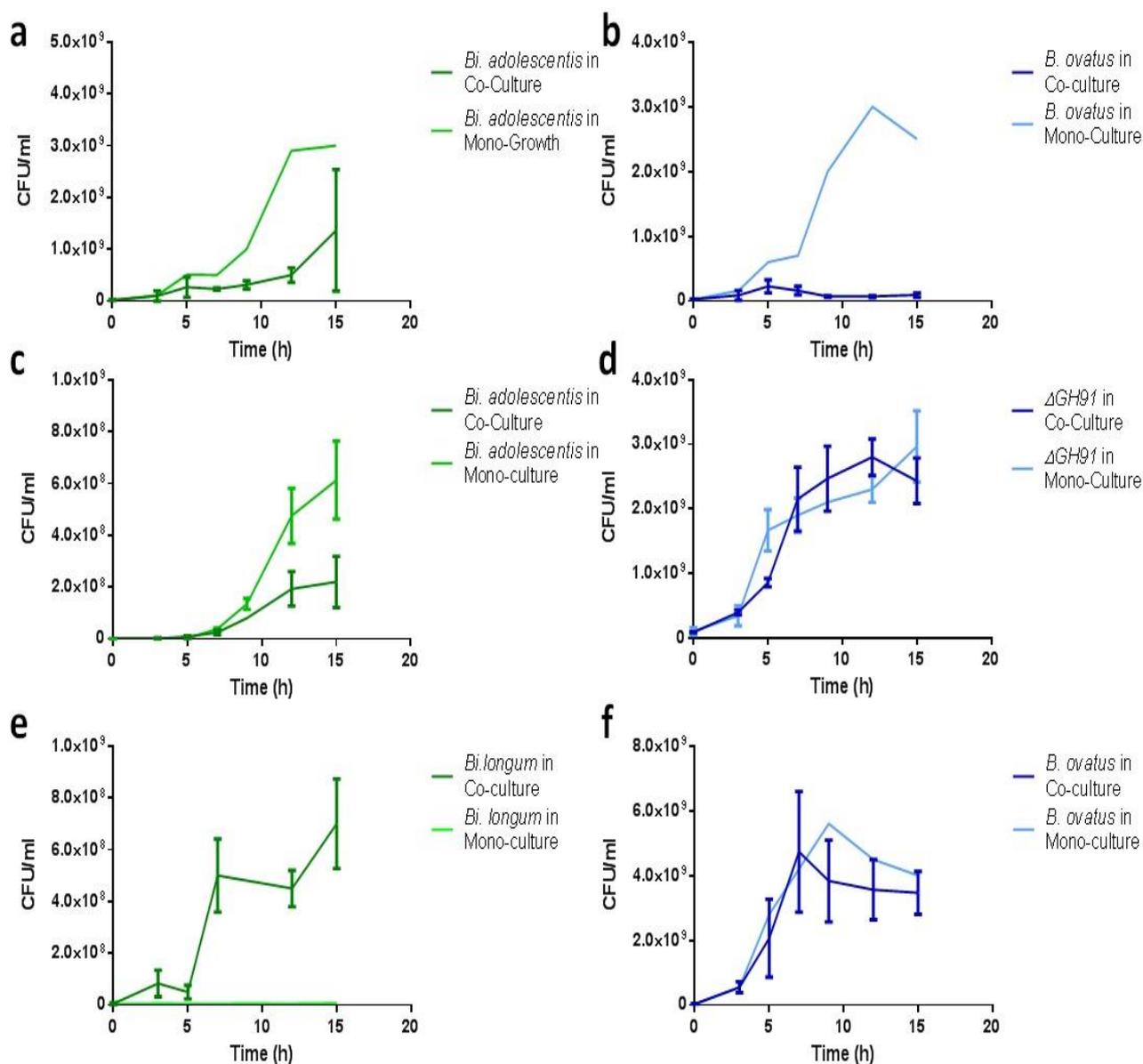


Figure 5.7 Inulin cross-feeding with *Bi. adolescentis* and *Bi. longum*. Comparison of *Bi. adolescentis* growth in mono-culture and co-culture with *B. ovatus* on inulin (a), and *B. ovatus* in mono-culture and co-culture with *Bi. adolescentis* on inulin (b). Comparison of *Bi. adolescentis* growth in mono-culture and co-culture with *B. ovatus* Δ GH91 mutant on inulin (c), and *B. ovatus* Δ GH91 mutant in mono-culture and co-culture with *Bi. adolescentis* on inulin (d). Comparison of *Bi. longum* growth in mono-culture and co-culture with *B. ovatus* on inulin (e), and *B. ovatus* in mono-culture and co-culture with *Bi. adolescentis* on inulin (f). Samples were taken from the growth cultures, and plated onto clostridial media for CFU evaluation. As in Chapter 3 bacteria were identified by colony morphology and resistance/susceptibility to gentamycin.

5.2.3.1.1 *B. ovatus* Co-culture with *Bi. adolescentis*

Bi. adolescentis and *B. ovatus* were co-cultured in BiMM+0.5 % inulin. Both *B. ovatus* and *Bi.*

adolescentis can utilise inulin as a growth substrate through different strategies. In co-culture with

Bi. adolescentis, *B. ovatus* fails to grow (Figure 5.7b). *B. ovatus* growth peaks at 5 h, at which point the ratio of *B. ovatus* to *Bi. adolescentis* is 50:50 (Figure 5.7a,b). After a further 5 h, *B. ovatus* cell density dropped to inoculation levels while *Bi. adolescentis* achieved a peak density of 1.3×10^9 CFU/ml (Figure 5.7a,b). Both *Bi. adolescentis* and *B. ovatus* in co-culture were unable to reach the cell density of their respective mono-cultures on inulin (Figure 5.7a,b), although *Bi. adolescentis* in co-culture was still growing at the final time point, and may achieve similar CFU as the mono-culture if allowed longer incubation (Figure 5.7a).

Samples of culture supernatant were taken at intervals during the *B. ovatus* – *Bi. adolescentis* co-culture on inulin and run on TLC (Figure 5.8a). Early time points were similar to that of *B. ovatus* in mono-culture although after 5 h there were no more oligosaccharides produced other than sucrose (Figure 5.8a). Unlike the *B. ovatus* mono-culture there was inulin polysaccharide present at the final 15 h sample (Figure 5.8a). dFA and low concentrations of fructose were present from 3 h onwards. From 5 h onwards the smear of long oligosaccharides became shorter (Figure 5.8a), implying *Bi. adolescentis* was using these glycans as growth substrates.

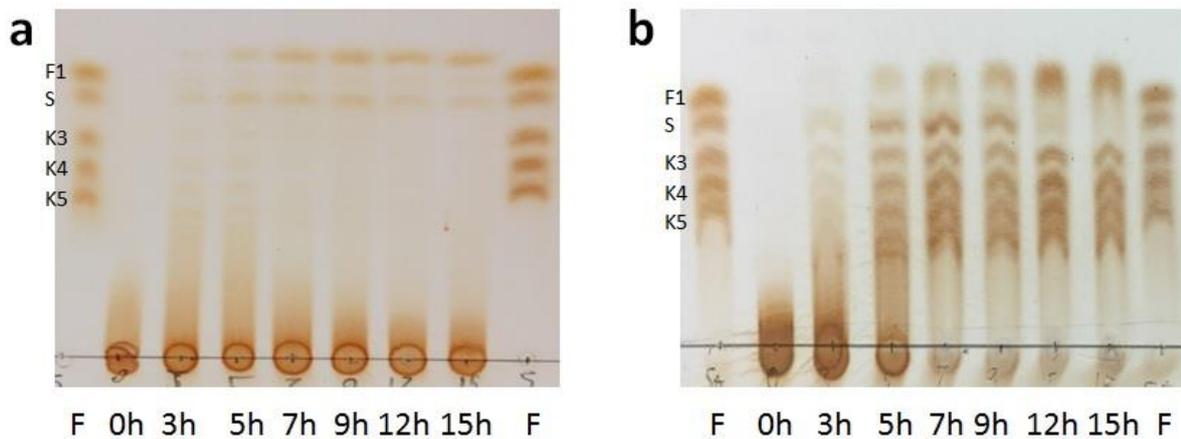


Figure 5.8 Supernatant glycan content of inulin co-cultures. Samples taken during *B. ovatus*-*Bi. adolescentis* co-culture on inulin (a) and *B. ovatus*-*Bi. longum* co-culture on inulin (b) were taken and run on TLC alongside known fructooligosaccharide standards (F), fructose, F1, sucrose, S, kestotriose, K3, kestotetraose, K4 and kestopentaose, K5.

5.2.3.1.2 $\Delta GH91$ *B. ovatus* mutant cross-feeding with *Bi. adolescentis*

The mutant *B. ovatus*, $\Delta GH91$, was co-cultured with *Bi. adolescentis* on inulin. $\Delta GH91$ showed similar growth in the co-culture as the mono-culture (Figure 5.7b). The *Bi. adolescentis* CFU was slightly reduced in co-culture when compared to mono-culture (Figure 5.7c).

5.2.3.1.3 *B. ovatus* cross-feeding with *Bi. longum*

A bacterium, *Bi. longum*, which is unable to utilise inulin was co-cultured with *B. ovatus* in BiMM+0.5 % inulin. Alone *Bi. adolescentis* cell density remained at inoculation levels, in co-culture *Bi. longum* was able to achieve cell density of 6.5×10^8 CFU/ml (Figure 5.7e), similar to the related bacterium *Bi. adolescentis*, which is able to utilise inulin. *B. ovatus* did not suffer any growth defect in co-culture with *Bi. longum*, achieving a similar cell density as the mono-culture (Figure 5.7f).

Samples of the co-growth supernatant were taken and evaluated for oligosaccharide content by TLC. The oligosaccharide profile throughout the growth was very similar to that of *B. ovatus* alone on inulin (Figure 5.8b), with the exception of the absence of sucrose from the 12 h and 15 h samples (Figure 5.8b).

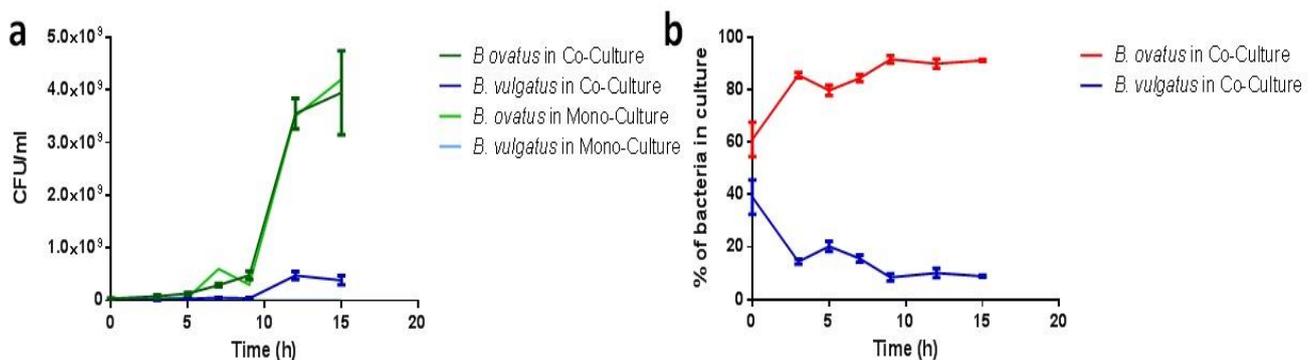


Figure 5.9 Inulin Cross-feeding with *B. vulgatus*. Co-culture of *B. vulgatus* and *B. ovatus* in MM with 0.5 % inulin measured by CFU/ml (a). Ratio of *B. ovatus* to *B. vulgatus* in the co-culture. Ratio was deduced by purification of genomic DNA at time points in the co-culture and quantification of specific tags inserted into the genomic DNA of the two organisms by qPCR (b).

5.2.3.1.4 *B. ovatus* cross-feeding with *B. vulgatus*

In mono-culture *B. vulgatus* was unable to utilise inulin but was able to utilise FOS-I (Figure 5.3c). In co-culture with *B. ovatus* on inulin *B. vulgatus* grew to 5×10^8 CFU/ml (Figure 5.9a), which equated to 10 % of the culture at 15 h (Figure 5.9b). From inoculation *B. vulgatus* dropped from 40 to 20 % of the co-culture, then from 5 h onward to 10 % in the 15 h sample (Figure 5.9b). *B. ovatus* in the co-culture reached a similar CFU as in mono-culture (Figure 5.9a).

5.2.3.2 Levan Cross-feeding

In co-culture with *B. thetaiotaomicron* on levan, *Bi. adolescentis* was able to grow (Figure 5.10a) suggesting the relatively low concentration of oligosaccharides present in the *B. thetaiotaomicron* growth supernatant (Figure 5.4d) was sufficient to support *Bi. adolescentis* growth. There was a 160-fold increase in final cell concentration from the *Bi. adloescentis* mono-culture on levan to the co-culture with *B. thetaiotaomicron* (Figure 5.10a). Conversely, *B. thetaiotomicron* cell density was reduced in the co-culture compared to the mono-culture (Figure 5.10b).

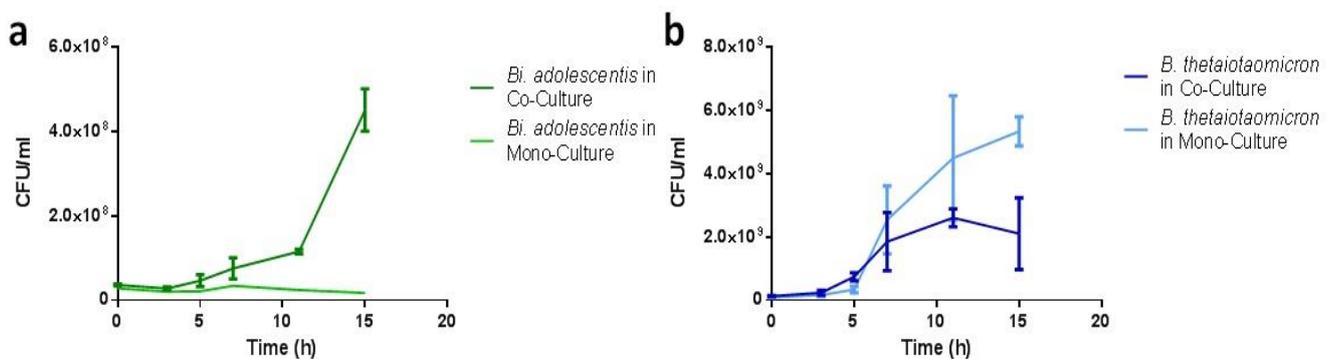


Figure 5.10 Crossfeeding of levan with *Bi. adolescentis*. Co-Culture of *Bi. adolescentis* and *B. ovatus* in MM with 0.5 % inulin measured by CFU/ml. Comparison of *Bi. adolescentis* growth in co-culture and mono-culture on inulin (a). Comparison of *B. ovatus* growth in co-culture and mono-culture on inulin (b).

5.2.4 Growth on Conditioned Media

The presence or absence of antimicrobial compounds released by *Bi. adolescentis* during growth on inulin with *B. ovatus* may explain why growth of the *Bacteroides* species was limited. To address this

hypothesis growth media were harvested from log phase mono- and co-cultures of *Bi. adolescentis* and *Bi. adolescentis* with *B. ovatus*. These media were then split, half boiled and half not boiled before being inoculated with *B. ovatus*. It was argued that boiling the media would inactivate any proteinaceous antimicrobial present. The *Bi. adolescentis* mono-culture conditioned media allowed growth of *B. ovatus*, when boiled or not boiled (Figure 5.11a), indicating there was no antimicrobial compound present. The co-culture conditioned media also allowed growth of *B. ovatus* (Figure 5.11b), again demonstrating no antimicrobial compound is present. Boiling the media caused slower growth in both mono-culture and co-culture conditioned media (Figure 5.11). This may be due certain components of the media being heat sensitive; cysteine and vitamin K are both known to be made unstable due to excessive boiling.

5.3 Discussion

5.3.1 *Bacteroides* and *Bifidobacterium* Fructan and FOS Utilisation

Utilisation of fructans and FOS by gut bacteria is a high interest area of research due to the positive health benefits described for the host in both *in vitro* studies and clinical trials (Kolida and Gibson, 2007). An inulin utilisation locus was identified due to upregulation in response to growth on the polysaccharide (Martens *et al.*, 2011). The utilisation system expressed from this locus has been characterised, revealing an efficient polysaccharide breakdown pathway (Shapiro 2015). Digestion of inulin into FOS-I is completed by a pair of GH91 enzymes, which together make a functional enzyme, covalently attached to cell surface. These oligosaccharides are bound and transported through a SusC homologue into the periplasm where digestion is completed by GH32 fructosidases. Most likely due to the short length of inulin polysaccharide, growth of the $\Delta GH91$ mutant shows that surface digestion of inulin is not required (Figure 5.3). This may also be why *Bi. adolescentis* is capable of utilising inulin (Figure 5.3) when *Bifidobacterium* are typically oligosaccharide users. There are two likely inulin/FOS utilisation loci in the *Bi. adolescentis* genome (Figure 5.6a). These loci were identified by the presence of a GH32 enzyme along with the components of an ABC-transporter and

an ESBP. The second of the two loci also includes a LacI-type regulator (Figure 5.6a). *Bi. longum* is able to utilise FOS-I but not inulin (Figure 5.3), despite possessing a similar locus to *Bi. adolescentis* (Figure 5.6). The different utilisation patterns observed during growth of *Bi. longum* and *Bi. adolescentis* on fructans and FOS in Figure 5.3 may be explained by *Bi. adolescentis* possessing two fructan targeting glycan utilisation loci whereas *Bi. longum* appears to express just one.

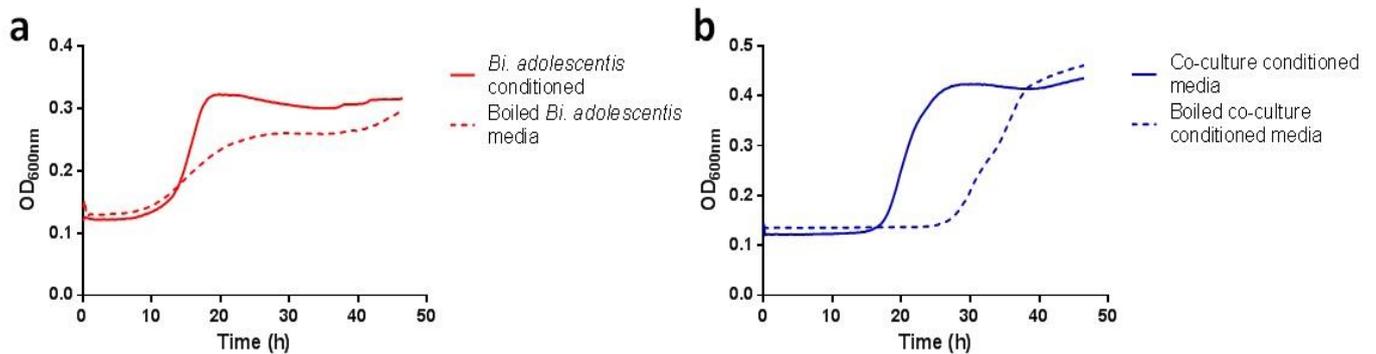


Figure 5.11 | Growth of *B. ovatus* on conditioned media. Comparison of *B. ovatus* growth on boiled and not boiled BiMM + 0.5 % inulin media conditioned by growth of *Bi. adolescentis* (a) and co-culture of *B. ovatus* and *Bi. adolescentis* (b). Growths were performed in an anaerobic chamber at 37 °C using an automatic plate reader.

As *B. ovatus*, *Bi. longum* and *Bi. adolescentis* are all capable of utilising levan oligosaccharides but not levan (Figure 5.3), and do not appear to possess specific levan utilisation loci, the GH32 enzymes present in the identified inulin/FOS utilisation systems are likely not linkage specific fructosidases.

Unlike *B. ovatus*, *B. vulgatus* requires digestion of inulin into oligosaccharides prior to transport into the periplasm where the only GH32 fructosidase expressed by *B. vulgatus* can complete degradation to fructose and sucrose. It is tempting to suggest that the reason *B. vulgatus* cannot utilise inulin but can utilise FOS-I is due to lack of surface inulinase activity, however Δ GH91 also lacks surface activity but can utilise inulin and FOS-I. Perhaps it is the combination of no surface activity and a SusD homologue which only binds short oligosaccharides, unable to recognise inulin as the *B. ovatus* inulin PUL SusD homologue.

5.3.2 *B. ovatus* shows restricted growth in co-culture with *Bi. adolescentis* on inulin but not with *Bi. longum*.

Both bacterial species in the *B. ovatus*-*Bi. adolescentis* co-culture were able to utilise inulin independently in mono-culture. However, when together in co-culture only *Bi. adolescentis* is able to achieve cell density similar to that of the mono-culture, while *B. ovatus* displayed little if any growth. Supernatant oligosaccharide data suggest each bacterium shows preference for different inulin/FOS-I fractions, as evidenced by *Bi. adolescentis* achieving high cell density while being unable to utilise much of the inulin available. *B. ovatus* on the other hand, utilises all inulin and inulin oligosaccharides leaving only the small molecule dFA in the supernatant at stationary phase. Lack of *B. ovatus* growth in co-culture despite plentiful accessible substrate would suggest *Bi. adolescentis* actively inhibits the growth of *B. ovatus*. Antimicrobial potential of *Bifidobacterium* has been previously explored. Antimicrobial compounds were found to be produced by *Bi. animalis* and *Bi. bifidum*, as non-proteinaceous pH dependent compounds (de Oliveira *et al.*, 2015). *Bi. adolescentis* has been shown to have antiviral effects on herpes simplex virus but there is no evidence of antibacterial activity (An *et al.*, 2012). Inulin supplemented media pre-conditioned with *Bi. adolescentis* alone and in combination with *B. ovatus* was used to investigate presence of an antimicrobial compound. However the inhibitory effects observed in the *B. ovatus* – *Bi. adolescentis* inulin co-culture was not observed, indicating the effect seen is not due to active killing of *B. ovatus*. This is further reinforced by growth of the $\Delta GH91$ mutant in co-culture with *Bi. adolescentis* similar to that of the mono-culture (Figure 5.7c,d), further indicating there is no antimicrobial activity.

Supernatant collected from the $\Delta GH91$ mutant grown on inulin showed only fructose despite growth similar to that of wild type *B. ovatus* on inulin (Figure 5.4b), while *Bi. adolescentis* gains a slight improvement in growth over the *Bi. adolescentis* mono-culture (Figure 5.7f). Stationary phase supernatants of *Bi. adolescentis* show the bacterium utilises longer FOS while leaving the shorter

FOS (Figure 5.5a), indicating *Bi. adolescentis* was not using the fructose produced by $\Delta GH91$ mutant during growth on inulin.

Media conditioned by partial growth of *B. ovatus*, *Bi. adolescentis* and a co-culture of the two bacterial on inulin was used to grow *B. ovatus* and *Bi. adolescentis* to investigate presence of antimicrobial compounds in the conditioned media (Figure 5.11). There was no obvious growth inhibition for either bacterium on the conditioned media, regardless of heat treatment, indicating the growth inhibition observed is not due to a direct killing mechanism.

Another explanation for the lack of *B. ovatus* growth on inulin is perhaps the simplest, *Bi. adolescentis* out competes *B. ovatus* on inulin. *Bi. adolescentis* glycan utilisation systems target oligosaccharides with ESBPs specific for the target oligosaccharide (Ejby *et al.*, 2013), whereas *B. ovatus* target polysaccharides (Martens *et al.*, 2011). Inulin is considered a polysaccharides, despite being relatively short when compared to other polysaccharides (Mensink *et al.*, 2015), which may be how *Bi. adolescentis* and $\Delta GH91$ are able to utilise inulin without a surface endo-acting inulinase (Figure 5.3). During growth on inulin, *B. ovatus* generates high concentrations of oligosaccharide which are utilised by *Bi. adolescentis* in the co-culture. At the same time, *Bi. adolescentis* is capable of utilising inulin giving it a competitive advantage over *B. ovatus* which is generating oligosaccharides that are being utilised preferentially by *Bi. adolescentis* due to the high affinity ESBP binding the oligosaccharide and trafficking glycan towards the associated ABC transporter. While *B. ovatus* is capable of utilising inulin without degradation at the cell surface, the action of the outer membrane GH91 inulinase allows access to the longest fraction of inulin, as evidenced by high molecular weight components of the polysaccharide remaining at stationary phase of $\Delta GH91$ cultures (Figure 5.4b). This suggests *B. ovatus* targets long chain inulin or even insoluble inulin which is usually discarded during inulin purification. If *B. ovatus* does indeed utilise long/insoluble inulin it may also generate oligosaccharides, hence opening up inaccessible material for FOS/short chain inulin users. This effect is shown clearly in the *B. ovatus* cross-feeding with *B. vulgatus* or *Bi. longum*,

where the former bacterium provides oligosaccharides for the latter two in the form of inulin breakdown products in the growth supernatant (Figure 5.9 & 5.10). Here *B. vulgatus* and *Bi. adolescentis* are completely reliant on *B. ovatus* for glycans as both lack the ability to breakdown inulin but are able to utilise FOS-I. In the gut the *B. ovatus* may provide the same function in the presence of resistant/insoluble inulin, providing other members of the gut microbiota with accessible FOS. This has been previously shown to occur when *Ruminococcus bromii* utilises resistant starch accessible products are released and used by other bacteria in the same culture (Ze *et al.*, 2012; 2015).

5.3.3 Levan Cross-feeding

Despite being composed of fructose, levan has not been labelled as a probiotic compound in the same way as inulin, and has not undergone the same level of investigation into the positive effects of utilisation by gut bacteria. During utilisation of levan, *B. thetaiotaomicron* releases relatively few oligosaccharides into the medium compared to *B. ovatus* inulin growth supernatant (Figure 5.4). The recipient bacterium used here, *Bi. adolescentis*, is capable of utilising levan oligosaccharides but is unable to use undigested levan (Figure 5.3). *Bi. adolescentis* grows in the co-growth, only introducing a slight drop in *B. thetaiotaomicron* cell density from the mono-culture (Figure 5.10). These data demonstrate the probiotic potential of levan and levan oligosaccharides in the gut. The altruistic surface degradation of levan by *B. thetaiotaomicron* allows for cross-feeding of oligosaccharides generated from action of the cell surface levanase, BT_1760 (Sonnenburg *et al.*, 2010). *Bacteroides spp.* have been shown to secrete enzymes in outer membrane vesicles (Elhenawy *et al.*, 2014). The vesicles are produced through budding of the outer membrane trapping periplasmic enzymes in the vesicles along with surface attached enzymes on the outer surface of the vesicles (Elhenawy *et al.*, 2014). These enzymes are able to degrade polysaccharides away from the *Bacteroides* cell surface, releasing oligosaccharide products that are not immediately available to the binding and transport

mechanisms required to sequester oligosaccharide products. BT_1760 is present in these vesicles (Elhenawy *et al.*, 2014) and could contribute to cross-feeding of levan oligosaccharides.

5.4 Conclusion

The probiotic nature of fructans and FOS is well described and their effect on specific bacteria and bacteria within a mixture of other gut microbes. Until recently, however, cross-feeding of glycans had not been explored. Recent published data has shown release of oligosaccharide during glycan utilisation which become available to bacteria occupying the same niche (Rakoff-Nahoum *et al.*, 2014). Data presented here corroborate these conclusions, and demonstrate cross-feeding between *Bacteroides* and *Bifidobacterium* on both inulin and levan type fructans. Interestingly, when co-cultured with a second inulin user *B. ovatus* is outcompeted, despite being able to assess a wider range of substrate than *Bi. adolescentis*. *B. thetaiotaomicron* is capable of supporting *Bi. adolescentis* in co-culture on levan, despite producing relatively few oligosaccharides. Use of Levan oligosaccharides as probiotic supplements has not been explored to the same extent as FOS-I. Data presented here shows the bifidogenic potential of levan through cross-feeding with *B. thetaiotaomicron* and potential for slow release of FOS-I by supplementation of food with inulin rather than FOS-I, through the action of *B. ovatus*.

5.5 Future work

Utilisation profiles show *Bi. longum* but not *Bi. adolescentis* is able to grow on fructose (Figure 5.3). The supernatant of *B. thetaiotaomicron* cultures grown on levan shows fructose in abundance (Figure 5.4). These data indicate *Bi. longum* may show greater growth during co-culture with *B. thetaiotaomicron* on levan than was found for *Bi. adolescentis*. If this hypothesis is correct it would demonstrate that levan oligosaccharide cross-feeding is not just limited to *Bi. adolescentis* but would include other *Bifidobacterium spp*, thus showing evidence of wider bifidogenic effects of these oligosaccharides.

The scope of this study did not include biochemical characterisation of the proteins encoded by the *Bi. adolescentis* and *Bi. longum* putative fructan utilisation loci. To confirm these loci are indeed responsible for fructan/FOS utilisation biochemical characterisation of recombinant enzymes and binding proteins would be required. Components of one *Bi. adolescentis* locus were already characterised by Sarah Shapiro (Shapiro 2015) although the second remains uncharacterised.

Generating genomic mutations in the genes suspected to confer the ability of *Bi. adolescentis* and *Bi. longum* would give valuable insight into the mechanisms *Bifidobacteria* uses to breakdown, capture and transport fructans. This is a problem however as there is not robust method of generating *Bifidobacterium* mutations. There are methods that have had very limited success. Currently there is a library of transposon (Tn5) mutations for *Bifidobacterium breve* (Ruiz *et al.* 2013), however this method is not targeted mutagenesis and would not allow probing of specific components of a glycan utilisation system. A method for site specific genetic insertion mutation method has been established, although this method may be species specific (O'Connell Motherway *et al.*, 2009). The mutation methodology includes protecting the transformed plasmid carrying the desired mutation from DNA modification systems of *Bi. breve* (O'Connell Motherway *et al.*, 2009). Development of a reliable targeted genomic mutation protocol would be an invaluable tool in studying the glycan utilisation systems of *Bi. adolescentis* and *Bi. longum*.

Chapter 6: Final Discussion

Bacteroides species have a general capacity to utilise complex carbohydrates in the human gut by employing conserved glycan utilisation systems based around a SusCD homologue complex that imports the target polymer into the periplasm (Martens *et al.*, 2009). These systems can possess multiple SusCD homologues, as shown in the xylan, arabinan and RGI polysaccharide utilisation systems (discussed in Chapter 3 and 4). This may reflect the structural variations within the same class of glycan allowing different structures to be degraded and transported using the same apparatus. The Large xylan PUL, for example, targets both the relatively simple wheat arabinoxylan (WX) and the complex corn arabinoxylan (CX), with CX being broken down into large oligosaccharides with complex glucuronoarabinoxylan structures. In contrast, the surface GH10 xylanase, BACOVA_04390, which is encoded by the Small xylan PUL, is able to make relatively frequent cuts in the WX backbone generating shorter oligosaccharides with arabinosyl substitutions. Each glycan utilisation system incorporates distinct SusCD homologue pairs with divergent sequences at the predicted ligand binding sites and it is reasonable to speculate that these complexes are highly specific recognising distinct structures even within the same class of glycan. This theory was tested in the arabinan utilisation system (Chapter 4) where the two SusC homologues were sequentially deleted from the genome. Surprisingly, branched and unbranched arabinan structures were utilised in both mutants, although to differing degrees. Arabinan with oligomeric arabinose sidechains have been previously described and purified from Quinoa seeds (Wefers *et al.*, 2014). Similar structures may be targeted by one of the two arabinan SusCD homologue complexes, which unfortunately was not tested here due to lack of substrate. This reflects the structural diversity of glycans within a single class due to sidechains extending from a conserved backbone, which impacts upon the complexity of the utilisation apparatus required to recognise and import these structures.

The xylan utilisation strategies employed by *Bacteroides. ovatus* (discussed in Chapter 3) are not invariant among the *Bacteroides* species in the human gut that use this hemicellulosic polysaccharide as a growth substrate. *B. xylanisolvens* XB1A, capable of utilising both simple and complex xylans (Despres *et al.*, 2016), possesses a similar two PUL xylan utilisation system, however, components shown to be key in the degradation of complex glucuronoarabinoxylans (GAXs) by *B. ovatus* are absent from the large PUL. PUL organisation is relatively well conserved between *B. xylanisolvens* and *B. ovatus* except for the *B. ovatus* genetic cluster encoding the enzymes belonging to families GH30, GH98 and GH115, which is replaced with a single gene for a GH5 subfamily 21 enzyme in *B. xylanisolvens*. The GH5 enzyme has been shown to display activity on CX similar to that of the GH98 enzyme, BACOVA_03433, while also being able to degrade WX into arabinoxylooligosaccharides (Despres *et al.*, 2016). This implies both *B. ovatus* and *B. xylanisolvens* large xylan PULs confer the ability to utilise complex GAXs without requiring extensive cell surface debranching prior to import. Interestingly incorporation of the GH98 xylanase gene into the large xylan PUL is rare, present in only five of 260 *Bacteroides* genomes analysed (Rogowski *et al.*, 2015). Unlike the xylan PULs of *B. ovatus* that act independently of one another to target either simple or complex xylans, the corresponding PULs of *B. xylanisolvens* XB1A are linked at the transcriptional level requiring both loci for growth on xylans (Despres *et al.*, 2016). Interestingly, cooperation of these PULs appears to extend to the proteomic level in which the large PUL encodes mostly surface localised GH enzymes, glycan binding proteins (including SusD homologues) and transporter proteins while the small PUL encodes enzymes involved in periplasmic glycan degradation; again this organisation greatly contrasts with the independent PUL organisation of *B. ovatus* (Despres *et al.*, 2016). This cooperation of xylan PULs of *B. xylanisolvens* XB1A, to my knowledge is highly unusual, and explains why regulation of the PULs are so tightly interlinked.

A common theme throughout each of the results chapters of this thesis is the importance of glycan cross-feeding among members of the gut microbiota. Chapter 3 presents data showing cross-feeding of xylooligosaccharides and arabinoxylooligosaccharides between *B. ovatus* and *Bifidobacterium adolescentis*. *Bi. adolescentis* was unable to utilise the highly complex CX-derived oligosaccharides, and thus fails to benefit from the depolymerisation of the xylan by *B. ovatus*. In cross-feeding experiments a mutant of *B. ovatus* lacking the essential GH98 surface xylanase was used in co-culture in place of *Bi. adolescentis* to represent a potential CX oligosaccharide utilising bacterium. The data clearly showed that *B. ovatus* was able to make CX-derived oligosaccharides available to other organisms in the human gut microbiota. This strategy was also used to demonstrate cross-feeding during growth of *B. thetaiotaomicron* on pectins described in Chapter 4. Again the potential recipient of glycan products was a mutant lacking an oligosaccharide generating surface enzyme which was required for growth on the cognate pectin. In Chapter 5 fructan cross-feeding was demonstrated with the highly studied prebiotic inulin and, to a lesser extent, levan. The degree to which cross-feeding occurs during *Bacteroides* utilisation of polysaccharides implies this is a common feature of the glycan utilisation systems of this genus, with the selfish utilisation of yeast mannan (Cuskin *et al.*, 2015) representing a rare event. The ubiquity of oligosaccharide release and the occurrence of cross-feeding during growth on the glycans tested implies that this sharing of nutrients is a positive selection pressure, which, if correct, suggests that there must be some benefit to the donor *Bacteroides spp.* It is possible cross-feeding is used to promote species diversity within the gut microbiota, enriching for bacteria unable to utilise polysaccharides, but which possess oligosaccharide utilisation systems.

Polysaccharides in the diet undergo mechanical mastication in the mouth, solubilisation and digestion by amylases in saliva before acid treatment and churning in the stomach. Although these processes can work to break apart the plant cell wall some structures enter the gut relatively intact, and the cytoplasmic storage glycans are inaccessible to the oligosaccharide utilising bacterium. Thus, species like *R. bromii* (Ze *et al.*, 2012), *B. thetaiotaomicron* and *B. ovatus*, that possess surface

(Sonnenburg *et al.*, 2010; Rogowski *et al.*, 2015) and secreted enzymes (Elhenawy *et al.*, 2014), are required to open up these structures, enabling oligosaccharide release. Presence of glycan utilisation systems which specifically target oligosaccharides over polysaccharides in the gut imply co-evolution with other species that are the donor organisms in cross-feeding processes.

Many studies have shown reduction of gut microbe abundance and diversity during antibiotic intervention, leading to potential *Clostridium difficile* infection, among other disorders (Rodriguez *et al.*, 2015). *C. difficile* infection is typified by frequent diarrhoea, which leads to decreased GI-transit time for dietary fibre and increased likelihood of the evacuation of gut microbes like *Bacteroides* (Rodriguez *et al.*, 2015). This is an extreme example, but *Bacteroides* oligosaccharide release may play a maintenance role in promoting bacterial diversity in the gut, hence avoiding even minor forms of dysbiosis. On the other hand, oligosaccharide release during glycan utilisation may just simply reflect an imbalance in oligosaccharide production at the cell surface and import which is widespread among *Bacteroides spp.*

Each *Bacteroides* annotated genome shows differences in predicted PUL content and hence different potential polysaccharide utilisation profiles. Chapter 3 describes the degradation and utilisation of xylans a polysaccharide found within the hemicellulose fraction of the plant cell wall. Similarly, Larsbrink *et al.* (2013) describe utilisation of xyloglucan, another hemicelluloitic polysaccharide, by *B. ovatus*. Interestingly, *B. thetaiotaomicron*, despite possessing upwards of 88 PULs (Martens *et al.*, 2009), is unable to utilise any hemicellulose glycan (Martens *et al.*, 2011). In contrast, *B. ovatus* was found to utilise all hemicellulose polysaccharides, with the exception of lichenin and laminarin (Martens *et al.*, 2011). In Chapter 5 growth of *B. ovatus* and *B. thetaiotaomicron* on fructans were tested. Here, and in the Martens *et al.* (2011) study, *B. thetaiotaomicron* was able to grow on levan but not on inulin, where the opposite was found to hold true for *B. ovatus*. Such differences in polysaccharide utilisation, even between relatively similar substrates, implies *B. thetaiotaomicron* and *B. ovatus* occupy different niches within the human gut. Lack of glycan utilisation systems

targeting hemicellulose polysaccharides may indicate *B. thetaiotaomicron* occupies an area of the gut in which little of this substrate is available. Inulin is a plant glycan entering the gut as part of the host diet (French, 1998). While levan can be also be a derived from plant material it is more commonly found in extracellular bacterial capsules of certain symbiotic species (Han, 1990). It should also be noted that *B. thetaiotaomicron* shows better growth on host-derived glycans such as mucin and chondroitin sulphate than *B. ovatus* (Martens *et al.*, 2011), implying a much tighter association of the former with the mucus layer of the gut than the latter species. This tighter mucus association may allow *B. thetaiotaomicron* access to capsule producing bacteria trapped within the mucus layer of the gut. While the preference for glycans more likely to be in the lumen of the gut suggests *B. ovatus* may have a looser association with the mucus layer or gut wall allowing it to occupy the lumen where the preferred substrates are more likely to be present. Interestingly, there is considerable overlap in the pectic polysaccharide utilisation loci of both *Bacteroides* discussed here. Each has the potential to utilise all available pectic polysaccharides, indicating pectins are prevalent in each species preferred niche. This may be due to the gel like quality of pectin, a thicker viscous glycan mixture which fulfils a space filling role dictating plant cell wall flexibility (Mohnen, 2008). These physical qualities along with the tight association of pectin with the recalcitrant structural polysaccharides (Mohnen, 2008) could allow pectin to be present in multiple niches in the gut, hence available to *B. ovatus* and *B. thetaiotaomicron*.

Final Conclusions

Dietary glycans offer a great challenge gut bacteria must overcome to access fermentable substrate to yield energy for growth. One such strategy for overcoming this challenge is offered by the *Bacteroides* glycan utilisation system, typified by the presence of SusCD-homologue complexes for glycan binding and transport and CAZymes to complete degradation of the target glycan. These systems, while sufficient for growth, are subject to product loss. These products can then be utilised by other members of the gut microbiota, which have developed specific mechanisms of oligosaccharide scavenging. In the case of *Bifidobacterium spp*, an extracellular solute binding protein with great affinity for oligosaccharides is employed to harvest short glycans from the gut lumen. In this thesis several rationales have been suggested for cross-feeding interactions, however, without studies involving cross-feeding in a simplified gut microbiota *in vivo*, these concepts remain speculation. Cross-feeding may be exploited by the food industry to produce slow-acting, longer term prebiotic treatments by administering the polysaccharide rather than oligosaccharide (eg. Inulin/Levan rather than FOS) enabling, for example, *B. ovatus*/*B. thetaiotaomicron* to generate FOS during utilisation of Inulin/Levan, causing slower FOS release giving longer acting prebiotics.

Several *Bacteroides* glycan utilisation systems have been characterised to date, a list to which this work adds, each presenting subtle variations upon the generic glycan utilisation system to overcome specific challenges present by the target glycan. Further studies of these systems may still yield interesting mechanisms of glycan degradation and utilisation.

References

- Aachary, A.A. and Prapulla, S.G. (2011) 'Xylooligosaccharides (XOS) as an Emerging Prebiotic: Microbial Synthesis, Utilization, Structural Characterization, Bioactive Properties, and Applications.', *Comprehensive Reviews in Food Science and Food Safety*, 10, pp. 2-16.
- Abrams, S.A., Griffin, I.J., Hawthorne, K.M., Liang, L., Gunn, S.K., Darlington, G. and Ellis, K.J. (2005) 'A combination of prebiotic short- and long-chain inulin-type fructans enhances calcium absorption and bone mineralization in young adolescents', *Am J Clin Nutr*, 82(2), pp. 471-6.
- Agger, J., Vikso-Nielsen, A. and Meyer, A.S. (2010) 'Enzymatic xylose release from pretreated corn bran arabinoxylan: differential effects of deacetylation and deferuloylation on insoluble and soluble substrate fractions', *J Agric Food Chem*, 58(10), pp. 6141-8.
- Agirre, J., Ariza, A., Offen, W.A., Turkenburg, J.P., Roberts, S.M., McNicholas, S., Harris, P.V., McBrayer, B., Dohnalek, J., Cowtan, K.D., Davies, G.J. and Wilson, K.S. (2016) 'Three-dimensional structures of two heavily N-glycosylated *Aspergillus sp.* family GH3 β -d-glucosidases.' *Acta Crystallogr D Struct Biol.*, 72(Pt2), pp. 254-265.
- Akalın, A.S. and Erişir, D. (2008) 'Effects of Inulin and Oligofructose on the Rheological Characteristics and Probiotic Culture Survival in Low-Fat Probiotic Ice Cream', *Journal of food science*, 73(4), pp. M184-M188.
- Allerdings, E., Ralph, J., Steinhart, H. and Bunzel, M. (2006) 'Isolation and structural identification of complex feruloylated heteroxylan side-chains from maize bran', *Phytochemistry*, 67(12), pp. 1276-86.
- Al-Tamimi, M.A., Palframan, R.J., Cooper, J.M., Gibson, G.R. and Rastall, R.A. (2006) 'In vitro fermentation of sugar beet arabinan and arabino-oligosaccharides by the human gut microflora', *J Appl Microbiol*, 100(2), pp. 407-14.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Miller, W. and Lipman, D.J. (1997) 'Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.' *Nucleic Acids Res.* 25(17), pp. 3389-3402
- Amaretti, A., Bernardi, T., Leonardi, A., Raimondi, S., Zanoni, S. and Rossi, M. (2013) 'Fermentation of xylo-oligosaccharides by *Bifidobacterium adolescentis* DSMZ 18350: kinetics, metabolism, and beta-xylosidase activities', *Appl Microbiol Biotechnol*, 97(7), pp. 3109-17.
- An, H.M., Lee do, K., Kim, J.R., Lee, S.W., Cha, M.K., Lee, K.O. and Ha, N.J. (2012) 'Antiviral activity of *Bifidobacterium adolescentis* SPM 0214 against herpes simplex virus type 1', *Arch Pharm Res*, 35(9), pp. 1665-71.
- Anderson, K.M., Ashida, H., Maskos, K., Dell, A., Li, S.-C. and Li, Y.-T. (2005) 'A Clostridial Endo- β -galactosidase That Cleaves Both Blood Group A and B Glycotopes: THE FIRST MEMBER OF A NEW GLYCOSIDE HYDROLASE FAMILY, GH98', *Journal of Biological Chemistry*, 280(9), pp. 7720-7728.
- Arora, T., Sharma, R. and Frost, G. (2011) 'Propionate. Anti-obesity and satiety enhancing factor?', *Appetite*, 56(2), pp. 511-5.

- Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T., Batto, J.M., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L., Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., Kurokawa, K., Leclerc, M., Levenez, F., Manichanh, C., Nielsen, H.B., Nielsen, T., Pons, N., Poulain, J., Qin, J., Sicheritz-Ponten, T., Tims, S., Torrents, D., Ugarte, E., Zoetendal, E.G., Wang, J., Guarner, F., Pedersen, O., de Vos, W.M., Brunak, S., Doré, J.; MetaHIT Consortium, Antolín, M., Artiguenave, F., Blottiere, H.M., Almeida, M., Brechot, C., Cara, C., Chervaux, C., Cultrone, A., Delorme, C., Denariáz, G., Dervyn, R., Foerstner, K.U., Friss, C., van de Guchte, M., Guedon, E., Haimet, F., Huber, W., van Hylckama-Vlieg, J., Jamet, A., Juste, C., Kaci, G., Knol, J., Lakhdari, O., Layec, S., Le Roux, K., Maguin, E., Mérieux, A., Melo Minardi, R., M'rini, C., Muller, J., Oozeer, R., Parkhill, J., Renault, P., Rescigno, M., Sanchez, N., Sunagawa, S., Torrejon, A., Turner, K., Vandemeulebrouck, G., Varela, E., Winogradsky, Y., Zeller, G., Weissenbach, J., Ehrlich, S.D. and Bork, P. (2011) 'Enterotypes of the human gut microbiome.' *Nature*, 473(7346), pp. 174-180
- Aspeborg, H., Coutinho, P.M., Wang, Y., Brummer, H. and Henrissat, B. (2012) 'Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5)' *BMC Evol Biol.*, 186(12).
- Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A. and Gordon, J.I. (2005) 'Host-bacterial mutualism in the human intestine', *Science*, 307(5717), pp. 1915-20.
- Bakker-Zierikzee, A.M., Alles, M.S., Knol, J., Kok, F.J., Tolboom, J.J. and Bindels, J.G. (2005) 'Effects of infant formula containing a mixture of galacto- and fructo-oligosaccharides or viable *Bifidobacterium animalis* on the intestinal microflora during the first 4 months of life', *Br J Nutr*, 94(5), pp. 783-90.
- Bansil, R. and Turner, B. (2006) 'Mucin structure, aggregation, physiological functions and biomedical applications', *Current Opinion in Colloid & Interface Science*, 11(2-3), pp. 164-170.
- Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.P., Clement, C., Ouhdouch, Y., van Wezel, G.P. (2016) 'Taxonomy, Physiology and Natural Products of *Actinobacteria*' *Microbiol Mol Biol Rev.* 80(1) pp. 1-43.
- Beguin, P. and Aubert, J.P. (1994) 'The biological degradation of cellulose', *FEMS Microbiol Rev*, 13(1), pp. 25-58.
- Ben David, Y., Dassa, B., Borovok, I., Lamed, R., Koropatkin, N.M., Martens, E.C., White, B.A., Bernalier-Bonadille, A., Duncan, S.H., Flint, H.J., Bayer, E.A. and Morais, S. (2015) 'Ruminococcal cellulosome systems from rumen to human.' *Environ Microbiol.* 17(9), pp. 3407-3428
- Bendtsen, J.D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) 'Improved prediction of signal peptides: SignalP 3.0.' *J Mol Biol.* 340(4), pp. 783-795
- Benigar, E., Dogsa, I., Stopar, D., Jamnik, A., Cigić, I.K. and Tomšič, M. (2014) 'Structure and dynamics of a polysaccharide matrix: aqueous solutions of bacterial levan', *Langmuir*, 30(14), pp. 4172-4182.
- Benus, R.F., van der Werf, T.S., Welling, G.W., Judd, P.A., Taylor, M.A., Harmsen, H.J. and Whelan, K. (2010) 'Association between *Faecalibacterium prausnitzii* and dietary fibre in colonic fermentation in healthy human subjects', *Br J Nutr*, 104(5), pp. 693-700.
- Berg JM, Tymoczko JL and L, S. (2002) *Biochemistry*. New York: W H Freeman.
- Blake, J.D., Clarke, M.L., Jansson, P.E. and McNeil, K.E. (1982) 'Fructan from *Erwinia herbicola*', *J Bacteriol*, 151(3), pp. 1595-7.
- Braithwaite, K.L., Barna, T., Spurway, T.D., Charnock, S.J., Black, G.W., Hughes, N., Lakey, J.H., Virden, R., Hazlewood, G.P., Henrissat, B. and Gilbert, H.J. (1997) 'Evidence that galactanase A from

Pseudomonas fluorescens subspecies *cellulosa* is a retaining family 53 glycosyl hydrolase in which E161 and E270 are the catalytic residues', *Biochemistry*, 36(49), pp. 15489-500.

Brett, C. and K., W. (1990) *Physiology and Biochemistry of Plant Cell Walls*. Springer Netherlands.

Brinkworth, G.D., Noakes, M., Clifton, P.M. and Buckley, J.D. (2009) 'Effects of a low carbohydrate weight loss diet on exercise capacity and tolerance in obese subjects', *Obesity (Silver Spring)*, 17(10), pp. 1916-23.

Buchanan, S.K., Smith, B.S., Venkatramani, L., Xia, D., Esser, L., Palnitkar, M., Chakraborty, R., van der Helm, D. and Deisenhofer, J. (1999) 'Crystal structure of the outer membrane active transporter FepA from *Escherichia coli*', *Nat Struct Biol*, 6(1), pp. 56-63.

Buriti, F.C.A. and Saad, S.M.I. (2014) 'Chilled milk-based desserts as emerging probiotic and prebiotic products', *Critical reviews in food science and nutrition*, 54(2), pp. 139-150.

Bush, M.S., Marry, M., Huxham, I.M., Jarvis, M.C. and McCann, M.C. (2001) 'Developmental regulation of pectic epitopes during potato tuberisation', *Planta*, 213(6), pp. 869-80.

Cani, P.D., Neyrinck, A.M., Fava, F., Knauf, C., Burcelin, R.G., Tuohy, K.M., Gibson, G.R. and Delzenne, N.M. (2007) 'Selective increases of bifidobacteria in gut microflora improve high-fat-diet-induced diabetes in mice through a mechanism associated with endotoxaemia', *Diabetologia*, 50(11), pp. 2374-2383.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V. and Henrissat, B. (2009) 'The Carbohydrate-Active EnZymes database (CAZY): an expert resource for Glycogenomics', *Nucleic Acids Res*, 37(Database issue), pp. D233-8.

Cao, Y., Rocha, E.R. and Smith, C.J. (2014) 'Efficient utilization of complex N-linked glycans is a selective advantage for *Bacteroides fragilis* in extraintestinal infections', *Proc Natl Acad Sci U S A*, 111(35), pp. 12901-6.

Carpita, N., McCann, M. and Griffing, L.R. (1997) 'Back to the Walls', *Plant Cell*, 9(3), pp. 281-282.

Cartmell, A., McKee, L.S., Pena, M.J., Larsbrink, J., Brumer, H., Kaneko, S., Ichinose, H., Lewis, R.J., Vikso-Nielsen, A., Gilbert, H.J. and Marles-Wright, J. (2011) 'The structure and function of an arabinan-specific alpha-1,2-arabinofuranosidase identified from screening the activities of bacterial GH43 glycoside hydrolases', *J Biol Chem*, 286(17), pp. 15483-95.

Cerdeno-Tarraga, A.M., Patrick, S., Crossman, L.C., Blakely, G., Abratt, V., Lennard, N., Poxton, I., Duerdent, B., Harris, B., Quail M.A., Barron, A., Clark, L., Corton, C., Doggett, J., Holden, M.T., Larke, N., Line, A., Lord, A., Norbertczak, H., Ormond, D., Price, C., Rabinowitsch, E., Woodward, J., Barrell, B., Parkhill, J. (2005) 'Extensive DNA inversions in the *B. fragilis* genome control variable gene expression' *Science*, 307(5714), pp. 1463-1465.

Chapla, D., Pandit, P. and Shah, A. (2012) 'Production of xylooligosaccharides from corncob xylan by fungal xylanase and their utilization by probiotics', *Bioresour Technol*, 115, pp. 215-21.

Cho, K.H., Cho, D., Wang, G.R. and Salyers, A.A. (2001) 'New regulatory gene that contributes to control of *Bacteroides thetaiotaomicron* starch utilization genes', *J Bacteriol*, 183(24), pp. 7198-205.

Christensen, E.G., Licht, T.R., Leser, T.D. and Bahl, M.I. (2014) 'Dietary xylo-oligosaccharide stimulates intestinal *bifidobacteria* and *lactobacilli* but has limited effect on intestinal integrity in rats', *BMC Res Notes*, 7, p. 660.

- Cid, M., Pedersen, H.L., Kaneko, S., Coutinho, P.M., Henrissat, B., Willats, W.G. and Boraston, A.B. (2010) 'Recognition of the helical structure of beta-1,4-galactan by a new family of carbohydrate-binding modules', *J Biol Chem*, 285(46), pp. 35999-6009.
- Cockburn, D.W. and Koropatkin, N.M. (2016) 'Polysaccharide Degradation by the Intestinal Microbiota and Its Influence on Human Health and Disease', *J Mol Biol*, 428(16), pp. 3230-52.
- Cockburn, D.W., Orlovsky, N.I., Foley, M.H., Kwiatkowski, K.J., Bahr, C.M., Maynard, M., Demeler, B. and Koropatkin, N.M. (2015) 'Molecular details of a starch utilization pathway in the human gut symbiont *Eubacterium rectale*', *Mol Microbiol*, 95(2), pp. 209-30.
- Cohen, S., Chang, A.C.Y. and Hsu L. (1972) 'Nonchromosomal Antibiotic resistance in Bacteria: Genetic Transformation of *Escherichia coli* by R-factor DNA.' *Proc Natl Acad Sci U S A*, 69(8) pp. 2110-2114
- Cosgrove, D.J. (2005) 'Growth of the plant cell wall', *Nat Rev Mol Cell Biol*, 6(11), pp. 850-61.
- Costabile, A., Kolida, S., Klinder, A., Gietl, E., Bauerlein, M., Froberg, C., Landschutze, V. and Gibson, G.R. (2010) 'A double-blind, placebo-controlled, cross-over study to establish the bifidogenic effect of a very-long-chain inulin extracted from globe artichoke (*Cynara scolymus*) in healthy human subjects', *Br J Nutr*, 104(7), pp. 1007-17.
- Cummings, J.H. and Macfarlane, G.T. (1991) 'The control and consequences of bacterial fermentation in the human colon', *J Appl Bacteriol*, 70(6), pp. 443-59.
- Cummings, J.H., Pomare, E.W., Branch, W.J., Naylor, C.P. and Macfarlane, G.T. (1987) 'Short chain fatty acids in human large intestine, portal, hepatic and venous blood', *Gut*, 28(10), pp. 1221-7.
- Cuskin, F., Lowe, E.C., Temple, M.J., Zhu, Y., Cameron, E.A., Pudlo, N.A., Porter, N.T., Urs, K., Thompson, A.J., Cartmell, A., Rogowski, A., Hamilton, B.S., Chen, R., Tolbert, T.J., Piens, K., Bracke, D., Vervecken, W., Hakki, Z., Speciale, G., Munoz-Munoz, J.L., Day, A., Pena, M.J., McLean, R., Suits, M.D., Boraston, A.B., Atherly, T., Ziemer, C.J., Williams, S.J., Davies, G.J., Abbott, D.W., Martens, E.C. and Gilbert, H.J. (2015) 'Human gut *Bacteroidetes* can utilize yeast mannan through a selfish mechanism', *Nature*, 517(7533), pp. 165-9.
- Davies, G. and Henrissat, B. (1995) 'Structures and mechanisms of glycosyl hydrolases', *Structure*, 3(9), pp. 853-9.
- de Lima, E.A., Machado, C.B., Zanphorlin, L.M., Ward, R.J., Sato, H.H. and Ruller, R. (2016) 'GH53 Endo-Beta-1,4-Galactanase from a Newly Isolated *Bacillus licheniformis* CBMAI 1609 as an Enzymatic Cocktail Supplement for Biomass Saccharification', *Appl Biochem Biotechnol*, 179(3), pp. 415-26.
- de Oliveira, C.P., da Silva, J.A. and de Siqueira-Junior, J.P. (2015) 'Nature of the antimicrobial activity of *Lactobacillus casei*, *Bifidobacterium bifidum* and *Bifidobacterium animalis* against foodborne pathogenic and spoilage microorganisms', *Nat Prod Res*, 29(22), pp. 2133-6.
- D'Elia, J.N. and Salyers, A.A. (1996) 'Contribution of a neopullulanase, a pullulanase, and an alpha-glucosidase to growth of *Bacteroides thetaiotaomicron* on starch', *J Bacteriol*, 178(24), pp. 7173-9.
- Despres, J., Forano, E., Lepercq, P., Comtet-Marre, S., Jubelin, G., Chambon, C., Yeoman, C.J., Berg-Miller, M.E., Fields, C.J., Materns, E., Terrapon, N., Henrissat, B., White, B.A. and Mosoni, P. (2016) 'Xylan degradation by the human gut *Bacteroides xylanisolvens*XB1A(T) involves two distinct gene clusters that are linked at the transcriptional level' *BMC Genomics*, 17:326

- Dharmani, P., Srivastava, V., Kisson-Singh, V. and Chadee, K. (2009) 'Role of intestinal mucins in innate host defense mechanisms against pathogens', *J Innate Immun*, 1(2), pp. 123-35.
- Donohoe, D.R., Collins, L.B., Wali, A., Bigler, R., Sun, W. and Bultman, S.J. (2012) 'The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation', *Mol Cell*, 48(4), pp. 612-26.
- Donohoe, D.R., Garge, N., Zhang, X., Sun, W., O'Connell, T.M., Bunger, M.K. and Bultman, S.J. (2011) 'The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon', *Cell Metab*, 13(5), pp. 517-26.
- Drakoularakou, A., Tzortzis, G., Rastall, R.A. and Gibson, G.R. (2009) 'A double-blind, placebo-controlled, randomized human study assessing the capacity of a novel galacto-oligosaccharide mixture in reducing travellers' diarrhoea', *Eur J Clin Nutr*, 64(2), pp. 146-152.
- Duncan, S.H., Belenguer, A., Holtrop, G., Johnstone, A.M., Flint, H.J. and Lobley, G.E. (2007) 'Reduced dietary intake of carbohydrates by obese subjects results in decreased concentrations of butyrate and butyrate-producing bacteria in feces', *Appl Environ Microbiol*, 73(4), pp. 1073-8.
- Duranti, S., Turrone, F., Lugli, G.A., Milani, C., Viappiani, A., Mangifesta, M., Gioiosa, L., Palanza, P., van Sinderen, D. and Ventura, M. (2014) 'Genomic characterization and transcriptional studies of the starch-utilizing strain *Bifidobacterium adolescentis* 22L', *Appl Environ Microbiol*, 80(19), pp. 6080-90.
- Egan, M., Motherway, M.O., Kilcoyne, M., Kane, M., Joshi, L., Ventura, M. and van Sinderen, D. (2014) 'Cross-feeding by *Bifidobacterium breve* UCC2003 during co-cultivation with *Bifidobacterium bifidum* PRL2010 in a mucin-based medium', *BMC Microbiol*, 14, p. 282.
- Ejby, M., Fredslund, F., Andersen, J.M., Vujicic Zagar, A., Rosager Henriksen, J., Andresen, T.L., Svensson, B., Slotboom, D.J. and Abou Hachem, M. (2016) 'An ATP-binding cassette transporter mediates the uptake of alpha-(1,6)-linked dietary oligosaccharides in *Bifidobacterium* and correlates with competitive growth on these substrates', *J Biol Chem*.
- Ejby, M., Fredslund, F., Vujicic-Zagar, A., Svensson, B., Slotboom, D.J. and Abou Hachem, M. (2013) 'Structural basis for arabinoxylo-oligosaccharide capture by the probiotic *Bifidobacterium animalis* subsp. *lactis* BI-04', *Mol Microbiol*, 90(5), pp. 1100-12.
- El Kaoutari, A., Armougom, F., Gordon, J.I., Raoult, D. and Henrissat, B. (2013) 'The abundance and variety of carbohydrate-active enzymes in the human gut microbiota', *Nat Rev Microbiol*, 11(7), pp. 497-504.
- Elhenawy, W., Debelyy, M.O. and Feldman, M.F. (2014) 'Preferential packing of acidic glycosidases and proteases into *Bacteroides* outer membrane vesicles', *MBio*, 5(2), pp. e00909-14.
- Engelsen, S.B., Cros, S., Mackie, W. and Perez, S. (1996) 'A molecular builder for carbohydrates: application to polysaccharides and complex carbohydrates', *Biopolymers*, 39(3), pp. 417-433.
- Falke, J.J. and Erbse, A.H. (2009) 'The piston rises again', *Structure*, 17(9), pp. 1149-51.
- Ferreira, R.B., Gill, N., Willing, B.P., Antunes, L.C., Russell, S.L., Croxen, M.A. and Finlay, B.B. (2011) 'The intestinal microbiota plays a role in *Salmonella*-induced colitis independent of pathogen colonization', *PLoS One*, 6(5), p. e20338.

- Finegold, S.M., Li, Z., Summanen, P.H., Downes, J., Thames, G., Corbett, K., Dowd, S., Krak, M. and Heber, D. (2014) 'Xylooligosaccharide increases bifidobacteria but not lactobacilli in human gut microbiota', *Food Funct*, 5(3), pp. 436-45.
- Fleischer, A., O'Neill, M.A. and Ehwald, R. (1999) 'The pore size of non-graminaceous plant cell walls is rapidly decreased by borate ester cross-linking of the pectic polysaccharide rhamnogalacturonan II', *Plant Physiology*, 121(3), pp. 829-838.
- French, A.D. (1988) 'Accessible conformations of the β -d-(2 \rightarrow 1)-and-(2 \rightarrow 6)-linked d-fructans inulin and levan', *Carbohydrate research*, 176(1), pp. 17-29.
- Frost, G., Sleeth, M.L., Sahuri-Arisoylu, M., Lizarbe, B., Cerdan, S., Brody, L., Anastasovska, J., Ghourab, S., Hankir, M., Zhang, S., Carling, D., Swann, J.R., Gibson, G., Viardot, A., Morrison, D., Louise Thomas, E. and Bell, J.D. (2014) 'The short-chain fatty acid acetate reduces appetite via a central homeostatic mechanism', *Nat Commun*, 5, p. 3611.
- Fuchs, A. (1987) 'Potentials for Non-Food Utilization of Fructose and Inulin', *Starch-Stärke*, 39(10), pp. 335-343.
- Fukuda, S., Toh, H., Hase, K., Oshima, K., Nakanishi, Y., Yoshimura, K., Tobe, T., Clarke, J.M., Topping, D.L., Suzuki, T., Taylor, T.D., Itoh, K., Kikuchi, J., Morita, H., Hattori, M. and Ohno, H. (2011) 'Bifidobacteria can protect from enteropathogenic infection through production of acetate', *Nature*, 469(7331), pp. 543-7.
- Furuse, S.U., Ohse, T., Jo-Watanabe, A., Shigehisa, A., Kawakami, K., Matsuki, T., Chonan, O. and Nangaku, M. (2014) 'Galacto-oligosaccharides attenuate renal injury with microbiota modification', *Physiol Rep*, 2(7).
- Gibson, G.R. and Roberfroid, M.B. (1995) 'Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics', *J Nutr*, 125(6), pp. 1401-12.
- Gibson, G.R., Probert, H.M., Loo, J.V., Rastall, R.A. and Roberfroid, M.B. (2004) 'Dietary modulation of the human colonic microbiota: updating the concept of prebiotics', *Nutr Res Rev*, 17(2), pp. 259-75.
- Gibson, L.J. (2012) 'The hierarchical structure and mechanics of plant materials.' *J R Soc Interface*. 9(76), pp. 2749-2766.
- Gilmore, M.S. and Ferretti, J.J. (2003) 'The thin line between gut commensal and pathogen', *Science*, 299(5615), pp. 1999-2002.
- Gu, S., Chen, D., Zhang, J.-N., Lv, X., Wang, K., Duan, L.-P., Nie, Y. and Wu, X.-L. (2013) 'Bacterial community mapping of the mouse gastrointestinal tract', *PLoS One*, 8(10), p. e74957.
- Han, Y.W. (1990) 'Microbial levan', *Adv Appl Microbiol*, 35, pp. 171-94.
- Hara, H., Onoshima, S. and Nakagawa, C. (2010) 'Difructose anhydride III promotes iron absorption in the rat large intestine', *Nutrition*, 26(1), pp. 120-7.
- Hendry, G.A.F. (1993) 'Evolutionary origins and natural functions of fructans—a climatological, biogeographic and mechanistic appraisal', *New phytologist*, 123(1), pp. 3-14.

- Hernandez, A., Lopez, J.C., Santamaria, R., Diaz, M., Fernandez-Abalos, J.M., Copa-Patino, J.L. and Soliveri, J. (2008) 'Xylan-binding xylanase Xyl30 from *Streptomyces avermitilis*: cloning, characterization, and overproduction in solid-state fermentation', *Int Microbiol*, 11(2), pp. 133-41.
- Herrero, M., de Lorenzo, V. and Timmis, K.N. (1990) 'Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria.' *J Bacteriol.* 172(11), pp. 6557-6567
- Higgins, M.A., Whitworth, G.E., El Warry, N., Randriantsoa, M., Samain, E., Burke, R.D., Vocadlo, D.J. and Boraston, A.B. (2009) 'Differential recognition and hydrolysis of host carbohydrate antigens by *Streptococcus pneumoniae* family 98 glycoside hydrolases', *J Biol Chem*, 284(38), pp. 26161-73.
- Higuchi, M., Iwami, Y., Yamada, T. and Araya, S. (1970) 'Levan synthesis and accumulation by human dental plaque', *Archives of oral biology*, 15(6), pp. 563-567.
- Hinz, S.W., Pastink, M.I., van den Broek, L.A., Vincken, J.P. and Voragen, A.G. (2005) '*Bifidobacterium longum* endogalactanase liberates galactotriose from type I galactans', *Appl Environ Microbiol*, 71(9), pp. 5501-10.
- Hoeppli, R.E., Wu, D., Cook, L. and Levings, M.K. (2015) 'The environment of regulatory T cell biology: cytokines, metabolites, and the microbiome', *Front Immunol*, 6, p. 61.
- Hooper, L.V., Midtvedt, T. and Gordon, J.I. (2002) 'How host-microbial interactions shape the nutrient environment of the mammalian intestine', *Annual review of nutrition*, 22(1), pp. 283-307.
- Hosseini, E., Grootaert, C., Verstraete, W. and Van de Wiele, T. (2011) 'Propionate as a health-promoting microbial metabolite in the human gut', *Nutr Rev*, 69(5), pp. 245-58.
- Howarth, N.C., Saltzman, E. and Roberts, S.B. (2001) 'Dietary fiber and weight regulation.' *Nutr Rev.* 59(5), pp. 129-139
- Ishii, T. and Matsunaga, T. (1996) 'Isolation and characterization of a boron-rhamnogalacturonan-II complex from cell walls of sugar beet pulp', *Carbohydrate Research*, 284(1), pp. 1-9.
- Ishikawa, H., Matsumoto, S., Ohashi, Y., Imaoka, A., Setoyama, H., Umesaki, Y., Tanaka, R. and Otani, T. (2011) 'Beneficial effects of probiotic bifidobacterium and galacto-oligosaccharide in patients with ulcerative colitis: a randomized controlled study', *Digestion*, 84(2), pp. 128-33.
- Jarvis, M.C. and Apperley, D.C. (1995) 'Chain conformation in concentrated pectic gels: evidence from ¹³C NMR', *Carbohydrate research*, 275(1), pp. 131-145.
- Jeong, J.Y., Yim, H.S., Ryu, J.Y., Lee, H.S. Seen, D.S. and Kang, S.G. (2012) 'One-step sequence- and ligation-independent cloning as a rapid and versatile cloning method for functional genomics studies.' *Appl Environ Microbiol.* 78(15), pp. 5440-5443
- Jew, S., AbuMweis, S.S. and Jones, P.J. (2009) 'Evolution of the human diet: linking our ancestral diet to modern functional foods as a means of chronic disease prevention', *J Med Food*, 12(5), pp. 925-34.
- Johnson, L.P., Walton, G.E., Psichas, A., Frost, G.S., Gibson, G.R. and Barraclough, T.G. (2015) 'Prebiotics Modulate the Effects of Antibiotics on Gut Microbial Diversity and Functioning in Vitro', *Nutrients*, 7(6), pp. 4480-97.

- Jordan, L.D., Zhou, Y., Smallwood, C.R., Lill, Y., Ritchie, K., Yip, W.T., Newton, S.M. and Klebba, P.E. (2013) 'Energy-dependent motion of TonB in the Gram-negative bacterial inner membrane', *Proc Natl Acad Sci U S A*, 110(28), pp. 11553-8.
- Juncker, A.S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H. and Krogh, A. (2003) 'Prediction of lipoprotein signal peptides in Gram-negative bacteria', *Protein Sci*, 12(8), pp. 1652-62.
- Kaiko, G.E., Ryu, S.H., Koues, O.I., Collins, P.L., Solnica-Krezel, L., Pearce, E.J., Pearce, E.L., Oltz, E.M. and Stappenbeck, T.S. (2016) 'The Colonic Crypt Protects Stem Cells from Microbiota-Derived Metabolites', *Cell*, 165(7), pp. 1708-20.
- Kajander, K., Myllyluoma, E., Rajilic-Stojanovic, M., Kyrönpalo, S., Rasmussen, M., Jarvenpää, S., Zoetendal, E.G., de Vos, W.M., Vapaatalo, H. and Korpela, R. (2008) 'Clinical trial: multispecies probiotic supplementation alleviates the symptoms of irritable bowel syndrome and stabilizes intestinal microbiota', *Aliment Pharmacol Ther*, 27(1), pp. 48-57.
- Kamada, N., Kim, Y.G., Sham, H.P., Vallance, B.A., Puente, J.L., Martens, E.C. and Nunez, G. (2012) 'Regulated virulence controls the ability of a pathogen to compete with the gut microbiota', *Science*, 336(6086), pp. 1325-9.
- Kampmann, C., Dicksved, J., Engstrand, L. and Rautelin, H. (2016) 'Composition of human faecal microbiota in resistance to *Campylobacter* infection', *Clin Microbiol Infect*, 22(1), pp. 61.e1-8.
- Kanauchi, O., Andoh, A. and Mitsuyama, K. (2013) 'Effects of the modulation of microbiota on the gastrointestinal immune system and bowel function', *J Agric Food Chem*, 61(42), pp. 9977-83.
- Kankaanpää, P.E., Salminen, S.J., Isolauri, E. and Lee, Y.K. (2001) 'The influence of polyunsaturated fatty acids on probiotic growth and adhesion', *FEMS Microbiol Lett*, 194(2), pp. 149-53.
- Karunatilaka, K.S., Cameron, E.A., Martens, E.C., Koropatkin, N.M. and Biteen, J.S. (2014) 'Superresolution imaging captures carbohydrate utilization dynamics in the human gut symbionts.' *Mbio*. 5(6) e02172
- Kien, C.L. (1996) 'Digestion, absorption, and fermentation of carbohydrates in the newborn', *Clinics in perinatology*, 23(2), pp. 211-228.
- Kim, H.B. and Isaacson, R.E. (2015) 'The pig gut microbial diversity: understanding the pig gut microbial ecology through the next generation high throughput sequencing', *Veterinary microbiology*, 177(3), pp. 242-251.
- Knights, D., Ward, T.L., McKinlay, C.E., Miller, H., Gonzalez, A., McDonald, D. and Knight, R. (2014) 'Rethinking "enterotypes"', *Cell host & microbe*, 16(4), pp. 433-437.
- Knol, J., Scholtens, P., Kafka, C., Steenbakkers, J., Gro, S., Helm, K., Klarczyk, M., Schopfer, H., Bockler, H.M. and Wells, J. (2005) 'Colon microflora in infants fed formula with galacto- and fructo-oligosaccharides: more like breast-fed infants', *J Pediatr Gastroenterol Nutr*, 40(1), pp. 36-42.
- Koh, A., De Vadder, F., Kovatcheva-Datchary, P. and Backhed, F. (2016) 'From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites', *Cell*, 165(6), pp. 1332-45.
- Kolida, S. and Gibson, G.R. (2007) 'Prebiotic capacity of inulin-type fructans', *J Nutr*, 137(11 Suppl), pp. 2503s-2506s.

- Koropatkin, N.M., Martens, E.C., Gordon, J.I. and Smith, T.J. (2008) 'Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices', *Structure*, 16(7), pp. 1105-15.
- Koshland, D.E., Jr., Budenstein, Z. and Kowalsky, A. (1954) 'Mechanism of hydrolysis of adenosinetriphosphate catalyzed by purified muscle proteins', *J Biol Chem*, 211(1), pp. 279-87.
- Laemmli, U.K. (1970) 'Cleavage of structural proteins during the assembly of the head of bacteriophage T4.' *Nature*, 227(5259) pp. 680-685
- Lagaert, S., Pollet, A., Delcour, J.A., Lavigne, R., Courtin, C.M. and Volckaert, G. (2010) 'Substrate specificity of three recombinant alpha-L-arabinofuranosidases from *Bifidobacterium adolescentis* and their divergent action on arabinoxylan and arabinoxylan oligosaccharides', *Biochem Biophys Res Commun*, 402(4), pp. 644-50.
- Lagaert, S., Van Campenhout, S., Pollet, A., Bourgois, T.M., Delcour, J.A., Courtin, C.M. and Volckaert, G. (2007) 'Recombinant expression and characterization of a reducing-end xylose-releasing exo-oligoxylanase from *Bifidobacterium adolescentis*', *Appl Environ Microbiol*, 73(16), pp. 5374-7.
- Larsbrink, J., Rogers, T.E., Hemsworth, G.R., McKee, L.S., Tazuin, A.S., Spadiut, O., Kliner, S., Pudlo, N.A., Urs, K., Koropatkin, N.M., Creagh, A.L., Haynes, C.A., Kelly, A.G., Cederholm, S.N., Davies, G.J., Martens, E.C. and Brumer, H. (2014) 'A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes', *Nature*, 506(7489), pp. 498-502.
- Lasken, R.S. and McLean, J.S. (2014) 'Recent advances in genomic DNA sequencing of microbial species from single cells.' *Nat Rev Genet*, 15(9), pp. 577-584.
- Lau, J.M., McNeil, M., Darvill, A.G. and Albersheim, P. (1985) 'Structure of the backbone of rhamnogalacturonan I, a pectic polysaccharide in the primary cell walls of plants', *Carbohydrate research*, 137, pp. 111-125.
- Leach, J.D. and Sobolik, K.D. (2010) 'High dietary intake of prebiotic inulin-type fructans in the prehistoric Chihuahuan Desert', *Br J Nutr*, 103(11), pp. 1558-61.
- Lee, J.H. and O'Sullivan, D.J. (2010) 'Genomic insights into *bifidobacteria*', *Microbiol Mol Biol Rev*, 74(3), pp. 378-416.
- Lefranc-Millot, C., Guerin-Deremaux, L., Wils, D., Neut, C., Miller, L.E. and Saniez-Degrave, M.H. (2012) 'Impact of a resistant dextrin on intestinal ecology: how altering the digestive ecosystem with NUTRIOSE(R), a soluble fibre with prebiotic properties, may be beneficial for health', *J Int Med Res*, 40(1), pp. 211-24.
- Lehtio, J., Sugiyama, J., Gustavsson, M., Fransson, L., Linder, M. and Teeri, T.T. (2003) 'The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules', *Proc Natl Acad Sci U S A*, 100(2), pp. 484-9.
- Ley, R.E., Bäckhed, F., Turnbaugh, P., Lozupone, C.A., Knight, R.D. and Gordon, J.I. (2005) 'Obesity alters gut microbial ecology', *Proc Natl Acad Sci U S A*, 102(31), pp. 11070-11075.
- Ley, R.E., Turnbaugh, P.J., Klein, S. and Gordon, J.I. (2006) 'Microbial ecology: human gut microbes associated with obesity', *Nature*, 444(7122), pp. 1022-1023.
- Lilius, G., Persson, M., Bulow, L. and Mosbach, K. (1991) 'Metal affinity precipitation of proteins carrying genetically attached polyhistidine affinity tails', *Eur J Biochem*, 198(2), pp. 499-504.

- Lim, K.S., Huh, C.S. and Baek, Y.J. (1993) 'Antimicrobial susceptibility of bifidobacteria', *J Dairy Sci*, 76(8), pp. 2168-74.
- Liu, C., Song, Y., McTeague, M., Vu, A.W., Wexler, H. and Finegold, S.M. (2003) 'Rapid identification of the species of the *Bacteroides fragilis* group by multiplex PCR assays using group- and species-specific primers', *FEMS microbiology letters*, 222(1), pp. 9-16.
- Liu, D., Wang, S., Xu, B., Guo, Y., Zhao, J., Liu, W., Sun, Z., Shao, C., Wei, X., Jiang, Z., Wang, X., Liu, F., Wang, J., Huang, L., Hu, D., He, X., Riedel, C.U. and Yuan, J. (2011) 'Proteomics analysis of *Bifidobacterium longum* NCC2705 growing on glucose, fructose, mannose, xylose, ribose, and galactose', *Proteomics*, 11(13), pp. 2628-38.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M. and Henrissat, B. (2014) 'The carbohydrate-active enzymes database (CAZy) in 2013', *Nucleic Acids Res*, 42(Database issue), pp. D490-5.
- Lorca, G.L., Barabote, R.D., Zlotopolski, V., Tran, C., Winnen, B., Hvorup, R.N., Stonestrom, A.J., Nguyen, E., Huang, L.W., Kim, D.S. and Saier, M.H., Jr. (2007) 'Transport capabilities of eleven gram-positive bacteria: comparative genomic analyses', *Biochim Biophys Acta*, 1768(6), pp. 1342-66.
- Louis, P., Hold, G.L. and Flint, H.J. (2014) 'The gut microbiota, bacterial metabolites and colorectal cancer', *Nature reviews Microbiology*, 12(10), pp. 661-672.
- Lowe, E.C., Basle, A., Czjzek, M., Firbank, S.J. and Bolam, D.N. (2012) 'A scissor blade-like closing mechanism implicated in transmembrane signaling in a *Bacteroides* hybrid two-component system', *Proc Natl Acad Sci U S A*, 109(19), pp. 7298-303.
- Lupton, J.R. (2004) 'Microbial degradation products influence colon cancer risk: the butyrate controversy', *J Nutr*, 134(2), pp. 479-82.
- Macfarlane, G.T., Steed, H. and Macfarlane, S. (2008) 'Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics', *J Appl Microbiol*, 104(2), pp. 305-44.
- Mackie, R.I., Sghir, A. and Gaskins, H.R. (1999) 'Developmental microbial ecology of the neonatal gastrointestinal tract', *The American journal of clinical nutrition*, 69(5), pp. 1035s-1045s.
- Martens, E.C., Chiang, H.C. and Gordon, J.I. (2008) 'Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont', *Cell Host Microbe*, 4(5), pp. 447-57.
- Martens, E.C., Koropatkin, N.M., Smith, T.J. and Gordon, J.I. (2009) 'Complex glycan catabolism by the human gut microbiota: the *Bacteroidetes sus*-like paradigm', *J Biol Chem*, 284(37), pp. 24673-7.
- Martens, E.C., Lowe, E.C., Chiang, H., Pudlo, N.A., Wu, M., McNulty, N.P., Abbott, D.W., Henrissat, B., Gilbert, H.J., Bolam, D.N. and Gordon, J.I. (2011) 'Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts', *PLoS Biol*, 9(12), p. e1001221.
- Martinez, I., Kim, J., Duffy, P.R., Schlegel, V.L. and Walter, J. (2010) 'Resistant starches types 2 and 4 have differential effects on the composition of the fecal microbiota in human subjects', *PLoS One*, 5(11), p. e15046.
- Matsui, I., Ishikawa, K., Matsui, E., Miyairi, S., Fukui, S. and Honda, K. (1991) 'Subsite structure of Saccharomycopsis alpha-amylase secreted from *Saccharomyces cerevisiae*', *J Biochem*, 109(4), pp. 566-9.

- McGregor, N., Morar, M., Fenger, T.H., Stogios, P., Lenfant, N., Yin, V., Xu, X., Evdokimova, E., Cui, H., Henrissat, B., Savchenko, A. and Brumer, H. (2016) 'Structure-Function analysis of a mixed linkage β -glucanase/xyloglucanase from the key ruminal bacteroidetes *Prevotella bryantii* B(1)4' *J Biol Chem*, 291(3), pp. 1175-1195.
- McNeil, M., Darvill, A.G. and Albersheim, P. (1980) 'Structure of plant cell walls X. Rhamnogalacturonan I, a structurally complex pectic polysaccharide in the walls of suspension-cultured sycamore cells', *Plant physiology*, 66(6), pp. 1128-1134.
- Mei, G.Y., Carey, C.M., Tosh, S., Kostrzynska, M. (2011) 'Utilization of different types of dietary fibres by potential probiotics', *Can J Microbiol*, 57(10), pp. 857-865.
- Mensink, M.A., Frijlink, H.W., van der Voort Maarschalk, K. and Hinrichs, W.L. (2015) 'Inulin, a flexible oligosaccharide I: Review of its physicochemical characteristics', *Carbohydr Polym*, 130, pp. 405-19.
- Metzker, M.L. (2010) 'Sequencing technologies—the next generation', *Nature reviews genetics*, 11(1), pp. 31-46.
- Mewis, K., Lenfant, N., Lombard, V. and Henrissat, B. (2016) 'Dividing the Large Glycoside Hydrolase Family 43 into Subfamilies: a Motivation for Detailed Enzyme Characterisation.' *Appl Environ Microbiol*, 82(6), pp. 1686-1692
- Michalak, M., Thomassen, L.V., Roytio, H., Ouwehand, A.C., Meyer, A.S. and Mikkelsen, J.D. (2012) 'Expression and characterization of an endo-1,4-beta-galactanase from *Emericella nidulans* in *Pichia pastoris* for enzymatic design of potentially prebiotic oligosaccharides from potato galactans', *Enzyme Microb Technol*, 50(2), pp. 121-9.
- Miller, T.L. and Wolin, M.J. (1996) 'Pathways of acetate, propionate, and butyrate formation by the human fecal microbial flora', *Appl Environ Microbiol*, 62(5), pp. 1589-92.
- Miyazaki, K., Miyamoto, H., Mercer, D.K., Hirase, T., Martin, J.C., Kojima, Y. and Flint, H.J. (2003) 'Involvement of the multidomain regulatory protein XynR in positive control of xylanase gene expression in the ruminal anaerobe *Prevotella bryantii* B(1)4', *J Bacteriol*, 185(7), pp. 2219-26.
- Mohnen, D. (2008) 'Pectin structure and biosynthesis', *Curr Opin Plant Biol*, 11(3), pp. 266-77.
- Moreira, L.R. and Filho, E.X. (2008) 'An overview of mannan structure and mannan-degrading enzyme systems', *Appl Microbiol Biotechnol*, 79(2), pp. 165-78.
- Muda, P., Seymour, G.B., Errington, N. and Tucker, G.A. (1995) 'Compositional changes in cell wall polymers during mango fruit ripening', *Carbohydrate Polymers*, 26(4), pp. 255-260.
- Mullis, K.B. and Faloona, F.A. (1987) 'Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction.' *Methods Enzymol*. 155, pp. 335-350
- Munarin, F., Tanzi, M.C. and Petrini, P. (2012) 'Advances in biomedical applications of pectin gels' *Int J Biol Macromol*, 51(4) pp. 681-689.
- Nakamura, A., Furuta, H., Maeda, H., Takao, T. and Nagamatsu, Y. (2002) 'Structural studies by stepwise enzymatic degradation of the main backbone of soybean soluble polysaccharides consisting of galacturonan and rhamnogalacturonan', *Bioscience, biotechnology, and biochemistry*, 66(6), pp. 1301-1313.

Naumoff, D.G. (2011) 'Hierarchical classification of glycoside hydrolases', *Biochemistry (Mosc)*, 76(6), pp. 622-35.

Ndeh, D. (2013) 'Novel glycan-targeted extracellular proteases from divergent mucosal microbes' Thesis, Newcastle University.

Newgard, C.B., An, J., Bain, J.R., Muehlbauer, M.J., Stevens, R.D., Lien, L.F., Haqq, A.M., Shah, S.H., Arlotto, M., Slentz, C.A., Rochon, J., Gallup, D., Ilkayeva, O., Wenner, B.R., Yancy, W.S., Jr., Eisenson, H., Musante, G., Surwit, R.S., Millington, D.S., Butler, M.D. and Svetkey, L.P. (2009) 'A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance', *Cell Metab*, 9(4), pp. 311-26.

Nurizzo, D., Turkenburg, J.P., Charnock, S.J., Roberts, S.M., Dodson, E.J., McKie, V.A., Taylor, E.J., Gilbert, H.J. and Davies, G.J. (2002) '*Cellvibrio japonicus* α -L-arabinanase 43A has a novel five-blade β -propeller fold', *Nature Structural & Molecular Biology*, 9(9), pp. 665-668.

Øbro, J., Harholt, J., Scheller, H.V. and Orfila, C. (2004) 'Rhamnogalacturonan I in *Solanum tuberosum* tubers contains complex arabinogalactan structures', *Phytochemistry*, 65(10), pp. 1429-1438.

O'Connell Motherway, M., Fitzgerald, G.F. and van Sinderen, D. (2011) 'Metabolism of a plant derived galactose-containing polysaccharide by *Bifidobacterium breve* UCC2003', *Microb Biotechnol*, 4(3), pp. 403-16.

O'Connell Motherway, M., Fitzgerald, G.F. and van Sinderen, D. (2011) 'Metabolism of a plant derived galactose-containing polysaccharide by *Bifidobacterium breve* UCC2003', *Microb Biotechnol*, 4(3), pp. 403-16.

Oomen, R.J.F.J., Doeswijk-Voragen, C.H.L., Bush, M.S., Vincken, J.P., Borkhardt, B., Van Den Broek, L.A.M., Corsar, J., Ulvskov, P., Voragen, A.G.J. and McCann, M.C. (2002) 'In muro fragmentation of the rhamnogalacturonan I backbone in potato (*Solanum tuberosum* L.) results in a reduction and altered location of the galactan and arabinan side-chains and abnormal periderm development', *The Plant Journal*, 30(4), pp. 403-413.

Ouwehand, A.C., Salminen, S. and Isolauri, E. (2002) 'Probiotics: an overview of beneficial effects' *Antonie Van Leeuwenhoek* 82(1-4), pp. 279-289.

Pabst, M., Fischl, R.M., Brecker, L., Morelle, W., Fauland, A., Kofeler, H., Altmann, F. and Leonard, R. (2013) 'Rhamnogalacturonan II structure shows variation in the side chains monosaccharide composition and methylation status within and across different plant species', *Plant J*, 76(1), pp. 61-72.

Park, J., Kim, M., Kang, S.G., Jannasch, A.H., Cooper, B., Patterson, J. and Kim, C.H. (2015) 'Short-chain fatty acids induce both effector and regulatory T cells by suppression of histone deacetylases and regulation of the mTOR-S6K pathway', *Mucosal Immunol*, 8(1), pp. 80-93.

Pastell, H., Westermann, P., Meyer, A.S., Tuomainen, P. and Tenkanen, M. (2009) 'In vitro fermentation of arabinoxylan-derived carbohydrates by bifidobacteria and mixed fecal microbiota', *J Agric Food Chem*, 57(18), pp. 8598-606.

Patel, S., Majumder, A. and Goyal, A. (2012) 'Potentials of exopolysaccharides from lactic Acid bacteria', *Indian J Microbiol*, 52(1), pp. 3-12.

Pauly, M., Gille, S., Liu, L., Mansoori, N., de Souza, A., Schultink, A. and Xiong, G. (2013) 'Hemicellulose biosynthesis.' *Planta*. 238(4), pp. 627-642

- Pena, M.J., Zhong, R., Zhou, G.K., Richardson, E.A., O'Neill, M.A., Darvill, A.G., York, W.S. and Ye, Z.H. (2007) 'Arabidopsis irregular xylem8 and irregular xylem9: implications for the complexity of glucuronoxylan biosynthesis', *Plant Cell*, 19(2), pp. 549-63.
- Pokusaeva, K., Fitzgerald, G.F. and van Sinderen, D. (2011) 'Carbohydrate metabolism in *Bifidobacteria*', *Genes Nutr*, 6(3), pp. 285-306.
- Pokusaeva, K., Neves, A.R., Zomer, A., O'Connell-Motherway, M., MacSharry, J., Curley, P., Fitzgerald, G.F. and van Sinderen, D. (2010) 'Ribose utilization by the human commensal *Bifidobacterium breve* UCC2003', *Microb Biotechnol*, 3(3), pp. 311-23.
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F. and Yamada, T. (2010) 'A human gut microbial gene catalogue established by metagenomic sequencing', *nature*, 464(7285), pp. 59-65.
- Rakoff-Nahoum, S., Coyne, M.J. and Comstock, L.E. (2014) 'An ecological network of polysaccharide utilization among human intestinal symbionts', *Curr Biol*, 24(1), pp. 40-9.
- Rakoff-Nahoum, S., Foster, K.R. and Comstock, L.E. (2016) 'The evolution of cooperation within the gut microbiota', *Nature*, 533(7602), pp. 255-9.
- Raman, R., Venkataraman, M., Ramakrishnan, S., Lang, W., Raguram, S. and Sasisekharan, R. (2006) 'Advancing glycomics: implementation strategies at the consortium for functional glycomics', *Glycobiology*, 16(5), pp. 82r-90r.
- Ramnani, P., Gaudier, E., Bingham, M., van Bruggen, P., Tuohy, K.M. and Gibson, G.R. (2010) 'Prebiotic effect of fruit and vegetable shots containing Jerusalem artichoke inulin: a human intervention study', *Br J Nutr*, 104(2), pp. 233-40.
- Reeves, A.R., Wang, G.R. and Salyers, A.A. (1997) 'Characterization of four outer membrane proteins that play a role in utilization of starch by *Bacteroides thetaiotaomicron*', *J Bacteriol*, 179(3), pp. 643-9.
- Reichardt, N., Duncan, S.H., Young, P., Belenguer, A., McWilliam Leitch, C., Scott, K.P., Flint, H.J. and Louis, P. (2014) 'Phylogenetic distribution of three pathways for propionate production within the human gut microbiota', *Isme j*, 8(6), pp. 1323-35.
- Rey, F.E., Faith, J.J., Bain, J., Muehlbauer, M.J., Stevens, R.D., Newgard, C.B. and Gordon, J.I. (2010) 'Dissecting the in vivo metabolic potential of two human gut acetogens', *J Biol Chem*, 285(29), pp. 22082-90.
- Ridley, B.L., O'Neill, M.A. and Mohnen, D. (2001) 'Pectins: structure, biosynthesis, and oligogalacturonide-related signaling', *Phytochemistry*, 57(6), pp. 929-67.
- Riviere, A., Selak, M., Lantin, D., Leroy, F. and De Vuyst, L. (2016) '*Bifidobacteria* and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut', *Front Microbiol*, 7, p. 979.
- Roberfroid, M.B. (2005) 'Introducing inulin-type fructans', *British Journal of Nutrition*, 93(S1), pp. S13-S25.
- Rodionov, D.A. (2007) 'Comparative genomic reconstruction of transcriptional regulatory networks in bacteria', *Chem Rev*, 107(8), pp. 3467-97.

- Rodionov, D.A., Mironov, A.A. and Gelfand, M.S. (2001) 'Transcriptional regulation of pentose utilisation systems in the *Bacillus/Clostridium* group of bacteria', *FEMS Microbiol Lett*, 205(2), pp. 305-14.
- Rodriguez, C., Taminiau, B., Van Broeck, J., Delmee, M. and Daube, G. (2015) '*Clostridium difficile* infection and intestinal microbiota interactions', *Microb Pathog*, 89, pp. 201-9.
- Rodriguez-Colinas, B., Fernandez-Arrojo, L., Ballesteros, A.O. and Plou, F.J. (2014) 'Galactooligosaccharides formation during enzymatic hydrolysis of lactose: towards a prebiotic-enriched milk', *Food Chem*, 145, pp. 388-94.
- Rogers, T.E., Pudlo, N.A., Koropatkin, N.M., Bell, J.S., Moya Balasch, M., Jasker, K. and Martens, E.C. (2013) 'Dynamic responses of *Bacteroides thetaiotaomicron* during growth on glycan mixtures', *Mol Microbiol*, 88(5), pp. 876-90.
- Rogowski, A., Briggs, J.A., Mortimer, J.C., Tryfona, T., Terrapon, N., Lowe, E.C., Basle, A., Morland, C., Day, A.M., Zheng, H., Rogers, T.E., Thompson, P., Hawkins, A.R., Yadav, M.P., Henrissat, B., Martens, E.C., Dupree, P., Gilbert, H.J. and Bolam, D.N. (2015) 'Glycan complexity dictates microbial resource allocation in the large intestine', *Nat Commun*, 6, p. 7481.
- Roy, C.C., Kien, C.L., Bouthillier, L. and Levy, E. (2006) 'Short-chain fatty acids: ready for prime time?', *Nutr Clin Pract*, 21(4), pp. 351-66.
- Russell, W.R., Gratz, S.W., Duncan, S.H., Holtrop, G., Ince, J., Scobbie, L., Duncan, G., Johnstone, A.M., Lobley, G.E., Wallace, R.J., Duthie, G.G. and Flint, H.J. (2011) 'High-protein, reduced-carbohydrate weight-loss diets promote metabolite profiles likely to be detrimental to colonic health', *Am J Clin Nutr*, 93(5), pp. 1062-72.
- Ryan, S.M., Fitzgerald, G.F. and van Sinderen, D. (2005) 'Transcriptional regulation and characterization of a novel beta-fructofuranosidase-encoding gene from *Bifidobacterium breve* UCC2003', *Appl Environ Microbiol*, 71(7), pp. 3475-82.
- Saier, M.H., Jr. (2000) 'Families of transmembrane sugar transport proteins', *Mol Microbiol*, 35(4), pp. 699-710.
- Salt, W.G., Samara, M.A. and Stretton, R.J. (1985) 'The influence of incubation time and medium composition on fatty acid production by some *Bacteroides* species', *Microbios*, 42(171S), pp. 273-85.
- Salyers, A.A. (1984) '*Bacteroides* of the human lower intestinal tract', *Annual Reviews in Microbiology*, 38(1), pp. 293-313.
- Sassone-Corsi, M. and Raffatellu, M. (2015) 'No vacancy: how beneficial microbes cooperate with immunity to provide colonization resistance to pathogens', *J Immunol*, 194(9), pp. 4081-7.
- Scheller, H.V. and Ulvskov, P. (2010) 'Hemicelluloses', *Annu Rev Plant Biol*, 61, pp. 263-89.
- Scott, K.P., Martin, J.C., Duncan, S.H. and Flint, H.J. (2014) 'Prebiotic stimulation of human colonic butyrate-producing bacteria and bifidobacteria, in vitro', *FEMS Microbiol Ecol*, 87(1), pp. 30-40.
- Selak, M., Riviere, A., Moens, F., Van den Abbeele, P., Geirnaert, A., Rogelj, I., Leroy, F. and De Vuyst, L. (2016) 'Inulin-type fructan fermentation by bifidobacteria depends on the strain rather than the species and region in the human intestine', *Appl Microbiol Biotechnol*, 100(9), pp. 4097-107.

- Seymour, G., Colquhoun, I., Dupont, M., Parsley, R. and Selvendran, R. (1990) 'Composition and structural features of cell wall polysaccharides from tomato fruits', *Phytochemistry*, 29(3), pp. 725-731.
- Shaikh, F.A., Randriantsoa, M. and Withers, S.G. (2009) 'Mechanistic analysis of the blood group antigen-cleaving endo-beta-galactosidase from *Clostridium perfringens*', *Biochemistry*, 48(35), pp. 8396-404.
- Shapiro, S. (2015) 'Exploring fructan utilisation by members of the human intestinal microbiota' Thesis, Newcastle University.
- Shigehisa, A., Sotoya, H., Sato, T., Hara, T., Matsumoto, H. and Matsuki, T. (2015) 'Characterization of a bifidobacterial system that utilizes galacto-oligosaccharides', *Microbiology*, 161(7), pp. 1463-70.
- Shipman, J.A., Berleman, J.E. and Salyers, A.A. (2000) 'Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*', *J Bacteriol*, 182(19), pp. 5365-72.
- Silk, D.B.A., Davis, A., Vulevic, J., Tzortzis, G. and Gibson, G.R. (2009) 'Clinical trial: the effects of a trans-galactooligosaccharide prebiotic on faecal microbiota and symptoms in irritable bowel syndrome', *Alimentary Pharmacology & Therapeutics*, 29(5), pp. 508-518.
- Silva, I.R., Jers, C., Meyer, A.S. and Mikkelsen, J.D. (2016) 'Rhamnogalacturonan I modifying enzymes: an update', *N Biotechnol*, 33(1), pp. 41-54.
- Simon, G.L. and Gorbach, S.L. (1986) 'The human intestinal microflora', *Digestive diseases and sciences*, 31(9), pp. 147-162.
- Sinnott, M.L. (1990) 'Catalytic mechanism of enzymic glycosyl transfer', *Chemical Reviews*, 90(7), pp. 1171-1202.
- Skorupski, K. and Taylor, R.K. (1996) 'Positive selection vectors for allelic exchange.' *Gene*, 169(1) pp. 47-52
- Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N. and Sonnenburg, J.L. (2010) 'Specificity of polysaccharide use in intestinal *Bacteroides* species determines diet-induced microbiota alterations', *Cell*, 141(7), pp. 1241-52.
- Sprinz, H., Kundel, D.W., Dammin, G.J., Horowitz, R.E., Schneider, H. and Formal, S.B. (1961) 'The response of the germfree guinea pig to oral bacterial challenge with *Escherichia coli* and *Shigella flexneri*', *Am J Pathol*, 39, pp. 681-95.
- St John, F.J., Gonzalez, J.M. and Pozharski, E. (2010) 'Consolidation of glycosyl hydrolase family 30: a dual domain 4/7 hydrolase family consisting of two structurally distinct groups', *FEBS Lett*, 584(21), pp. 4435-41.
- Stock, A.M., Robinson, V.L. and Goudreau, P.N. (2000) 'Two-component signal transduction', *Annu Rev Biochem*, 69, pp. 183-215.
- Studier, F.W. and Moffatt, B.A. (1986) 'Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes.' *J Mol Biol*. 189(1), pp. 113-130

- Suzuki, T. and Hara, H. (2004) 'Various nondigestible saccharides open a paracellular calcium transport pathway with the induction of intracellular calcium signaling in human intestinal Caco-2 cells', *J Nutr*, 134(8), pp. 1935-41.
- Tabachnikov, O. and Shoham, Y. (2013) 'Functional characterization of the galactan utilization system of *Geobacillus stearothermophilus*', *Febs j*, 280(3), pp. 950-64.
- Terrapon, N., Lombard, V., Gilbert, H.J. and Henrissat, B. (2015) 'Automatic prediction of polysaccharide utilization loci in Bacteroidetes species', *Bioinformatics*, 31(5), pp. 647-55.
- Thornton, D.J., Rousseau, K. and McGuckin, M.A. (2008) 'Structure and function of the polymeric mucins in airways mucus', *Annu Rev Physiol*, 70, pp. 459-86.
- Topping, D.L. and Clifton, P.M. (2001) 'Short-chain fatty acids and human colonic function: roles of resistant starch and non-starch polysaccharides', *Physiological reviews*, 81(3), pp. 1031-1064.
- Tremaroli, V. and Bäckhed, F. (2012) 'Functional interactions between the gut microbiota and host metabolism', *Nature*, 489(7415), pp. 242-249.
- Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A. and Affourtit, J.P. (2009) 'A core gut microbiome in obese and lean twins', *nature*, 457(7228), pp. 480-484.
- Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R. and Gordon, J.I. (2006) 'An obesity-associated gut microbiome with increased capacity for energy harvest', *Nature*, 444(7122), pp. 1027-31.
- Turner, K.W. and Roberton, A.M. (1979) 'Xylose, arabinose, and rhamnose fermentation by *Bacteroides ruminicola*', *Appl Environ Microbiol*, 38(1), pp. 7-12.
- Urbanikova, L., Vrsanska, M., Morkeberg Krogh, K.B., Hoff, T. and Biely, P. (2011) 'Structural basis for substrate recognition by *Erwinia chrysanthemi* GH30 glucuronoxylanase', *Febs j*, 278(12), pp. 2105-16.
- Van de Wiele, T., Boon, N., Possemiers, S., Jacobs, H. and Verstraete, W. (2004) 'Prebiotic effects of chicory inulin in the simulator of the human intestinal microbial ecosystem', *FEMS Microbiology Ecology*, 51(1), pp. 143-153.
- van den Broek, L.A., Hinz, S.W., Beldman, G., Vincken, J.P. and Voragen, A.G. (2008) '*Bifidobacterium* carbohydrases-their role in breakdown and synthesis of (potential) prebiotics', *Mol Nutr Food Res*, 52(1), pp. 146-63.
- van den Broek, L.A., Lloyd, R.M., Beldman, G., Verdoes, J.C., McCleary, B.V. and Voragen, A.G. (2005) 'Cloning and characterization of arabinoxylan arabinofuranohydrolase-D3 (AXHd3) from *Bifidobacterium adolescentis* DSM20083', *Appl Microbiol Biotechnol*, 67(5), pp. 641-7.
- Van der Meulen, R., Makras, L., Verbrugghe, K., Adriany, T. and De Vuyst, L. (2006) 'In vitro kinetic analysis of oligofructose consumption by *Bacteroides* and *Bifidobacterium* spp. indicates different degradation mechanisms', *Appl Environ Microbiol*, 72(2), pp. 1006-12.
- van Loo, J., Coussement, P., de Leenheer, L., Hoebregs, H. and Smits, G. (1995) 'On the presence of inulin and oligofructose as natural ingredients in the western diet', *Crit Rev Food Sci Nutr*, 35(6), pp. 525-52.

- Veereman-Wauters, G., Staelens, S., Van de Broek, H., Plaskie, K., Wesling, F., Roger, L.C., McCartney, A.L. and Assam, P. (2011) 'Physiological and Bifidogenic Effects of Prebiotic Supplements in Infant Formulae', *Journal of Pediatric Gastroenterology and Nutrition*, 52(6), pp. 763-771.
- Vidal, S., Doco, T., Williams, P., Pellerin, P., York, W.S., O'Neill, M.A., Glushka, J., Darvill, A.G. and Albersheim, P. (2000) 'Structural characterization of the pectic polysaccharide rhamnogalacturonan II: evidence for the backbone location of the aceric acid-containing oligoglycosyl side chain', *Carbohydrate research*, 326(4), pp. 277-294.
- Vijn, I. and Smeekens, S. (1999) 'Fructan: more than a reserve carbohydrate?', *Plant Physiol*, 120(2), pp. 351-60.
- Wachtershauser, A. and Stein, J. (2000) 'Rationale for the luminal provision of butyrate in intestinal diseases', *Eur J Nutr*, 39(4), pp. 164-71.
- Walker, A.W., Ince, J., Duncan, S.H., Webster, L.M., Holtrop, G., Ze, X., Brown, D., Stares, M.D., Scott, P., Bergerat, A., Louis, P., McIntosh, F., Johnstone, A.M., Lobley, G.E., Parkhill, J. and Flint, H.J. (2011) 'Dominant and diet-responsive groups of bacteria within the human colonic microbiota', *Isme j*, 5(2), pp. 220-30.
- Wang, J., Shoemaker, N.B., Wang, G.R. and Salyers, A.A. (2000) 'Characterization of a *Bacteroides* mobilizable transposon, NBU2, which carries a functional lincomycin resistance gene', *J Bacteriol*, 182(12), pp. 3559-71.
- Watson, D., O'Connell Motherway, M., Schoterman, M.H., van Neerven, R.J., Nauta, A. and van Sinderen, D. (2013) 'Selective carbohydrate utilization by lactobacilli and bifidobacteria', *J Appl Microbiol*, 114(4), pp. 1132-46.
- Weaver, C.M., Martin, B.R., Nakatsu, C.H., Armstrong, A.P., Clavijo, A., McCabe, L.D., McCabe, G.P., Duignan, S., Schoterman, M.H. and van den Heuvel, E.G. (2011) 'Galactooligosaccharides improve mineral absorption and bone properties in growing rats through gut fermentation', *J Agric Food Chem*, 59(12), pp. 6501-10.
- Wexler, H.M. (2007) '*Bacteroides*: the Good, the Bad and the Nitty-Gritty' *Clin Microbiol Rev*, 20(4), pp. 593-621
- Whisner, C.M., Martin, B.R., Schoterman, M.H., Nakatsu, C.H., McCabe, L.D., McCabe, G.P., Wastney, M.E., van den Heuvel, E.G. and Weaver, C.M. (2013) 'Galacto-oligosaccharides increase calcium absorption and gut bifidobacteria in young girls: a double-blind cross-over trial', *Br J Nutr*, 110(7), pp. 1292-303.
- Wolf, S., Mouille, G. and Pelloux, J. (2009) 'Homogalacturonan methyl-esterification and plant development', *Molecular plant*, 2(5), pp. 851-860.
- Wood, T.M. (1988) 'Preparation of crystalline, amorphous, and dyed cellulase substrates', in *Methods in Enzymology*. Academic Press, pp. 19-25.
- Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S.A., Bewtra, M., Knights, D., Walters, W.A. and Knight, R. (2011) 'Linking long-term dietary patterns with gut microbial enterotypes', *Science*, 334(6052), pp. 105-108.
- Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C., Hooper, L.V. and Gordon, J.I. (2003) 'A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis', *Science*, 299(5615), pp. 2074-6.

- Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C., Hooper, L.V. and Gordon, J.I. (2003) 'A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis', *Science*, 299(5615), pp. 2074-6.
- York, W.S., Kumar Kolli, V.S., Orlando, R., Albersheim, P. and Darvill, A.G. (1996) 'The structures of arabinoxyloglucans produced by solanaceous plants', *Carbohydr Res*, 285, pp. 99-128.
- Yu, X., Yin, J., Li, L., Luan, C., Zhang, J., Zhao, C. and Li, S. (2015) 'Prebiotic Potential of Xylooligosaccharides Derived from Corn Cobs and Their In Vitro Antioxidant Activity When Combined with *Lactobacillus*', *J Microbiol Biotechnol*, 25(7), pp. 1084-92.
- Zachar, Z. and Savage, D.C. (1979) 'Microbial interference and colonization of the murine gastrointestinal tract by *Listeria monocytogenes*', *Infect Immun*, 23(1), pp. 168-74.
- Ze, X., Ben David, Y., Laverde-Gomez, J.A., Dassa, B., Sheridan, P.O., Duncan, S.H., Louis, P., Henrissat, B., Juge, N., Koropatkin, N.M., Bayer, E.A. and Flint, H.J. (2015) 'Unique Organization of Extracellular Amylases into Amylosomes in the Resistant Starch-Utilizing Human Colonic Firmicutes Bacterium *Ruminococcus bromii*', *MBio*, 6(5), pp. e01058-15.
- Ze, X., Duncan, S.H., Louis, P. and Flint, H.J. (2012) '*Ruminococcus bromii* is a keystone species for the degradation of resistant starch in the human colon', *ISME J*, 6(8), pp. 1535-43.
- Zhang, X., Rogowski, A., Zhao, L., Hahn, M.G., Avci, U., Knox, J.P. and Gilbert, H.J. (2014a) 'Understanding how the complex molecular architecture of mannan-degrading hydrolases contributes to plant cell wall degradation', *J Biol Chem*, 289(4), pp. 2002-12.
- Zhang, X. (2015) 'Insight into the enzymatic degradation of mannan (in situ) and pectin utilization by *Bacteroides thetaiotaomicron*' Thesis, Newcastle University
- Zhang, Z., Liu, Q. and Hendrickson, W.A. (2014b) 'Crystal structures of apparent saccharide sensors from histidine kinase receptors prevalent in a human gut symbiont', *FEBS J*, 281(18), pp. 4263-79.

Appendix A

A.1 Protein localisation based on LipoP 1.0 analysis

LipoP analysis of protein sequences expressed as recombinant proteins in this project was used to predict the cellular localisation when expressed in *Bacteroides*. There are 3 possible locations for each protein discussed here, surface attached lipoprotein, periplasm facing lipoprotein and soluble in the periplasm. LipoP uses analysis of N-terminal signal sequences to predict protein localisation in gram negative bacteria (Figure A.1). The analysis output gives 4 predicted classes, Spl, signal peptide I, SpII, lipoprotein signal peptide II, TMH, n-terminal transmembrane helix and CYT, cytoplasmic. The lipoprotein signal typically included a cysteine at position 20 which is used to covalently link the protein to the lipid membrane. Figure A.1 shows an example result of BACOVA_04390 and BACOVA_04387 signal sequence analysis with LipoP 1.0. The LipoP results gives potential cleavage sites and a prediction of the signal sequence present. When multiple possible sites are found the software generates a graph showing the most likely cleavage site (Figure A.1). The results show presence of a lipoprotein signal (SpII) at the N-terminal of BACOVA_04390 and signal sequence (Spl) at the N-terminal of BACOVA_04387.

BACOVA_04390

```
# Sequence SpII score=13.4585 margin=13.659413 cleavage=19-20 Pos+2=D
# Cut-off=-3
Sequence      LipoP1.0:Best  SpII   1      1      13.4585
Sequence      LipoP1.0:Margin SpII   1      1      13.659413
Sequence      LipoP1.0:Class  CYT    1      1      -0.200913
Sequence      LipoP1.0:Class  SpI    1      1      -2.61098
Sequence      LipoP1.0:Signal CleavII 19     20     13.4585 # LILTS|CDDNK Pos+2=D
```

BACOVA_04387

```
# Sequence SpI score=11.4217 margin=11.622613 cleavage=24-25
# Cut-off=-3
Sequence      LipoP1.0:Best  SpI    1      1      11.4217
Sequence      LipoP1.0:Margin SpI    1      1      11.622613
Sequence      LipoP1.0:Class  CYT    1      1      -0.200913
Sequence      LipoP1.0:Signal CleavI 24     25     11.3294 # GEVFA|KDGSS
Sequence      LipoP1.0:Signal CleavI 20     21     7.36119 # SFSYG|EVFAK
Sequence      LipoP1.0:Signal CleavI 22     23     1.02515 # SYGEV|FAKDG
Sequence      LipoP1.0:Signal CleavI 18     19     0.868677 # MFSFS|YGEVF
Sequence      LipoP1.0:Signal CleavI 19     20     0.432756 # FSFSY|GEVFA
Sequence      LipoP1.0:Signal CleavI 25     26     -1.96794 # EVFAK|DGSSL
```

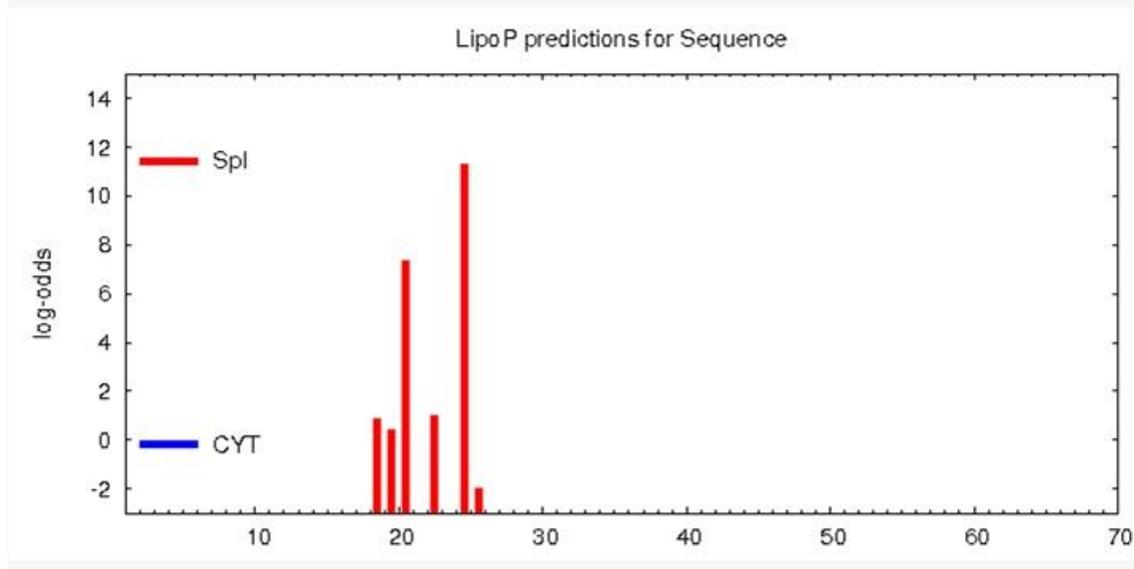


Figure A.1 LipoP 1.0 results of sequence analysis of BACOVA_04390 and BACOVA_04387.

Protein	Predicted function/Activity	PUL	Predicted Cellular location
BACOVA_04390	GH10/CBM	Xylan	Surface
BACOVA_04387	GH10		Periplasm
BACOVA_03432	GH30		Surface
BACOVA_03433	GH98		Surface
BT_4667	GH2	Galactan	Periplasm
BT_4668	GH53		Surface
BT_4669	SGBP		Surface
BT_4670	SusD		Surface
BT_4673	HTCS		Periplasm
BT_0348	GH51	Arabinan	Periplasm
BT_0360	GH43		Surface
BT_0361	SusD		Surface
BT_0363	SusD		Surface
BT_0365	SGBP		Surface
BT_0366	HTCS		Periplasm
BT_0367	GH43		Surface
BT_0368	GH51		Periplasm
BT_0369	GH43		Periplasm
BT_4151	GH2		RGI
BT_4156	GH2	Periplasm	
BT_4160	GH35	Periplasm	
BT_4181	GH2	Periplasm	

Table A.1 Prediction of protein localisation discussed in this project

A.2 Validation of GH10 kinetic analysis of XOS hydrolysis

In chapter 3.2.2.3 characterisation of the pair of GH10 enzymes, BACOVA_04390 and BACOVA_04387, by substrate depletion requires the concentration of substrate to be such that it is below the K_M of the enzymes to ensure reaction rates observed are within the linear phase of the reaction curve (Michaelis-Menten).

$$k \cdot t = \ln \left(\frac{[S_0]}{[S_t]} \right)$$

(Equation A.1)

Where: $k = k_{cat}/K_M$, t = time, and $[S_0]$ and $[S_t]$ represent the substrate concentration at time 0 and t , respectively. This relationship is only valid when the concentration of enzyme, $[E]$, is \ll than substrate concentration, $[S]$, $\ll K_M$ (Matsui *et al.*, 1991).

To ensure the substrate concentration used was indeed below the K_M of the enzymes different concentrations of substrate was used with the same concentration of enzyme (0.05 μM). Initial scoping experiments suggested the highest activity of each enzyme would be on xylohexaose (X6, Data not shown) so this was used as the substrate to investigate which enzyme concentration to use in the assays (Table A.2).

Enzyme	Substrate concentration (μM)	Enzyme concentration (nM)	Rate ($\text{min}^{-1} \text{mM}^{-1}$)
BACOVA_04390	75	50	110
	50	50	102
	25	50	120
BACOVA_04387	75	50	3208
	50	50	3250
	25	50	3256

Table A.2 | Rate of substrate depletion of different concentrations of X6 by BACOVA_04390 and BACOVA_04387.

Although amount of products differs as substrate concentration was changed the rates remained consistent (Table A.2), indicating that the reaction was indeed within the linear phase of Michaelis-Menten kinetics curve.

A.3 *Bacteroides* qPCR Tag insertion

Insertion of unique sequences, referred to as 'Tags', into *Bacteroides* genome were performed as described in chapter 2.9.4. PCR from the resulting genomes were run on agarose gels (Figure A.2). These PCR products were produced using primers to target either Tag sequence or ATT site 1. The presence of 200 bp Tag1 or Tag11 product, indicate these products have successfully been inserted into the genome while the ATT site 1 products indicate the tag had inserted into ATT site 2. These tags were used as targets for qPCR to identify different *Bacteroides* strains in a co-culture.

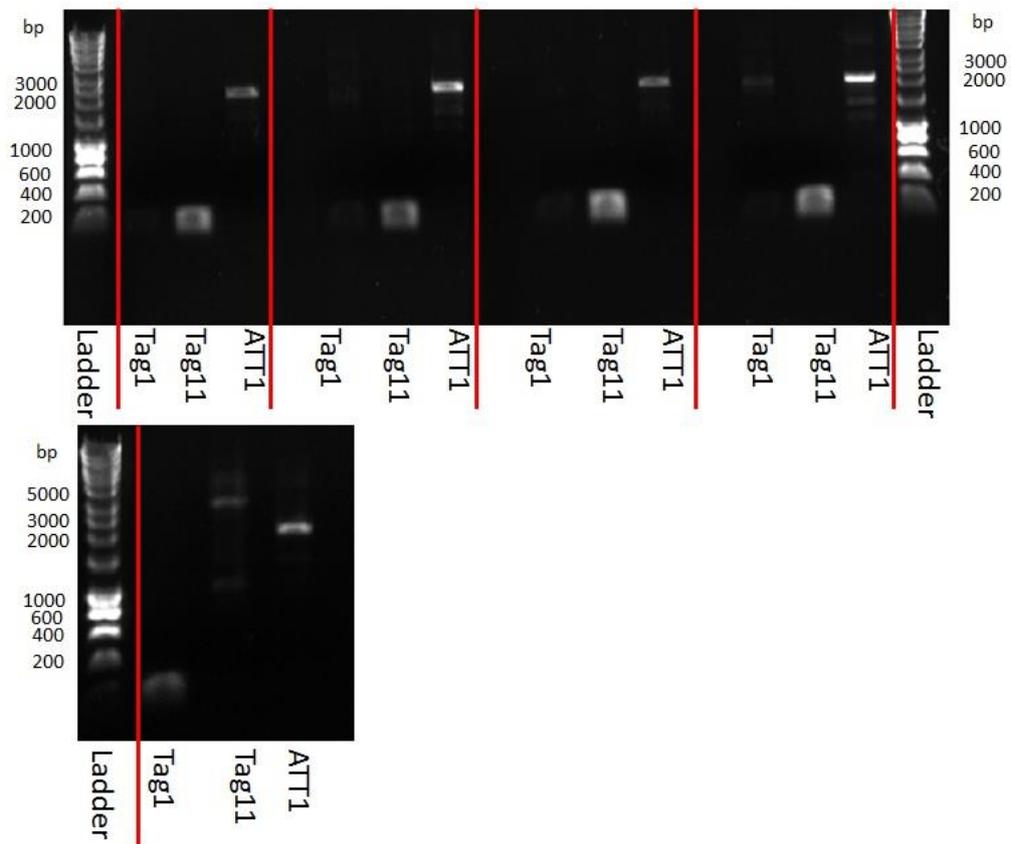


Figure A.2 Examples of agarose (0.8 %) DNA gel showing conformation of *tag1/11* insertion into *Bacteroides* genome.

A.4 Recombinant pectin PUL Protein Expression

Chapter 4 gives examples of a few expression gels of recombinant proteins used in that chapter.

Figure A.3, Figure A.4 and Figure A.5 show a full list of expression gels of all proteins used in chapter

4.

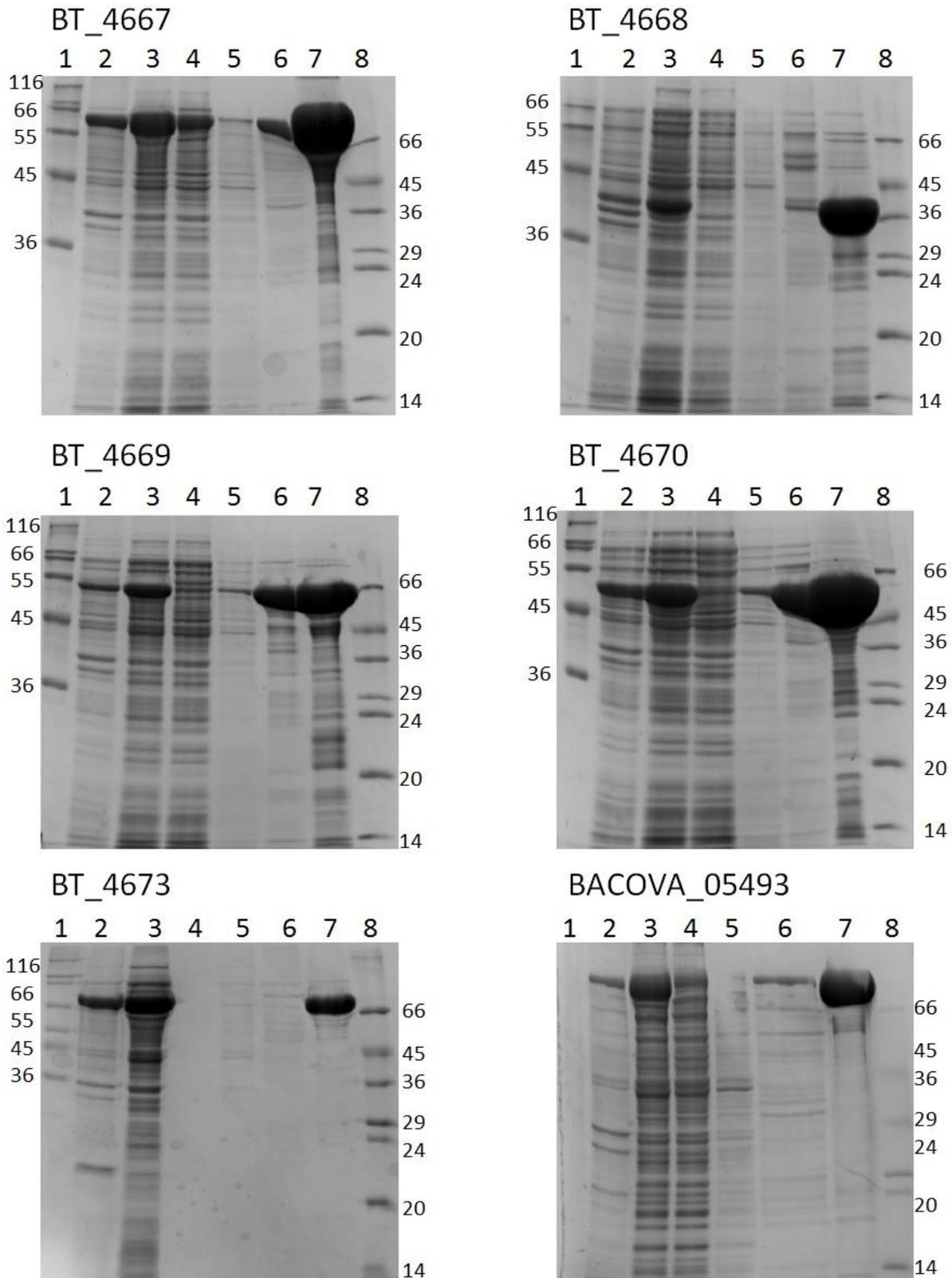


Figure A.3 Galactan PUL protein expression SDS-PAGE. Fractions collected from IMAC were subjected to SDS-PAGE to evaluate protein content. Lane 1, High/wide molecular weight marker, 2, cell pellet, 3, cell lysate, 4, flow through, 5, wash fraction, 6, elution fraction, 7, elution fraction, 8, low/wide molecular weight marker. Values of weight marker are given in kDa.

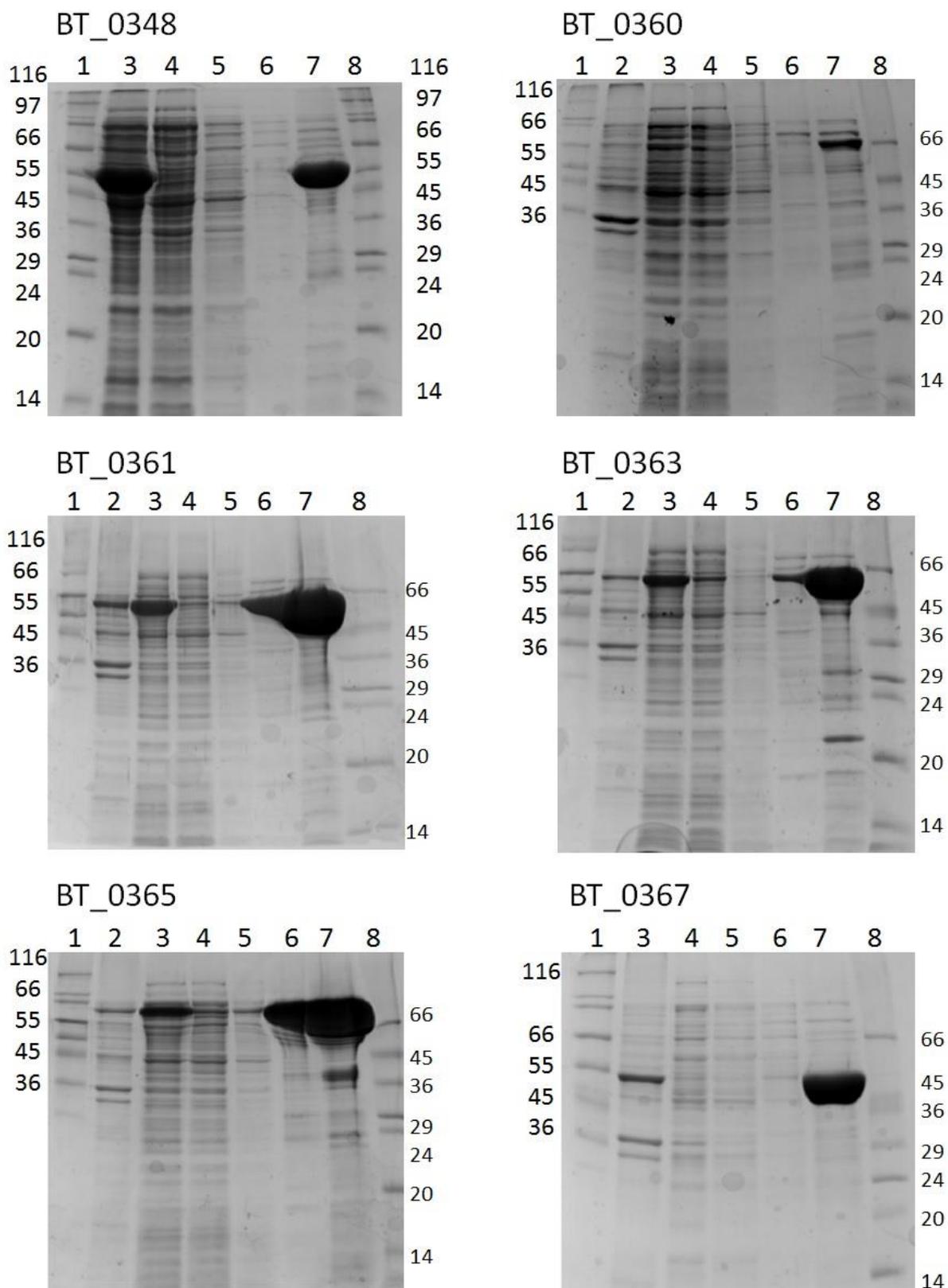


Figure A.4 Arabinan PUL protein expression SDS-PAGE. Fractions collected from IMAC were subjected to SDS-PAGE to evaluate protein content. Lane 1, High/wide molecular weight marker, 2, cell pellet, 3, cell lysate, 4, flow through, 5, wash fraction, 6, elution fraction, 7, elution fraction, 8, low/wide molecular weight marker. Values of weight marker are given in kDa.

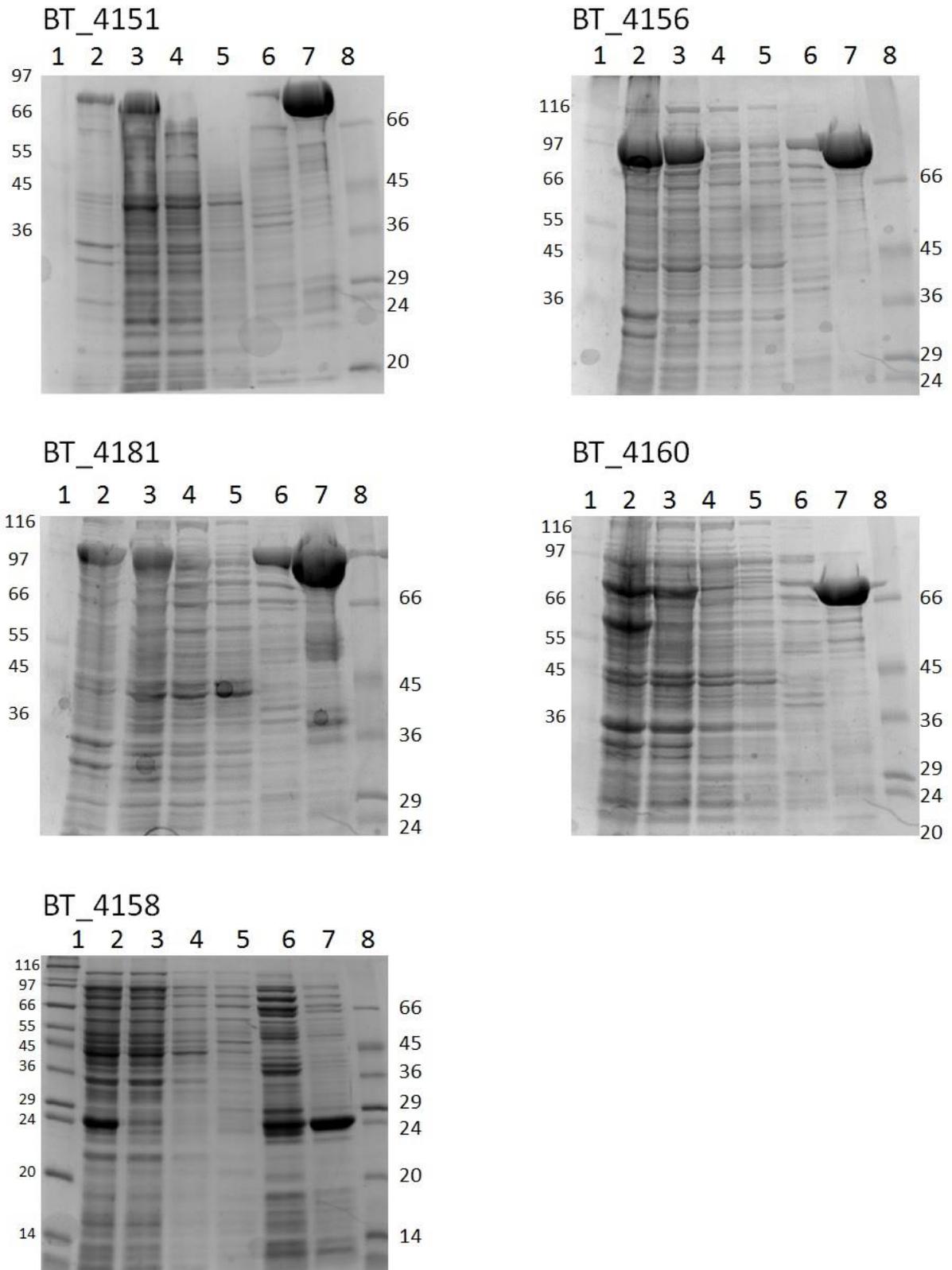


Figure A.5 RGI PUL protein expression SDS-PAGE. Fractions collected from IMAC were subjected to SDS-PAGE to evaluate protein content. Lane 1, High/wide molecular weight marker, 2, cell pellet, 3, cell lysate, 4, flow through, 5, wash fraction, 6, elution fraction, 7, elution fraction, 8, low/wide molecular weight marker. Values of weight marker are given in kDa.

A.5 Validation of BT_4668 (GH53) kinetic analysis of galactooligosaccharide hydrolysis

In Chapter 4.2.2.3 kinetic analysis of BT_4668 action was performed using HPAEC to assay for depletion of each oligosaccharide substrate in a reaction mixture. As mentioned in Appendix A.2 this relationship of rate of substrate depletion by enzyme concentration is only valid in the linear phase of the reaction where the enzyme is at maximal possible activity giving a direct readout of k_{cat}/K_M of the reaction. This is validated by ensuring the substrate concentration is sufficiently below the K_M of the enzyme on that specific substrate. Preliminary investigation showed the highest activity, thus likely the lowest K_M would be when the enzyme is acting on galactohexaose (Gal6). The concentration of substrate was altered and rate of substrate depletion was measured for each reaction (Table A.3). Indeed for each of the concentrations used the rate of substrate depletion remained the same indicating the reaction remained in the linear phase of the reaction, according to Michaelis-Menten reaction kinetics.

Enzyme	Substrate concentration (μM)	Enzyme concentration (nM)	Rate ($\text{min}^{-1} \text{mM}^{-1}$)
BT_4668	75	50	23.3
	50	50	18.6
	25	50	20.2

Table A.3 Rate of substrate depletion of different concentrations of Gal6 by BT_4668.

Appendix B: Primer list

Primers used in this project are listed in table B.1.

Primer Name	Sequence	Use
BT_4668E180AF	GAATGGGTACAGGTAGGTAATGCGACAACCTCCCGCATGATG	Mut
BT_4668E180AR	CATCATGCCGGGAGTTGTCGCATTACCTACCTGTACCCATTC	Mut
BT_4668E292AF	ACAAACCTGTCATGATTTGCGCGATAGGTATGCCTTATGATCAG	Mut
BT_4668E292AR	CTGATCATAAGGCATACCTATCGCGCAAATCATGACAGGTTTGT	Mut
BT_4668pEXF	CTCGACTCTAGAAACGGGTGTAGTGAA	CpEX
BT_4668pEXR	CTCGACCAGCTGTTATTGGATTTTAAAAG	CpEX
BT_4667EAF	Gtggagtatcggtaatcggtccccacgcaatgc	Mut
BT_4667EAR	GCATTGCGTGGGGACCGCATTACCGATACTCCAC	Mut
BT_4667pEXF	ATAAGTAGTCGACATGCTACAGGCACAGCGCAGCGAG	CpEX
BT_4667pEXR	CGCGGTATCTAGATTATTTTCGTCAGAAGACGGATACTTCC	CpEX
BT_4669FF	ATAAGTAGTCGACAACCTGAAAGAACTGGTTGCCGAA	SOE
BT_4669FR	CATAATCCACTCTATTA AAAAGATACGATATCAGTTTCAATG	SOE
BT_4669BF	TTAAAAAGATACGATATCAGTTTCAATGAAAATAGCAAGC	SOE
BT_4669BR	GTAGGTATCTAGAATTATAGCCATTCGGGGCTTG	SOE
BT_4670FF	ATAAGTAGGATCCAGATCTGGCATTCTTAGAAAAC	SOE
BT_4670FR	GTATCTTGCACCATACGATTATTGTTAGTATATTA	SOE
BT_4670BF	GTATGGTGCAAAGATACGATATGAAAGTTTTT	SOE
BT_4670BR	CTAAATATCTAGATATCGTTGTTTCCCGTTGTA	SOE
BT_4673pETF	ATAAGTAGTaatatttatctcggtgctga	CpET
BT_4673pETR	CGCGGTATCTcggtggaagaatccggatattcagaata	CpET
BACOVA_05493FF	AACATTGAGtgcacAATAAAGATTCGCTGCTG	SOE
BACOVA_05493FR	CCGATAAATTTGATAGCTCATTCTTAAAGTTCCCGG	SOE
BACOVA_05493BF	CGGGGAACCTTAAGAATGAGCTATCAAATTTATCGG	SOE
BACOVA_05493BR	GGCGGCCGCTCTAGA TGTAATCTTCAGTTTATCC	SOE
GalPULFF	ATAGCTGGATCCTTGTA AAAAGAAGAAGGC	SOE
GalPULFR	CTAAAATGCCTTCTCACTTTTTTATTGGATT	SOE
GalPULBF	AAAAGTGAGAAGGCATTTTAGGGCTTTTCG	SOE
GalPULBR	TTCAATCTAGATGTCGTTCCGGGTTTATA	SOE
BT_0365FF	AGTAATGTCGACCTTCGTGAACAGTATCATGTAC	SOE
BT_0365FR	ACTAAACTTTAAGGTAAGTAGTAACGGCCATTTCTTGTC	SOE
BT_0365BF	GGCCGTTACTACTT ACCTTAAAGTTTAGTATACATGAAAA	SOE
BT_0365BR	GATATA TCTAGA CATACTTTGGTACCGCCCTGTA	SOE
BT_0362FF	AGTAATGTCGACGCCGATGCAGCGCAAAG	SOE
BT_0362FR	CTGTTTTATTGAATGATCAGTTAATTGTTCCAACCCGGG	SOE
BT_0362BF	TGGAACAATTAAGTATCATTCAATAAAAACAGCAAAGAA	SOE
BT_0362BR	GATATA TCTAGA GCCTACCGGCAGGAATACTTTTTTC	SOE
BT_0364FF	AGTAAT GGATCC GGTGCATTCATTTATGGTGC	SOE
BT_0364FR	CATGTCTTATATCTGGTATACTAACTTTAAGGTTTAATTC	SOE
BT_0364BF	AAAGTTTAGTATACCAGATATAAGACATGAAAAAGACAA	SOE
BT_0364BR	GATATA TCTAGA TACATCATTACGCCAGAAGATC	SOE
BT_0360DAF	GGCAGCGTGCAGCGACGCGGATGAAAACCTCCGCATCAGG	Mut
BT_0360DAR	CCTGATGCGGAGTTTTTCATCCGCGTGCCTGCACGCTGCC	Mut
BT_0360pEXF	AGTAATGGATCCCAGCGACGACGATGAAAAC	CpEX

BT_0360pEXR	GCAAGATCTAGATTATTGGAGTTTCTTCCCC	CpEX
BT_0367DAF	CTCCTTCTGCTAATTCTTGGGATGCGAATTATTTATCAGTAGCC	Mut
BT_0360DAR	GGCTACTGATAAATAATTCGCATCCCAAGAATTAGCAGAAGGAG	Mut
BT_0367pEXF	CGTGCTGGATCCAATTCTTGGGATGATAATTATTTATC	CpEX
BT_0367pEXR	GTAATATCTAGAGATTCTTTTTCCCCAGACAGAACG	CpEX
BT_4156EAF	GGAGTAACGGAAACGCGGGCGGCTGGAAT	Mut
BT_4156EAF	ATTCCAGCCGCCCGCTTCCGTTACTCC	Mut
BT_4156pEXF	GATAACATTCGAGTCGACatcgaagtccctgtaac	CpEX
BT_4156pEXR	TGGCGGCCCTCTAGAgTcaaacatcagattcag	CpEX
BACOVA_04390AE1F	ggcaagttgattaaagatccGCCttggatattaaggttaacac	Mut
BACOVA_04390AE1R	GTGTTAACCTTAATATCCAAGGCGGATACTTTAATCAACTTGCC	Mut
BACOVA_04390AE2F	CatgcatgggacgtagtcaatGCGccaatggatgacggaaaaac	Mut
BACOVA_04390AE2R	GTTTTCCGTCATCCATTGGCGCATTGACTACGTCCCATGCATG	Mut
BACOVA_04390pEXF	CGCGGATCCATGGAGTGGTATAAAGACCCTAC	CpEX
BACOVA_04390pEXR	CCCTCTAGATTCCAAATCTCCGGTAAAGTCTC	CpEX
BACOVA_03433E361AF	tcttggttcaattattgtGCCcagtttgggggtatgatg	Mut
BACOVA_03433E361AR	CATCATACCCCCAAAACCTGGGCACAATAATTGAAACCAAGA	Mut
BACOVA_03433D467AF	tcagtacggcatcggtttcGCG caatgcggtggacagagg	Mut
BACOVA_03433D467AR	TCTGTCCAGCCGCATTGCGCGAAACGGATGCCGTACT	Mut
BACOVA_03433pEXF	cgcGTCGACTGTATTCTTGGTTCTGTAAAGATGAC	CpEX
BACOVA_03433pEXR	cccTCTAGATTTTCTTTCGATAACAATGTTGTCGAG	CpEX
Tag1Forward	ATGTCGCCAATTGTCACTTTCTC	qPCR
Tag11Forward	ATGCCGCGGATTTATTGGAAGAA	qPCR
TagReverse	CACAATATGAGCAACAAGGAATCC	qPCR
ATT1Forward	CCTTTCACCGCTTTCAACG	Diag
ATT1Reverse	TCAACTAAACATGAGATACTAGC	Diag

Table B.1 Primers used in this Project. The primers are classified by use, Mut, site directed mutagenesis, CpEX, cloning into pExchange, CpET, cloning into pET, SOE, for overlap extension PCR, qPCR, denotes which primers were used to amplify tags in qPCR and, Diag, for primers used in diagnostic PCR.

Appendix C: Chemicals, Media, Enzymes and Substrate Suppliers

C.1 Chemicals

Amersham-Boehringer Mannheim

2'-Deoxyadenosine 5'-triphosphate (dATP)

2'-Deoxycytidine 5'-triphosphate (dCTP)

2'-Deoxyguanosine 5'-triphosphate (dGTP)

2'-Deoxythymidine 5'-triphosphate (dTTP)

BioGene

Electrophoresis grade agarose

British Drug Houses (BDH)

Acetic Acid (Glacial)

Acrylamide solution (40% w/v; Electran)

Boric acid

Bromophenol blue

Citric Acid

Calcium Chloride

Chloroform

Dimethylformamide

Ethanol (industrial grade)

Hydrochloric acid

Isopropanol

Magnesium Chloride

Magnesium sulphate

Methanol

Polyethelene glycol MW 400 (PEG-400)

Polyethelene glycol MW 550 (PEG-5500)

Polyethelene glycol MW 1000 (PEG-1000)
Polyethelene glycol MW 20000 (PEG-20000)
Sodium acetate
Sodium Chloride
Sulphuric acid

Fisions

46/48% w/v NaOH
Sodium acetate trihydrate

James Burrough (F.A.D.) Ltd

Ethanol

Megazyme

Sugar beet arabinan
Debranched arabinan
Linear Arabinan
Arbinooligosaccharides (DP 2-8)
Arabinose
RGI
Galactan from potato
Galactan from lupin
Galactose
Lactose
Inulin from chicory
Fructan from onion
Fructan from wheat
Fructan from agave
Neosugar (FOS)

Raftilose (FOS)

Kestose

Kestotetraose

Kestopentaose

Fructotriose

Melford Laboratories

Isopropyl- β -D-thiogalactosidase (IPTG)

HEPES

G.E. Healthcare

Agarose (ultrapure)

Sigma-Aldrich

3,5-Dinitrosalicylic acid (DNSA)

Ammonium persulphate

Ampicillin

Bis tris propane

Bovine serum albumin, fraction V (BSA)

Chloramphenicol

Coomassie brilliant blue G

D-Glucose

di-Sodium hydrogen phosphate

Ethelene diamine tetra-acetic acid, disodium salt (EDTA)

Ethidium bromide

Ethylene glycol

Glycerol

Imidazole

Kanamycin

Tetracycline

Levan from *E. herbicola*

Levan from *Z. mobilis*

N,N,N',N'-Tetramethylethylenediamine (TEMED)

Nicotinamide adenine dinucleotide-reduced

Phenol

Polyethylene glycol MW 3350 (PEG-3350)

Sodium bicarbonate

Sodium carbonate

Sodium dihydrogen orthophosphate

Sodium dodecyl sulphate (SDS)

Sucrose (nuclease free)

Trizma base (Tris)

B-Mercaptoethanol

C.2 Media

Difco

Bacto[®]tryptone/peptone

Bacto[®]yeast extract

Oxoid

Bacteriological Agar

Sigma-Aldrich

LB Broth

Clostridial Media

C.3 Enzymes

MBI Fermentase

DNA restriction endonucleases

Invitrogen

Bacteriophage T4 DNA ligase

Novagen

KOD HotStart DNA polymerase

Stratagene

Dpn1 restriction endonuclease

C.4 Kits

Plasmid Mini Kit

Plasmid Midi Kit

Qiaquick Gel Extraction Kit

Qiaquick PCR Purification Kit

Sigma-Aldrich

GeneElute™ Bacterial Genomic DNA Kit

Stratagene

QuikChange™ Site-Directed Mutagenesis Kit

Megazyme

D-Mannose/D-Fructose/D-Glucose assay Kit

Galactose/Arabinose assay kit

Acetic acid detection kit

Roche

LightCycler® 480 SYBR Green I Master