Human Milk Oligosaccharides: Bioactive Sugars to Exploit the Bacterial Sweet Tooth

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering

2023

Jessica E. Lloyd

School of Natural Sciences, Department of Chemistry

List of Contents

Word Count: 44,187

Nomenclature

IUPAC-approved symbol nomenclature and abbreviations for the graphical representation of glycans.

○ Hex	● Glc	● Man	● Gal	● Gul	● Alt	● All	● Tal	● Ido
□ HexNAc	■ GlcNAc	■ ManNAc	■ GalNAc	■ GulNAc	■ AltNAc	■ AllNAc	■ TalNAc	■ IdoNAc
◨ HexN	◨ GlcN	◨ ManN	◨ GalN	◨ GulN	◨ AltN	◨ AllN	◨ TalN	◨ IdoN
◇ HexA	◆ GlcA	◆ ManA	◆ GalA	◆ GulA	◆ AltA	◆ AllA	◆ TalA	◆ IdoA
△ ndHex	▲ Qui	▲ Rha	▲ Fuc		▲ 6dAlt		▲ 6dTal	
◭ dHexNAc	◭ QuiNAc	◭ RhaNAc	◭ FucNAc		◭ 6dAltNAc		◭ 6dTalNAc	
▭ ddHex	▬ Oli	▬ Tyv		▬ Abe	▬ Par	▬ Dig	▬ Col	
☆ Pen		★ Ara	★ Lyx	★ Xyl	★ Rib			
◇ dNon-ulosonic	◆ Kdn				◆ Neu5Ac	◆ Neu5Gc	◆ Neu	
◇ ddNon-ulosonic	◆ Pse	◆ Leg		◆ Aci	◆ 4eLeg			
⬡ Unknown	⬢ Bac	⬢ LDManHep	⬢ Kdo	⬢ Dha	⬢ DDManHep	⬢ MurNAc	⬢ MurNGc	⬢ Mur
⬠ Assigned	⬟ Api	⬟ Fruc	⬟ Tag	⬟ Sor	⬟ Psi		⬟ Non-sugar	

List of Abbreviations

2'FL	2'-Fucosyllactose
3'SL	3'-Sialyllactose
3-APM	3-Azidopropyl 2,3,4,6-tetra-*O*-acetyl-α-D-Mannopyranoside
3FL	3-Fucosyllactose
6'SL	6'-Sialyllactose
ACN	Acetonitrile
AcOH	Acetic Acid
ADMP	2-Azido-1,3-Dimethylimidazolinium Hexafluorophosphate
ADP	Adenosine Diphosphate
AMP	Adenosine Monophosphate
ATP	Adenosine 5'-Triphosphate
BCN	N-[(1R,8S,9s)-Bicyclo[6.1.0]non-4-yn-9-ylmethyloxycarbonyl]-1,8-diamino-3,6-dioxaoctane
BMO	Bovine Milk Oligosaccharide
BSA	Bovine Serum Albumin
CAD	Charged Aerosol Detector
Cds	Cadmium Sulphide
CE	Capillary Electrophoresis
CFU	Colony Forming Unit
CHO	Chinese Hamster Ovary
ConA	Concanavalin A
CPS	Capsular Polysaccharide
CSH	Cell Surface Hydrophobicity
CV	Column Volume
D-Ara	D-Arabinose
DCM	Dichloromethane
DC-SIGN	Dendritic cell-specific ICAM3-grabbing non-integrin
DFL	Difucosyllactose
DHAP	Dihydroxyacetone Phosphate
DIPEA	N,N-Diisopropylethylamine
DIPEA	N,N-Diisopropylethylamine
DMC	2-Chloro-1,3-Dimethylimidazolinium Chloride
DMF	Dimethylformamide
DMSO	Dimethyl Sulfoxide
DP	Degrees of Polymerisation
DSLNT	Disialyllactose-*N*-tetraose
DTSP	3,3'-Dithiodipropionic acid di(N-Hydroxysuccinimide Ester)
E. coli	*Escherichia coli*
EFSA	European Food Safety Authority
ELISA	Enzyme-Linked Immunosorbent Assay
ELSD	Evaporative Light Scattering

ESI	Electrospray Ionisation
EtOAc	Ethyl Acetate
EtOH	Ethanol
FDA	Food and Drug Administration
FHA	Filamentous Hemagglutinin
FITC	Fluorescein Isothiocyanate
FKP	L-Fucokinase/GDP-L-Fucose Pyrophosphorylase
Fuc	Fucose
fucA	L-Fuculose 1-Phosphate Aldolase
fucI	L-Fucose Isomerase
fucK	L-Fuculose Kinase
fucO	L-1,2-Propanediol Oxidoreductase
fucP	L-Fucose Permease
fucR	L-Fucose Operon Activator
FucT	Fucosyltransferase
FUT2	α-1,2-Fucosyltranferase encoding gene
FUT3	α-1,3/4-Fucosyltransferase encoding gene
FX	GDP-Fucose Synthase
GAG	Glycosaminoglycan
Gal	Galactose
GalNAc	N-Acetyl-D-Galactosamine
GALT	Gut-Associated Lymphoid Tissue
GalT	β1-4-Glactosyltransferase
GBP	Glycan-Binding Protein
GBS	Group B *Streptococcus*
GDP	Guanosine Diphosphate
GDP-Ara	GDP-D-Arabinose
GDP-Fuc	GDP-L-Fucose
GDP-Gal	GDP-LGalactose
GFET	Graphene Field-Effect Transistor
Glc	Glucose
GlcNAc	N-Acetyl-D-Glucosamine
GMD	GDP-D-Man-4,6-Dehydratase
GMP	Guanosine Monophosphate
GPC	Gel Permeation Chromatography
GTP	Guanosine 5'-Triphosphate
H. mustelae	*Helicobacter mustelae*
H. pylori	*Helicobacter pylori*
HBHHA	Heparin-Binding Hemagglutinin Adhesin
HILIC	Hydrophilic Interaction Liquid Chromatography
HIV	Human Immunodeficiency Virus
HmFucT	*Helicobacter mustelae* α1,2-Fucosyltransferase
HMO	Human Milk Oligosaccharide
HPAEC-PAD	High pH Anion-Exchange Chromatography with Pulsed Amperometric Detection

HpFucT	*Helicobacter pylori* α1,2-Fucosyltransferase
HPLC	High Pressure Liquid Chromatography
HR-MS	High Resolution Mass Spectrometry
HRP	Horseradish Peroxidase
IBS	Irritable Bowel Syndrome
IPA	Isopropanol
IPTG	Isopropyl β-D-1-Thiogalactopyranoside
Lac-AP	3-Aminopropyl Lactose
Lac-N$_3$	1-Azido Lactose
LacNAc	N-Acetyl-D-Lactosamine
LacNB	Lacto-*N*-Biose
Lac-PEG$_4$-N$_3$	1-Azido-PEG$_4$-3-Aminopropyl Lactoside
Lac-triF	3-Trifluoroacetamidopropyl Lactoside
LC	Liquid Chromatography
Le	Lewis
L-Fuc	L-Fucose
L-Gal	L-Galactose
LNDFH-I	Lacto-*N*-Difucohexaose I
LNFP	Lacto-*N*-Fucopentaose
LNFPI	Lacto-*N*-Fucopentaose I
LNFPII	Lacto-*N*-Fucopentaose II
LNFPIII	Lacto-*N*-Fucopentaose III
LNFPV	Lacto-*N*-Fucopentaose V
LNnT	Lacto-*N*-Neotetraose
LNT	Lacto-*N*-Tetraose
LNT1	Lacto-*N*-Triose I
LNT2	Lacto-*N*-Triaose II
LOS	Lipooligosaccharide
LPS	Lipopolysaccharide
LR-MS	Low Resolution Mass Spectrometry
M1	HMO Mix 1
M2	HMO Mix 2
Man-CDs	Mannose-Coated Cds Quantum Dots
MeOH	Methanol
MQ	Milli-Q Water
MS	Mass Spectrometry
MW	Molecular Weight
Neu5Ac	N-Acetylneuraminic acid
Neu5GC	N-glycolylneuraminic acid
OD	Optical Density
PAK	*Pseudomonas aeruginosa K*
***Para*-LNnH**	*Para*-Lacto-*N*-Neohexaose
PBASE	1-Pyrenebutanoic Acid Succinimidyl Ester
PBS	Phosphate Buffered Saline

PEG	Poly (ethylene glycol)
PMO	Porcine Milk Oligosaccharide
PNC	Platelet-Neutrophil Complex
PPE	Poly(p-phenylene ethynylene)
ppm	Parts Per Million
PTFE	Polytetrafluoroethylene
QCM	Quartz Crystal Microbalance
RBF	Round Bottom Flask
Rf	Retention Factor
RI	Refractive Index
RSV	Respiratory Syncytial Virus
RT	Room Temperature
S. cerevisiae	*Saccharomyces cerevisiae*
SAX	Strong Anion-Exchange Chromatography
SCFAs	Short Chain Fatty Acids
SDS-PAGE	Sodium Dodecyl-Sulphate Polyacrylamide Gel Electrophoresis
Se	Secretor
SM	Starting Material
TFA	Trifluoroacetic Acid
TLC	Thin Layer Chromatography
UDP	Uridine-5'-Diphosphate
UDP-Gal	Uridine-Diphospho-Galactose
UEA-I	*Ulex europaeus* Agglutinin I
WHO	World Health Organisation
WT	Wild-Type
α1,2-FucT	α1,2-Fucosyltransferase
α1,3-FucT	α1,3-Fucosyltransferase

Abstract

Carbohydrates, or glycans, are ubiquitous in nature and present across all kingdoms of life - bacteria, fungi, viruses, yeast, plants, animals, and humans. They are essential to many biological processes. Fucose is a deoxyhexose; an unusual sugar, which is abundant in host mucosal surfaces, where it decorates terminal positions of mucin carbohydrate ligands; microbial cell surface structures and is found in some dietary sources [1]. Fucosylated glycans are involved in a variety of physiological processes including immunity, brain development, cancer and host-microbe interactions. As such, opportunities for the detection, prevention, and intervention of bacterial infection through diet or novel therapeutics are possible due the cell surface, species-specific carbohydrates and carbohydrate-binding proteins presented by pathogens. Human milk oligosaccharides (HMOs) are associated with a range of diverse health benefits, acting as prebiotics for commensal bacteria, receptor decoys for pathogens and as immune stimulants. Their bioactive properties provide a unique opportunity to impact bacteria-associated disease in humans and animals.

Over 200 different structures of HMO have been identified in human breast milk, significantly more than are present in the milk of livestock and most primates. However, their biosynthesis is largely unknown and there are many challenges that impede either their purification from donated samples or chemical and/or enzymatic synthesis. The aim of this research is to 1) develop improved liquid chromatography strategies for the characterisation and purification of HMOs; 2) investigate the synthesis of fucosylated HMOs and novel, non-natural HMO analogues; 3) assess the ability of HMOs to modulate bacterial growth and 4) develop a selective and sensitive microplate-based assay to detect glycan-bacterial interactions in live bacteria.

Liquid chromatography methods were optimised for the characterisation and purification of HMOs, including for the analysis of fucosylation reactions. To enable benchtop analysis and purification of HMOs, a method for the use of HILIC cartridges was developed, which successfully resolved mixtures of HMOs. Investigations with α1,2-fucosyltranferase from *H. mustelae,* showed that it is a promiscuous enzyme that can accept and convert a variety of di-, tri-, tetra- and pent- saccharides, including core HMO structures and fucosylated glycans and functionalised lactosides. This enzyme is capable of synthesising HMO analogues and the procedure can be scaled up when used in combination with L-fucosekinase/GDP-L-fucose pyrophosphorylase (FKP) in a two-step, one pot reaction. Furthermore, FKP can convert D-Ara and L-Gal into their respective GDP-sugar

nucleotides, and these can be processed as donors by α1,2-fucosyltransferase to synthesise non-natural HMO analogues. Microbiological assays revealed that *Escherichia coli* is able to utilise L-fucose as a sole energy source but not HMOs, nor do HMOs affect microbial growth when co-fed with fucose. Finally, an ELISA-like binding assay was developed which allowed selective binding of E. coli to glycan-functionalised microplates. The methodology can be adapted to gold nanoparticles, quantum dots and gold/graphene chips for use as highly specific biosensors for detection of pathogens or identification of pathogen-specific glycan motifs for the development of targeted drug therapies.

Declaration

I declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given the University of Manchester certain rights to use such Copyright, including for administrative purposes.

Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, the University Library's regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in the University's policy on Presentation of Theses.

This PhD thesis is dedicated to my grandparents, Dr Joyce Martin and Dr Judith Finch, without whose unwavering support and good examples, this thesis would not be possible.

Thank you, Auntie Joyce and Grandma.

1935 – 2021

1936 – 2022

Acknowledgement

First and foremost, I would like to thank my supervisor, Prof Rob Field and giving me the opportunity to be part of the group and work alongside so many remarkable people. Thank you for your expert supervision, patience, and encouragement throughout a turbulent four years. Without your support, I am certain that my imposter syndrome would have gotten the better of me and this thesis wouldn't exist – so thank you for your unwavering optimism. And thank you to Iceni Glycoscience (formerly Iceni Diagnostics), the industrial sponsors for this studentship.

I would also like to thank every member of the Field group, past and present, who has encouraged and supported me throughout my time in the group. I would especially like to thank Dr Martin Rejzek for your valuable guidance and for teaching me your very thorough approach to experimental work. To Martin, Sergey, Giulia, Irina, Peterson and Sanaz – your expertise was invaluable, as were the many coffee breaks and timely lunch reminders! And to the Manchester contingent; Alex, Evaldas and Iakovia; the laughs, friendship and banana bread were a huge help during the final stretch, including the trials and tribulations of moving labs (twice!) and trying to complete a PhD during a global pandemic. A special thank you must also go to Iakovia for all her help with the NMR characterisation shown in this thesis – I couldn't have done it without you, quite literally. I hope to cross paths with you all again someday soon.

The unconditional support of my entire family for the course of my scientific career has been wonderful. Thank you, Mum and Dad, for all your support and encouragement, and for your expert dog-sitting during long days in the lab. Thank you also to my grandparents, Auntie Joyce and Grandma, to whom this thesis is dedicated. You encouraged me to pursue my academic career and always showed great enthusiasm for my research, even to the end.

And, finally, thank you to Joe. If not for this PhD, we never would have met.

Acknowledgement of Contributions

I would like to thank the following individuals for their contributions to this PhD thesis:

- Dr Irina Ivanova and Dr Rositsa Karamanska of the John Innes Centre, for chemically synthesising the BSA-mannose, BSA-2'FL and FITC glycoconjugates used in this PhD thesis.

- Dr Abubaker Mohamed of the Manchester Institute of Biotechnology and Department of Material, The University of Manchester, for conducting the mannose-functionalised biosensor studies with *Escherichia coli*.

Chapter 1

Chapter 1 – Introduction

1.1 Human Breast Milk

Human breast milk is widely regarded as the 'gold standard' diet for infants, providing complete postnatal nutrition during early neonate and further development, fined tuned over millions of years of evolution [2]. Breast milk contains many complex proteins, lipids and carbohydrates, the composition of which changes dramatically over the course of lactation in relation to the needs of the infant [3]. Contained within breast milk are a myriad of biologically active compounds that possess diverse roles aiding the development of the immune system and intestinal microbiota [4]. Human milk oligosaccharides (HMOs) are one such biologically active component that exert prebiotic effects, as well as acting as anti-adhesion agents and immune system modulators [5]. HMOs are non-digestible, complex carbohydrates [6] that are highly abundant in human milk (20-25 g/L in colostrum and 10-15 g/L in mature milk) being the third largest solid component after lactose and lipids [7, 8]. The composition of HMOs in breast milk is dependent on multiple factors including, maternal diet [9], genotype and lactation time course [4].

1.2 Human Milk Oligosaccharides (HMOs)

Over 200 different structures of human milk oligosaccharides (HMOs) have been identified in human breast milk [10]. HMOs consist of five monosaccharides: D-glucose (Glc), D-galactose (Gal), N-acetylglucosamine (GlcNAc), L-fucose (Fuc), and sialic acid (N-acetylneuraminic acid, Sia) in humans [3]. The reducing end is formed of lactose (Galβ1-4Glc) and can be elongated by the addition of β1-3- or β1-6-linked lacto-*N*-biose (Galβ1-3GlcNAc) to give a type 1 chain or the addition of N-acetyllactosamine repeat units (Galβ1-4GlcNAc) to give a type 2 chain, which can be further extended as seen in Figure 1A. Thus, a large number of core structures can be formed in either a linear configuration, *para*-HMO, or branched configuration, *iso*-HMO, introduced by a β1-6 linkage between two disaccharide units [11]. Fucosylation in α1-2, α1-3 or α1-4 linkage and/or sialylation in α2-3 or α2-6 linkage of the lactose or the elongated polylactosamine backbone gives further diversification of HMO structures [3] (see Figure 1.1B). As such, HMOs can be classified into three major categories: fucosylated neutral HMOs (35-50%), non-fucosylated neutral HMOs (42-55%) or sialylated acidic HMOs (12-14%) [12].
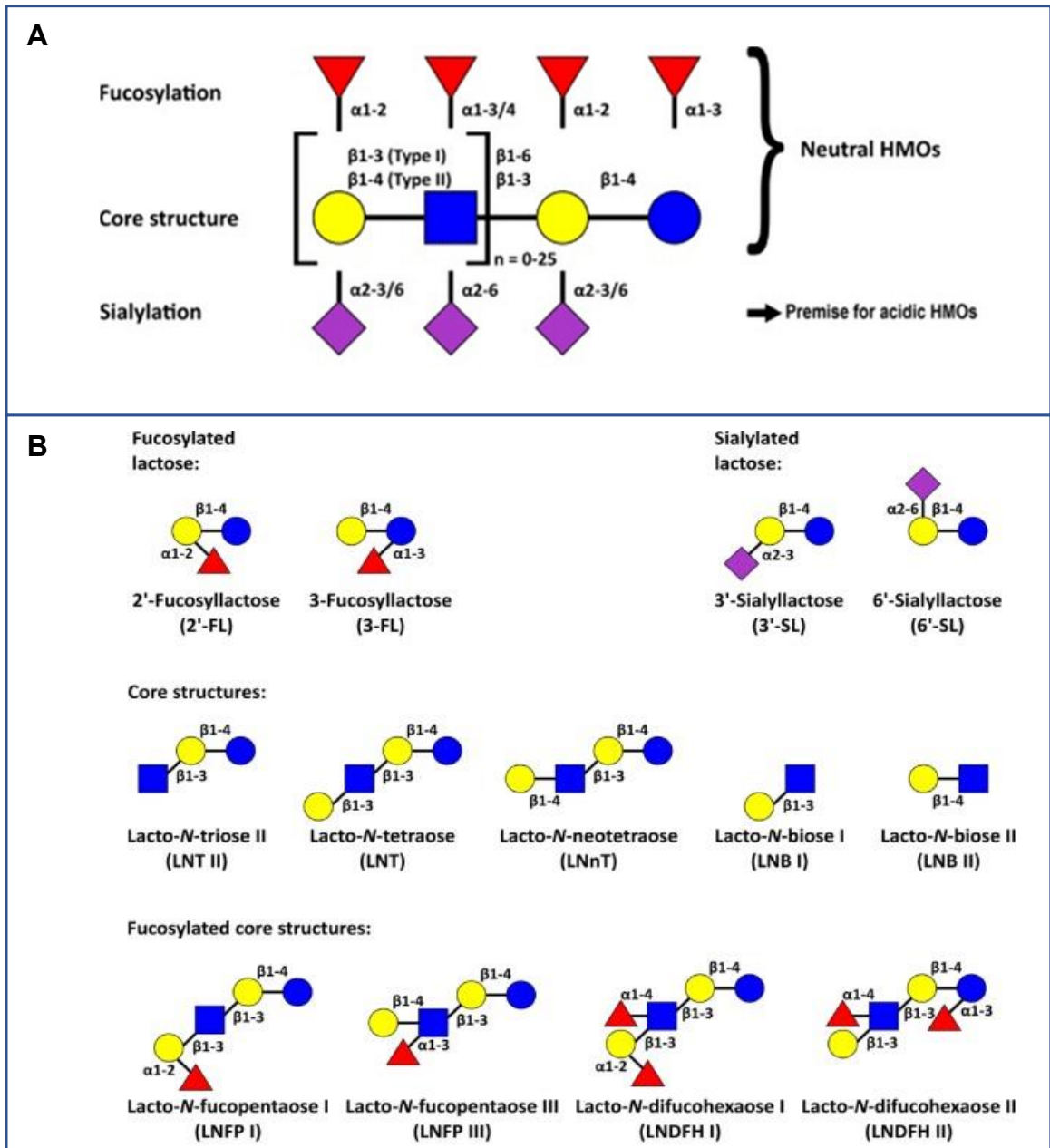
Figure 1.1: Structural diversity of HMOs. (**A**) The reducing end is formed of lactose (Galβ1-4Glc) and can be elongated by the addition of β1-3- or β1-6-linked lacto-N-biose (Galβ1-3GlcNAc) or N-acetyllactosamine repeat units (Galβ1-4GlcNAc). The polylactosamine backbone can then be further modified by the addition on fucose or neuraminic acid [13]. (**B**) Common HMO structures. Adapted with permission from [14].

Comparatively, 50-70% of total HMOs are fucosylated with only 10-20% sialylated structures present. However, the reverse is true for bovine milk oligosaccharides (BMOs) and porcine milk oligosaccharides (PMOs), where the majority of free milk oligosaccharides are acidic, sialylated structures (Table 1.1) [15, 16]. Novel milk oligosaccharides have also been found in the milk of the yak (grunniose and vakose), camels (medalose and ramelose) and buffalo (bubaliose) with rare linkages, such as α1-3Gal side chains and α1-3GalNAc at the reducing end Glc etc. [17-20] (see Figure 1.2).

Table 1.1: Comparison of free oligosaccharides in humans vs. bovine milk oligosaccharides (BMO) and porcine milk oligosaccharides (PMO) [15, 21].

	HMO	BMO	PMO
Number of oligosaccharides	200+	~40	~40
Percentage of fucosylation (%)	50-70	~1	1-3
Percentage of sialylation (%)	10-20	~70	~70
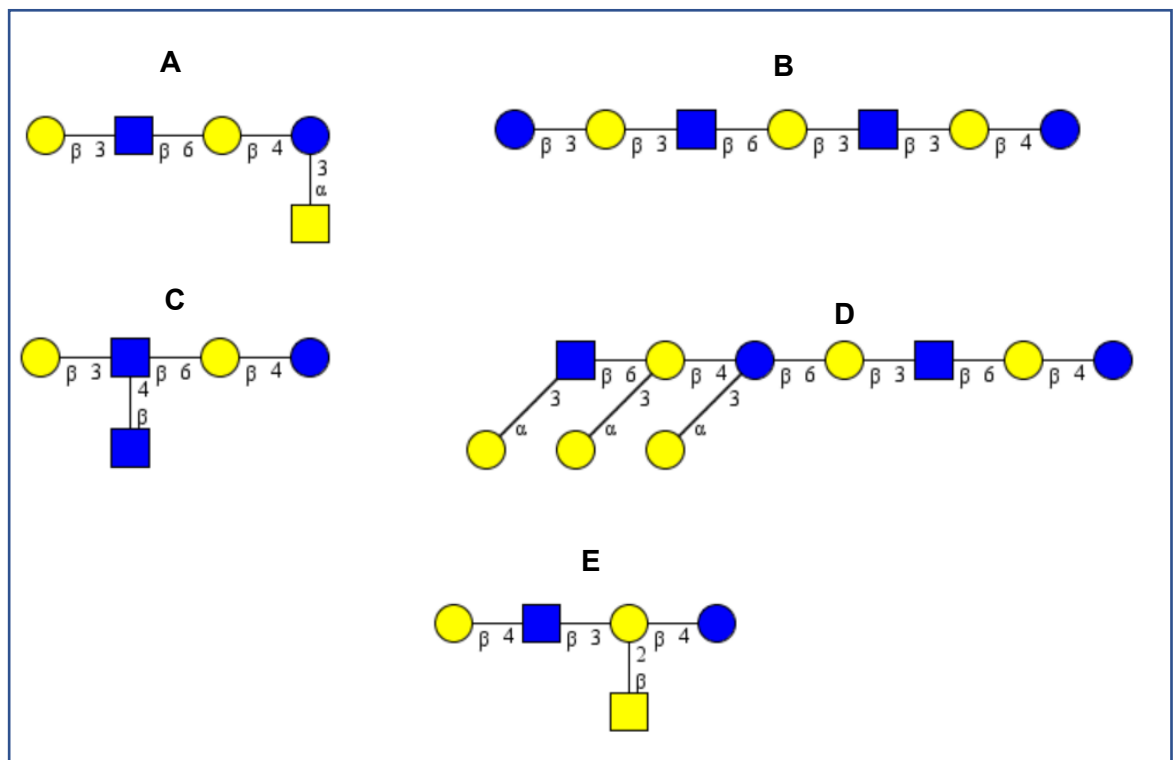
Figure 1.2: Structures of novel oligosaccharides found in animal milk. (A) Grunniose and **(B)** vakose from yak milk; **(C)** medalose and **(D)** ramelose from camel milk; and **(E)** bubaliose from buffalo milk.

1.3 HMO Profile

As previously mentioned, over 200 different structures of HMO have been identified in human breast milk [10], and the variety of structures present in a maternal mother's milk is dependent on her genotype, as well as the duration of lactation. HMOs are synthesised by the lactocytes, milk-secreting cells that line the alveoli of the human mammary gland [22], and the profile of HMOs produced by individuals is the result of the expression of specific transferase enzymes, α1,2-fucosyltransferase *(FUT2)* and α1,3/4-fucosyltransferase (*FUT3*). The Secretor gene (Se), *FUT2*, is responsible for the addition of α1,2-linked fucose in an HMO structure, while the Lewis gene (Le), *FUT3*, catalyses the addition of fucose via a α1,3/4-linkage. The different expression of these enzymes results in a four main phenotypes and subsequent HMO profiles: Se^+/Le^+, Se^-/Le^+, Se^+/Le^-, Se^-/Le^- (see Table. 1.2) [23]. It follows that not all mothers produce the same complement of HMOs. Se^+/Le^+ mothers can produce the full spectrum of HMO structures, whereas Se^-/Le^+ and Se^+/Le^- mothers are deficient in *FUT2* and/or *FUT3* respectively. However, women with a Se^-/Le^- profile that express neither *FUT2* nor *FUT3* still produce milk which contains fucosylated HMOs, like 3-fucosyllactose (3-FL, Galβ1-4(Fucα1-3)Glc) or lacto-*N*-fucopentaose III (LNFP III). This suggests that other Se- and Le-independent fucosyltransferases (FucTs) (*FUT4, 5, 6, 7* or *9*) may also be involved. Furthermore, α1,2-fucosylated structures found in the milk of non-Secretor women towards the end of lactation suggests a role for *FUT1* in HMO synthesis [24].

Table 1.1: Human milk oligosaccharide (HMO) secretor phenotypes. Taken from [38].

Genes	Lewis+	Lewis-
Secretor +	Se^+/Le^+	Se^+/Le^-
	Able to secrete all HMOs	Able to secrete 2'FL, 3FL, LNFPI, LNFPIII
Secretor -	Se^-/Le^+	Se^-/Le^-
	Able to secrete 3FL, LNFPII, LNFPIII	Able to secrete 3FL, LNFPIII, LNFPV

1.4 HMO Bioactivity

HMOs were originally identified as a "bifidus factor" in human milk, due to the presence of a bacterial isolate found predominantly in the faeces of breastfed infants, termed *Bacillis bifidus communis* by Tissier [25]. However, in recent years it has been increasingly apparent that HMOs are responsible for more than the growth of beneficial bacteria in the infant gut. Briefly, HMOs are known prebiotics due to their use as metabolic substrates by commensal bacteria, giving them advantageous growth over potential pathogens [26]. HMOs prevent pathogen attachment by acting as soluble glycan receptor decoys and directly modulate the gene expression of epithelial cells. They also modulate the immune system by altering lymphocyte cytokine production and reduce selectin-mediated cell-cell interactions by decreasing leukocyte rolling on endothelial cells leading to reduced mucosal leukocyte infiltration and activation. HMOs are also thought to provide an essential source of sialic acid for development and cognition in the brain (see Figure 1.3). Studies have also shown synergy of HMOs with some antibiotics against Group B Streptococcus [27, 28].
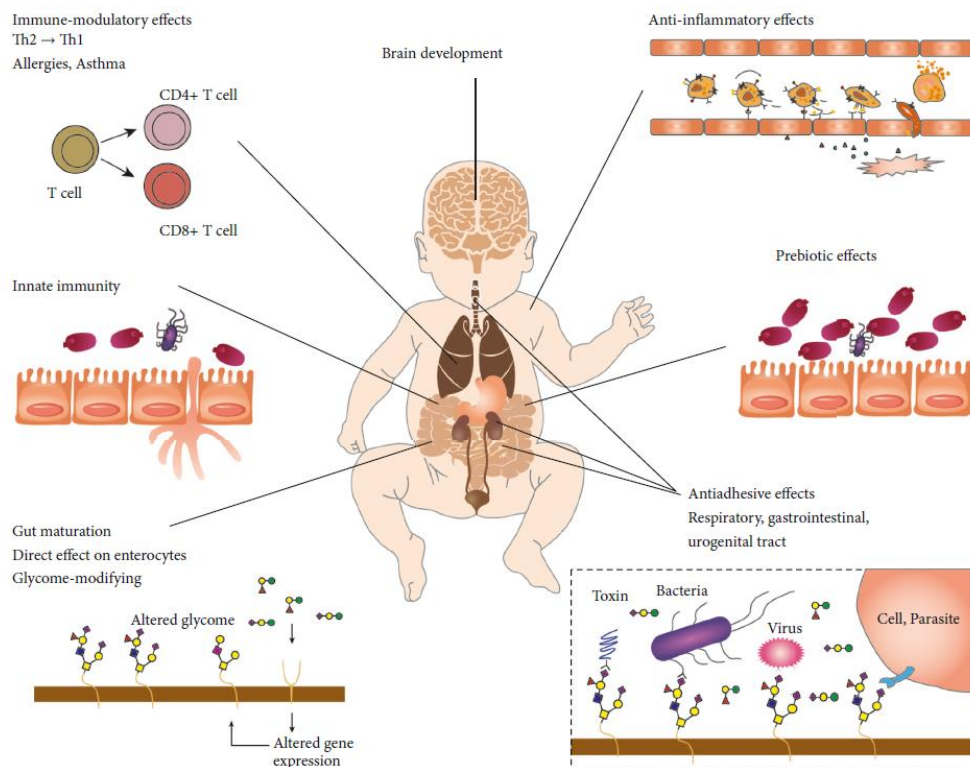
Figure 1.3: Schematic summarising the postulated beneficial effects of HMOs on the breastfed infant. HMOs serve as prebiotic agents, selectively enriching bifidogenic gut bacteria. They have also been shown to act as antivirals and antimicrobials due to their antiadhesive properties. HMOs interact with immune cells and affect the expression of pro/anti-inflammatory cytokines. Expression of proteins at the glycocalyx and mucus layers is also altered by HMOs and their metabolites. HMOs are also though to aid brain development and cognition by providing a source of sialic acid. Reproduced with permission from [29].

1.4.1 Prebiotics

HMO are probably best known for their prebiotic effects in breast-fed infants, where they exert a strong bifidogenic effect, characterised by the proliferation of specific strains including *Bifidobacterium infantis, B. breve* and *B. bifidum,* to shape the infant's gut microbiota by selectively stimulating the growth of specific bacteria, or commensal bacteria (see Figure 1.4) [30, 31]. The concept of prebiotics was first introduced in the 1990s by Gibson and Roderbroid, who defined prebiotics as "a non-digestible food ingredient that beneficially affects the host by selectively stimulating the growth and/or activity of one or a limited number of bacteria in the colon, and thus improves host health" [32]. Since then, the accepted definition has been revised but the premise remains the same. To qualify as a prebiotic, compounds must comply with several criteria: they must be resistant to gastric activity, enzyme hydrolysis and gastrointestinal absorption. Due to their low absorption rate (~1%) and high concentrations of intact HMOs found in the distal small intestine and colon, HMOs qualify as prebiotics [33, 34]. What's more, evidence suggests that breast-fed infants have a higher abundance of beneficial bifidobacteria compared with formula-fed infants [35].
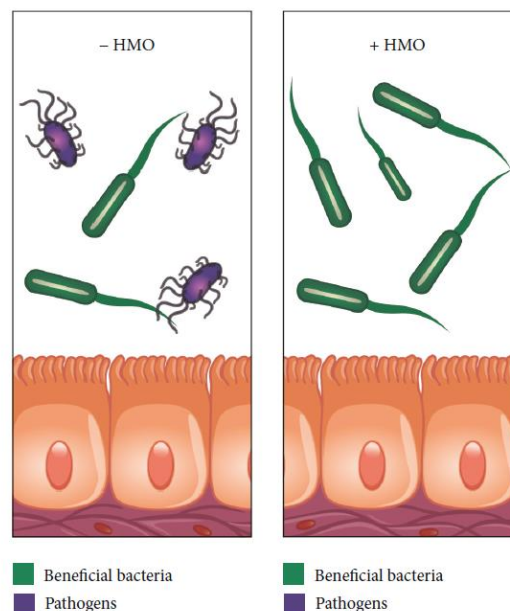
Figure 1.4: Illustration of the prebiotic effect of HMOs on the gut microbiota. In presence of HMOs, the growth of beneficial commensal bacteria is selectively stimulated and preventing the growth of pathogens. While in the absence of HMOs, there is an increased presence of pathogens. Taken with permission from [29].

Notably, *Bifidobacterium bifidum* was shown to grow on human milk glycan fractions [36]. This trend was then observed in *B. longum* subsp. *infantis,* which was shown to grow with HMOs as the only carbon source and ultimately consume mono- and disaccharide degradation products [37]. Furthermore, other *Bifidobacterium* species, such as *B. breve* and *B. lactis,* have been confirmed to selectively consume fucosylated and sialylated HMOs [38, 39]. *B. breve* also metabolises lacto-*N*-tetraose (LNT; Galβ1-3GlcNAc-Lac, type I), but not lacto-*N*-neotetraose (LNnT; Galβ1-4GlcNAcLac, type 2) [11].

In order to utilise HMOS, the single monosaccharides need to be released, and the breakdown of each linkage requires specific enzymes that exist in microbes [40]. First, the fucose and sialic acid "decorations" are removed by α-L-fucosidases and sialidases so that the core structure can be accessed by β-galactosidases and β-N-acetylglucosaminidases [41-43]. The genome sequence of *B. longum* subsp. *infantis* also revealed entire gene clusters evolved for milk utilisation within the infant microbiome, controlling the expression of glycosidases, transporters and glycan-binding proteins. In a single 43kbp gene cluster four glycosyl hydrolases were found, which are predicted to cleave milk oligosaccharides into its constituent monosaccharides: a sialidase, a fucosidase, an N-acetyl-β-hexosaminidase and a β-galactosidase [44]. Full characterisation of *Bifidobacterium* strains and genes, which utilise HMOs can be found in Appendices 1.1.

Although *Bifidobacterium* and *Bacteroides* are the most comprehensively studied genera of gut colonisers, they are not the only species that utilise HMOs. *Lactobacillus, Escherichia, Klebsiella, Enterococcus, Staphylococcus*, and *Clostridium* have also been reported to utilise HMOs, though the data is sporadic and strain specific [45, 46]. Many infant gut commensals, considered non-HMO utilizers, have also been tested for growth on more abundantly present HMOs. *Clostridium perfringens, Escherichia coli, Enterococcus faecalis,* some *Lactobacillus species (L. acidophilus, L. casei, L. fermentum, L. plantarum, L. rhamnosus, L. salivarius),* and *Staphylococcus epidermidis* and *Streptococcus thermophilus* have been used in growth curve experiments where they sustain moderate or little growth on HMOs, such as 2'FL, 3FL and LNnT [47-52]. However, strains that cannot utilise HMOs can often metabolise the digestion products, i.e., monosaccharides and disaccharides (Fuc GlcNAc and Neu5Ac), released into the environment by other HMO utilising species.

Metabolites from HMO digestion by microbes can also influence intestinal health. The release of short-chain fatty acids (SCFAs) can mediate beneficial effects by lowering the pH of the intestinal environment, which prevents the growth of pathobionts [53, 54].

1.4.2 Decoy Receptors

HMOs also directly affect the ability of pathogens to invade and colonise the infant gut epithelial by acting as antiadhesive antimicrobials, in addition to indirectly providing a competitive advantage to non-pathogenic commensals [3]. Adherence to mucosal surfaces is often vital for viral, bacterial and protozoan pathogens to colonise the host and cause disease and, frequently, this is initiated by lectin-glycan interactions. This is supported by the existence of a vast range of different lectin families (see Table 1.3) and can be specific to different species and strains of microorganism [55, 56]. HMOs prevent the attachment of pathogens to the intestinal epithelial surface by acting as receptor decoys by mimicking the cell surface glycans (see Figure 1.5).

Studies using *in vitro* cultured epithelial cells have shown that *Campylobacter jejuni* binds type 2, α1,2-fucosylated H-antigens and soluble α1-2-fucosylated HMOs block binding of *C. jejuni* to cultured cells in human intestinal mucosa explants and reduce colonisation in mice [57, 58]. Furthermore, in a study examining episodes of *C. jejuni*-associated diarrhoea in 100 mother-infant pairs, there were significantly fewer cases of *C. jejuni*-associated diarrhoea in infants who received high concentrations of 2'FL from the nursing mother. Similarly, where the mother's milk contained high concentrations of lacto-*N*-difucohexaose I (LNDFH-I), another α1,2-fucosylated HMO, cases of calicivirus diarrhoea were also less frequently observed [59].

Additionally, HMOs also appear to offer protection to nasopharyngeal regions and the upper respiratory tract, with breast-fed infants having a lower risk of developing infections caused by *Streptococcus pneumoniae, Pseudomonas aeruginosa, Haemophilus influenzae* and respiratory syncytial virus (RSV). This is due to human milk covering mucosal surfaces of the infant's nasopharyngeal regions and occasionally the upper respiratory tract during episodes of aspiration [60, 61]. Furthermore, HMOs are absorbed and excreted in the urine and sialylated HMOs have been shown to reduce hemagglutination induced by uropathogenic *E. coli* [62].

Viruses also employ this mechanism of glycan-mediated attachment [63, 64]. HMOs serve as soluble decoy receptors due to their resemblance of mucosal cell surface glycans, thus preventing pathogens from binding and reducing the risk of infection [65]. Similarly, viruses also express glycans to bind to host lectins. They are known glycan-virus binding interactions between HMOs and influenza, rotavirus, RSV, human immunodeficiency virus (HIV), and norovirus [66]. Rotavirus has been shown to bind H1 antigens and its precursor, Lacto-N-biose, via its VP4 spike protein [67] and the viral capsid of norovirus binds the structural mimics of human blood group antigens, 2'FL and 3FL [68]. While HIV, expresses

the envelope glycoprotein gp120 to facilitate binding to DC-SIGN (dendritic cell-specific ICAM3-grabbing non-integrin) on human dendritic cells [69-71]. This interaction is important for HIV transmission from mother-to-child during breastfeeding by aiding HIV entry via mucosal barriers [72] and could explain the surprisingly low (10-20%) mother-to-child transmission of HIV when breastfeeding [73]. These findings have accelerated the research into the antiviral potential of HMOs, including for the treatment of COVID-19 in recent months [74, 75].

Table 1.3: Examples of bacterial adhesins with glycans relevant to systems effected by HMOs. Adapted from [76].

Microorganism	Adhesin	Glycan-Receptor Specificity	Site of Infection
Bordetella pertussis	Filamentous hemagglutinin (FHA)	Sulfated glycolipids, heparin	Ciliated epithelial in respiratory tract
Campylobacter jejuni	Flagella, LPS	H antigen	Intestinal cells
Escherichia coli	P fimbriae	Galα1-4Galβ-	Urinary tract
	Type-1 fimbriae	Manα1-3(Manα6Manα1-6) Man	Urinary tract
	K99 fimbriae	Gangliosides GM3, Neu5Gcα2-3Galβ1-4Glc	Intestinal cells
Haemophilus influenza	HMW1 adhesin	Neu5Acα2-3Galβ1-4GlcNAcβ-, heparan sulphate	Respiratory epithelial
Helicobacter pylori	BabA	Sialyl Lewis x	Stomach
	SabA	Lewis b	Stomach/duodenum
Mycobacterium tuberculosis	Heparin-binding hemagglutinin adhesin (HBHHA)	Heparan sulphate	Respiratory epithelial
Neisseria gonorrhoeae	Opa proteins protein	LacCer; Neu5Acα2-3Galβ1-4GlcNAcβ-, syndecans, heparan sulphate	Genital tract
Pseudomonas aeruginosa	Type IV pili	Aisalo GM1 and GM2	Respiratory tract
Streptococcus pneumoniae	Carbohydrate-binding modules of β-galactosidase, BgaA	Lactose or *N*-acetyl-lactosamine	Respiratory tract
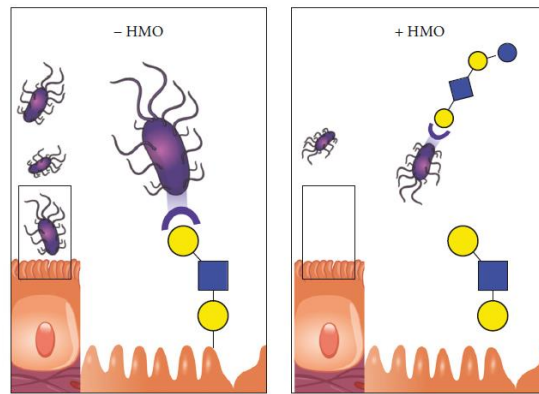
Figure 1. 5: Illustration of the antiadhesive properties of HMOs. They directly reduce microbial infections by acting as antiadhesive antimicrobials to prevent pathogen attachment at the intestinal cell surface. Taken with permission from [29].

1.4.3 Intestinal Epithelial Cell Modulation

The intestinal epithelium forms a functional barrier between the circulatory system and the gut's luminal content and is, therefore, critical for the passage of nutrients and macromolecules, and prevention of paracellular translocation of bacteria [77]. The intestinal barrier is comprised of a mucus layer covering the intestinal epithelia, which acts to separate the gut microbiota from the submucosa. Intercellular junctional complexes that maintain epithelium barrier function regulate the entry of luminal nutrients, ions and water. Microbes can indirectly alter the epithelial permeability through effects on host immune cells and the release of cytokines [78]

Observation from *in vitro* studies suggest that HMOs directly interact with intestinal epithelial cells to affect gene expression, which subsequently affects cell surface glycosylation and reprograms the cell cycle leading to modulation of cell growth, differentiation, and apoptosis [79, 80]. Incubation of human intestinal epithelial cell lines with 3'-SL lowers the expression of sialyltransferases ST3Gal1, ST3Gal2 and ST3Gal4 leading to decreased α2,3- and α2,6-sialylation and lower enteropathogenic *E. coli* binding [81]. Additionally, a series of experiments by Kunz *et al.,* revealed that incubation of intestinally transformed HT-29 cells with neutral and acidic HMOs induces differentiation while neutral oligosaccharides strongly induce apoptosis [79]. It does this by increasing mucin production and enhancing barrier function and the cellular level, coupled with modulation of the immune system by promoting a regulatory T cell enriched gut-associated lymphoid tissue (GALT) population [53, 54].

1.4.4 Immune System Modulation

HMOs have been detected in faeces and urine of breasted infants, but also directly in peripheral blood [82-84]. HMOs are thought to directly modulate the immune response by either acting locally on cells of the mucosa-associated lymphoid tissues or on a systemic level. Sialylated HMOs promote lymphocyte maturation and balanced Th1/Th2-cytokine production to achieve low level immunity [85]. Furthermore, 3'SL and 6'SL have been shown to enhance bacterial clearance in *Pseudomonas aeruginosa K* (PAK)-infected mouse model and promote phagocytosis of PAK. They are also effective at promoting phagocytosis of bacteria, intracellular ROS generation and lysosomal activity [86].

However, it is currently unknown which receptors and signalling pathways transduce HMO-mediated effects on lymphocyte cytokine production or macrophage stimulation. The immune system utilises several different types of lectins and HMOs may interfere with specific lectin-glycan binding specificities. For example, selectins are C-type lectins and cell adhesion molecules that mediate the leukocyte trafficking and platelet– neutrophil complex (PNC) formation, which mediate leukocyte migration from the blood stream to subendothelial regions of inflammation [87]. In an *in vitro* flow model with TNFα-activated human endothelial cells, sialylated HMOs reduce leukocyte rolling and adhesion and therefore, reduce inflammation [88].

1. 5 Challenges and Opportunities for the Study of HMOs

The plethora of bioactive properties that HMOs offer gives a huge range of opportunities for perspective interventions and treatments: from infant nutrition to adult diet supplementation, glycan-based probes, new anti-infective strategies and use in antibiotic cocktails to potentiate the effect of antibiotics to combat antibiotic resistance. The discovery of HMO protection against rotavirus and norovirus infections can provide an alternative to current therapies.

The greatest barrier to research in the field of HMO glycobiology is the limited availability of HMOs. They are currently sourced from human milk samples, synthesised chemically, enzymatically, or by microbial cell factories. Firstly, there is a limited supply of donor milk and the isolation procedures employed to purify individual oligosaccharides from pooled samples is both tedious and labour intensive [89]. As such, many studies use pooled milk samples to draw conclusions about the bioactive properties of HMOs. A failing of this

approach is that specific structure-activity relationships cannot be elucidated without single-entity compounds. This is evident in a recent example where pooled HMOs were found to reduce incidence of NEC in a neonatal rat model, however, after further investigation it was revealed that this protection was the sole action of disialyllactose-*N*-tetraose (DSLNT) [90].

There are also low concentrations of milk oligosaccharides in mammals (0.7 to 1 g/L in bovine colostrum compared to 5 – 20 g/L in human colostrum) and their oligosaccharide structures are less diverse and complex, even in human primates [91, 92]. It follows that the use of animal models in preclinical studies does not translate well to humans as they produce a different panel of oligosaccharides and are therefore, not exposed to HMOs, which leads to differences in response [89]. Complex HMO structures are also extremely difficult to access chemically; requiring organic solvents, toxic reagents, poor atom economy, and complex and lengthy protecting and de-protecting steps to achieve chemoselectivity and regioselectivity. Advances in chemoenzymatic synthesis of HMOs are needed create structurally defined HMOs and HMO libraries to facilitate initial research studies for identification of targets. A better understanding of how HMOs are produced in the human mammary gland would guide chemoenzymatic synthesis and inform the bioengineering process. Bioengineered microorganisms also have the advantage of nearly ceiling less scale-up opportunities [89].

1.6 HMO Biosynthesis

Relatively little is understood about the synthesis of HMOs, and little is known about how the difference in fucosylation, and sialylation occurs between human and animal milk. It is known that the lack of fucosylated milk oligosaccharides in other non-human milk is also due to a lack of *FUT* expression. Thus, there is an enhanced selective advantage arising from milk oligosaccharide content due to increased complexity, variability and concentration of HMOs compared to other species [16]. One might speculate that fucosylated HMOs are integral to human evolution and perhaps relate to long childhood and ex-utero brain development of infants compared to great apes and other mammals. For example, the brain volume of a chimpanzee foetus is only half that of a human foetus at 16 weeks of gestation and shows greater acceleration in brain growth [93].

HMO biosynthesis takes place in the Golgi of mammary gland epithelial cells and is likely an extension of lactose biosynthesis, where the UDP-Gal:N-acetylglucosamine β1-4-

galactosyltransferase (GalT) is in complex with α-lactalbumin, referred to as 'lactose synthase', transfers UDP-Gal to GlcNAc moieties to form reducing end lactose [94, 95]. Following the formation of the terminal lactose, elongation and branching occur via a variety of transferase enzymes, which are subsequently sialylated and/or fucosylated much like the synthesis of human blood group antigens. Fucosylation is regulated by the expression of fucosyltransferase (*FUT*) enzymes [96], where *FUT2* encodes expression of α1,2-fucosyltranserfase and FUT3 encodes α1,3/4-fucosyltransferase expression. However, the order of biosynthetic events is largely unknown.

1.6.1 Enzymatic Synthesis of HMOs

Due to their high structural diversity and complexity, as a result of the high number of possible monosaccharide combinations and linkages, the identification and quantification of HMOs pose a challenge to researchers. However, the correct identification of HMOs is essential, as their specific structures lead to different biological effects [97, 98]. Therefore, strides have been taken in recent years, towards the chemical and enzymatic synthesis of HMOs.

Carbohydrates are polyhydroxy compounds, which contain large numbers of the same functional groups; therefore, their chemical synthesis is notoriously difficult, requiring organic solvents, toxic reagents, high atom economy, and complex and lengthy protecting and de-protecting steps to achieve chemoselectivity and regioselectivity. Additionally, glycosyl donors can be connected by either α or β glycosidic linkages, which creates added complexity to the synthesis. This is poignant when we consider that anomeric mixtures cannot be used directly for biological research [99]. In recent years, the Seeberger group have developed an 'automated glycan assembly' – based on the solid-phase synthesis paradigm like automated polynucleotide synthesis, which allows access to polysaccharides of up to 100 monosaccharide units via a 201 automated iterative cycles of coupling, capping and deprotection [100].

Fortunately, enzymatic methods for the synthesis of HMOs have been proving effective. Unlike chemical methods, they can be done in mild, aqueous solutions on non-protected substrates and catalyse glycosylations in a highly selective manner [101]. Many different enzymatic approaches have been investigated, including the engineering of various model host strains such as *E. coli* [102-104] and yeast [105] to create cell factories via whole cell microbial transformations and loop engineering [106, 107]. The production of 2'FL has been approved to be produced by microbial fermentation since 2015 and is regarded,

along with LNnT as safe for incorporation to infant formula by the US Food and Drug Administration (FDA) and European Food Safety Authority (EFSA) (see Figure 1.6) [108-110]. However, these approaches are a very young technology and, to date, have mainly been used for the production of 2'FL, 3FL, LNT and LNnT, and more recently LNFPIII and lacto-*N*-fucopentaose V (LNFPV) [111]. As such, they are limited by the efficiency of lactose import and metabolism, balancing heterologous gene expression and cell growth and lack of knowledge regarding key enzymes of the biosynthetic routes to more complex HMOs [108]. Multi-enzyme cascades and modular enhancement strategies have also been improved for enhanced efficiently of α1,2/3-fucosyltransferases [112-115].
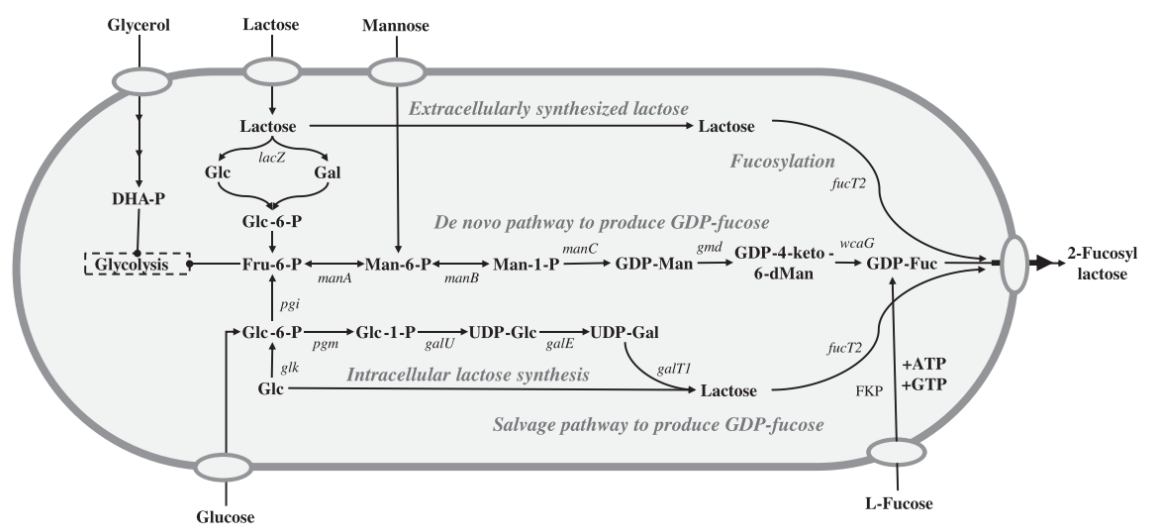
Figure 1.6: Gene targets of metabolic engineering to produce 2'FL in E. coli. Three elements for producing 2'FL in *E. coli* are illustrated. Either extracellularly fed or intracellularly synthesized lactose can be used as a backbone. De novo and/or salvage pathways for producing GDP-L-fucose can be used. Fucosyltranferases from various sources can be used for fucosylation of lactose. Reproduced with permission from [89].

Fucosyltransferase enzymes are responsible for the addition of fucose residues at specific locations and via specific linkages during glycan synthesis. To date, thirteen human fucosyltransferase genes have been cloned and characterised (see Table 1.4). Mammalian fucosyltranferases catalyse the transfer of fucose from a donor substrate, GDP-L-fucose, to oligosaccharide chains via an α1,2- α1,3-, α1,4- or α1,6-glycosidic linkage [116]. While α1-3-fucosyltransferases have been more extensively studied for their use in the synthesis of Lewis antibodies [115, 117-121], by comparison the α1-2-fucosyltransferases are largely understudied.

Table 1.4: Genes and enzymes involved in the fucosylation of fucose-containing glycans in humans.

Gene	Gene ID	Enzyme	Linkage	Product	Ref
FUT1	HGNC:4012	Fucosyltransferase 1 (H blood group)	α1,2	ABH Lewis antigens	[122, 123]
FUT2	HGNC:4013	Fucosyltransferase 2	α1,2		[122, 123]
FUT3	HGNC:4014	Fucosyltransferase 3 (Lewis blood group)	α1,3/4	Lewis antigens	[123, 124]
FUT4	HGNC:4015	Fucosyltransferase 4	α1,3		[123, 124]
FUT5	HGNC 4016	Fucosyltransferase 5	α1,3/4		[123, 124]
FUT6	HGNC:4017	Fucosyltransferase 6	α1,3		[123, 124]
FUT7	HGNC:4018	Fucosyltransferase 7	α1,3		[123, 124]
FUT8	HGNC:4019	Fucosyltransferase 8	α1,6	N-glycan core fucosylation	[125, 126]
FUT9	HGNC:4020	Fucosyltransferase 9	α1,3		[123, 124]
FUT10	HGNC:19234	Fucosyltransferase 10	α1,3		[127]
FUT11	HGNC:19233	Fucosyltransferase 11	α1,3		[127]
POFUT1	HGNC:14988	Protein O-fucosyltransferase 1		Polypeptides fucosylation	[128-130]
POFUT2	HGNC:14683	Protein O-fucosyltransferase 2			[128-130]

1.6.2 α1,2-Fucosyltransferases

The glycosyltransferase family 11 includes the human α1,2-fucosyltransferase (α1,2-FucT) genes (*FUT1* and *FUT2*) [131] and a large number of putative α1,2-FucTs with yet unknown acceptor specificities. All α1,2-FucTs add a fucose residue via an α1,2-linkage but different species have unique substrate specificities (see Figure 1.7). For instance, α1,2-FucT from humans has been shown to have a preference towards type 2 LacNAc (βGal1-4GlcNAc) or Lewis antigen, Lewis X (βGal1-4(αFuc1-3)βGlcNAc), whereas α1,2-FucT from *H. pylori* prefers type 1 glycans such as LacNAc (βGal1-3GlcNAc) or Lewis A (βGal1-3(αFuc1-4)βGlcNAc) as well as a wide variety of type 1 oligosaccharide substrates [132]. The same is true of α1,2-FucTs enzymes from enteropathogenic *E.* coli O128, O86

and O127, which all display preferences towards type 1 lacto-*N*-biose or mucin-type core 1 O-glycan, respectively [133].
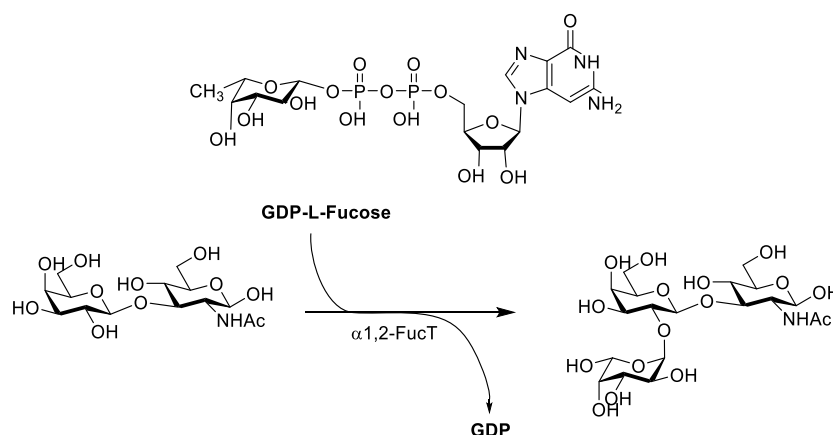
Figure 1.7: Reaction mechanism of α1,2-fucosyltranferases. Catalyses the transfer of L-fucose, from GDP-L-fucose to the terminal galactose of both O- and N-linked glycans.

The α1,2-FucT from *Helicobacter mustelae* (HmFucT) that we will use for this thesis is commercially available, however, there is no published data about its substrate specificity or crystal structure. Despite this, α1,2-FucT from *H. pylori* (HpFucT) and HmFucT share 77.9% sequence homology therefore, it can be assumed that they will demonstrate similar substrate specificity due to remarkably similar AlphaFold structures (see Figure 1.8A). No structural information is available for the α1,2-FucTs, including annotation of the catalytic site, however the catalytic site is conserved as among other fucosyltransferases. ConSurf analysis revealed a highly conserved region in a pocket of the protein – the likely location of the catalytic site (see Figure 1.8B) [134-139].

A *H. pylori*   *H. mustelae*

B

1	2	3	4	5	6	7	8	9

Variable   Average   Conserved

Insufficient Data
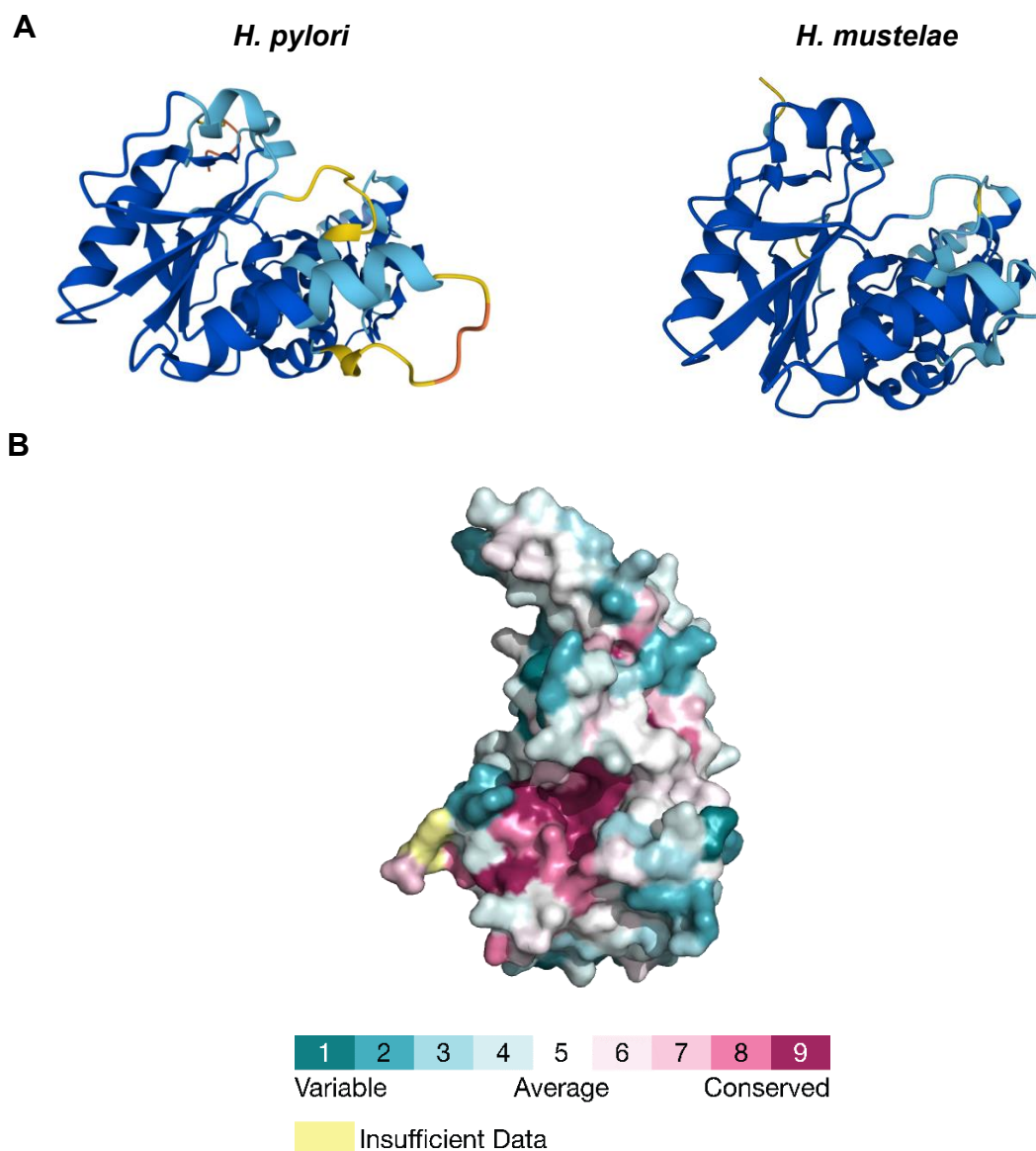
Figure 1.8: Summary of α1,2-fucosyltranferase structure for *H. pylori* and *H. mustelae* putative enzymes. AlphaFold structure prediction of (**A**) HpFucT (left) and HmFucT (right) [140, 141]. (**B**) PyMOL representation of HmFucT showing the protein surface from AlphaFold with the evolutionary conservation of amino acid residues from ConSurf [134-139].

1.6.3 L-Fucokinase/GDP-L-Fucose Phosphorylase (FKP)

Although fucosyltranferases are useful tools for HMO synthesis, the donor required for the enzymatic reaction to occur, GDP-L-fucose, is very expensive commercially meaning that use of a commercial donor makes the scalable synthesis of HMO analogues unachievable. Therefore, we must consider an alternative, cost-effective approach to the acquisition of a suitable donor. One such method is to utilise the bi-functional enzyme, L-fucokinase/GDP-L-fucose phosphorylase (FKP) and use it in a one-pot reaction with the α1,2-

fucosyltransferase to simultaneously catalyse the formation of the donor, which can then be used to catalyse the fucosylation of HMO analogues in one step.

GDP-L-fucose is synthesised via two different metabolic pathways *in vivo*: the de novo pathway and the salvage pathway. The de novo pathway is constitutively active and evolutionarily conserved and involves the conversion of GDP-D-mannose to GDP-L-fucose via three enzymatic reactions. On the other hand, the salvage pathway involves the scavenging of free L-Fuc from the diet from which the enzymes L-fucokinase and GDP-L-fucose pyrophosphorylase can catalyse the formation of GDP-L-fucose (as illustrated in Figure 1.6). FKP is found in the salvage pathway of *Bacteroides fragilis,* which mammals harbour as a symbiont in their intestines, for the formation of GDP-L-fucose from L-fucose [142]. The overall structure of FKP exhibits a tetrameric organisation, where each monomer is composed of an N-terminal GDP-fucose pyrophosphorylase domain and a C-terminal fucokinase domain joined by a linker (see Figure 1.9).
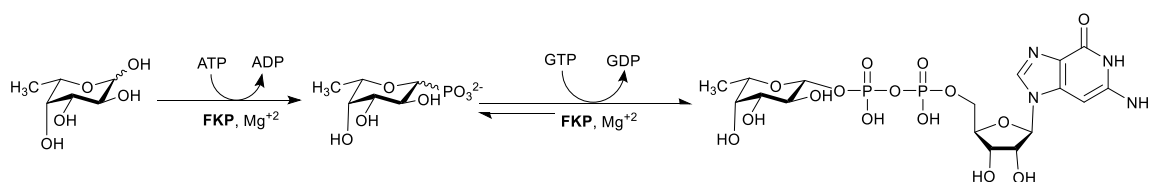
Figure 1.9: Formation of the sugar nucleotide GDP-L-fucose by FKP from *Bacteroides fragilis*. Adapted from [143].

Previous studies have exploited the fucose salvage pathway using FKP for the incorporation of non-natural analogues into fucosylated glycans. In 2006, the Bertozzi and Wong groups incorporated azido and alkynyl functionalised at position 6 on fucose giving rise to non-natural analogues for the synthesis of carbohydrate probes [144, 145]. The kinetic parameters of FKP catalysis of non-natural C-5 derivatised fucose analogues was described in 2009 and were generally well-tolerated (see Figure 1.10). The largest loss in catalytic efficiency (~43-fold) was seen with removal of the C-5 methyl group, which indicated that the interaction between the enzyme active site and the C-5 methyl group of L-fucose was found to be critical for catalytic efficiency. The alkyne-, alkene-, and fluorine-substituted analogues were phosphorylated by FKP almost as efficiently as the natural substrate L-fucose. [146].
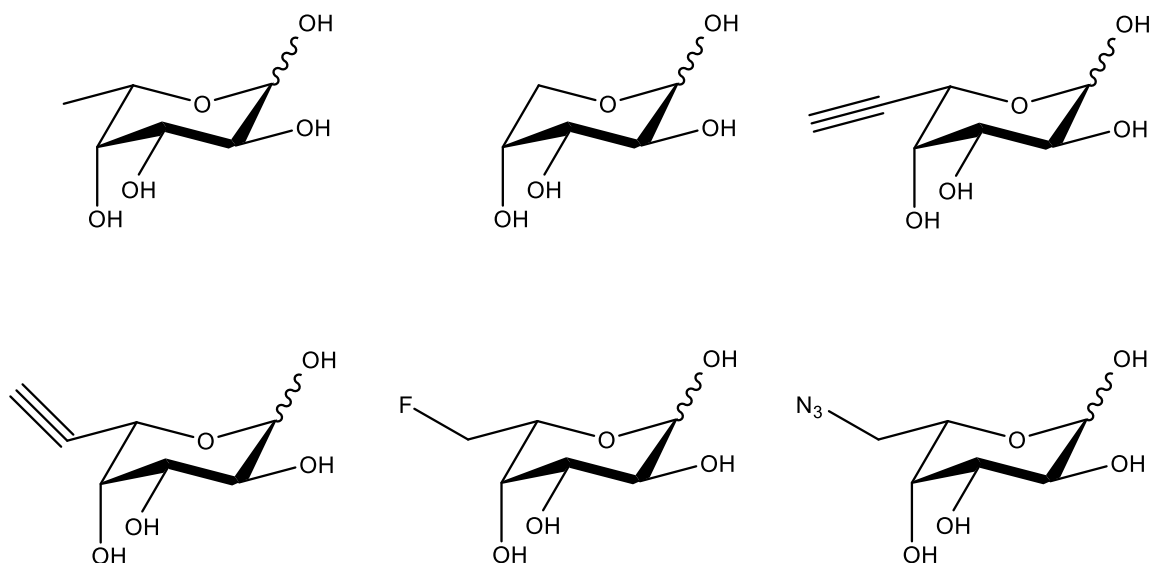
Figure 1.10: Fucokinase activity of *B. fragilis* FKP with C-5 substituted fucose analogues. *B. fragilis* FKP fucokinase activity and specificity were determined in the presence of 5 mM ATP using a simple coupled enzymatic assay with spectrophotometric readout. The production of ADP by B. fragilis FKP is coupled to the oxidation of NADH in the presence of pyruvate kinase and lactate dehydrogenase. The consumption of NADH, and accordingly the fucokinase activity, is monitored by the change in absorbance at 340 nm [146].

1.7 Fucose in Nature

Fucose is a deoxyhexose; an unusual sugar that is present in a variety of glycolipids and glycoproteins produced by mammalian cells and found in many of the animal kingdoms [147]. It is abundant on host mucosal surfaces, where it decorates terminal positions of mucin carbohydrate ligands; microbial cell surface structures and is found in some dietary sources [1]. Fucosylated glycans are involved in a variety of physiological processes including immunity, brain development, cancer and host-microbe interactions in humans [147].

Perhaps the most well-known fucosylated glycans are the ABO blood group antigens, encoded by fucosyltranferases *FUT1* and *FUT2* to synthesise the H-antigen (see Figure 1.11) [148]. The blood group antigens also have an important role in mediating host-microbe interactions as L-fucose is found abundantly in the mammalian gut [149]. For example, *Helicobacter pylori* is able to utilise host expression of Lewis[b] antigen to recognise and attach to gastric epithelial tissue, leading to peptic ulcers and gastric cancer. This method of pathogen adhesion is also used by *Vibrio cholera*, rotavirus and norovirus [67, 150, 151], which bind to Lewis[y] [152], H-antigens [67], and 2'FL and 3FL [68], respectively.
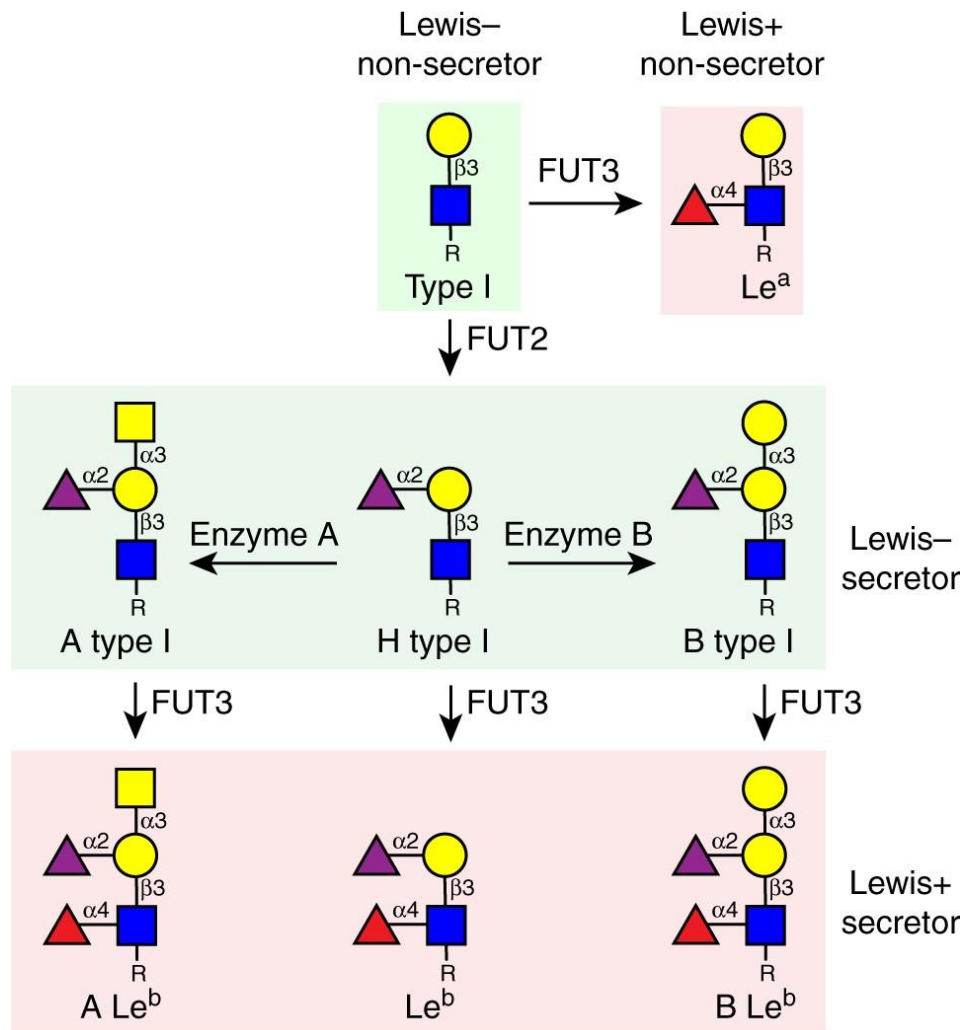
Figure 1.11: Biosynthesis pathway of human ABO blood group antigens. Reproduced with permission from [153].

Fucosylated glycans are also displayed by both commensal bacteria and pathogens to mimic host glycans and evade the immune system [149]. The activity of *FUT2* also promotes normal microbial gut diversity and composition [154], while its upregulation serves as a protective mechanism to increase tolerance to infection and maintain host-microbiome symbiosis [155, 156]. Fucosylated oligosaccharides also expedite the establishment of a healthy microbiota and provide protection from infection [157, 158].

Many fucosylated glycans on glycoproteins serve as important cancer biomarkers [147]. Change in fucosylation of glycoproteins such as fucosylated alpha-fetoprotein (AFP) is one of the most representative types of glycan-related cancer biomarkers [159]. Sialyl Lewis antigens are also upregulated in a variety of cancer types and serve as prognostic factors for increased risk of metastasis [160]. Zhang *et al.,* demonstrated that by preventing terminal fucosylation by knocking down *FUT1* and *FUT4*, tumour growth was

inhibited [161]. Additionally, in studies that aim to interfere with the synthesis of the fucosyltransferase donor, GDP-L-fucose, fucose analogues 2-fluoro-L-fucose and alkynyl-fucose were able to inhibit GDP-fucose synthesis and therefore liver cancer HepG2 cell proliferation, migration and tumour formation, and hepatoma cell migration and invasion, respectively [162, 163] (see Figure 1.12).
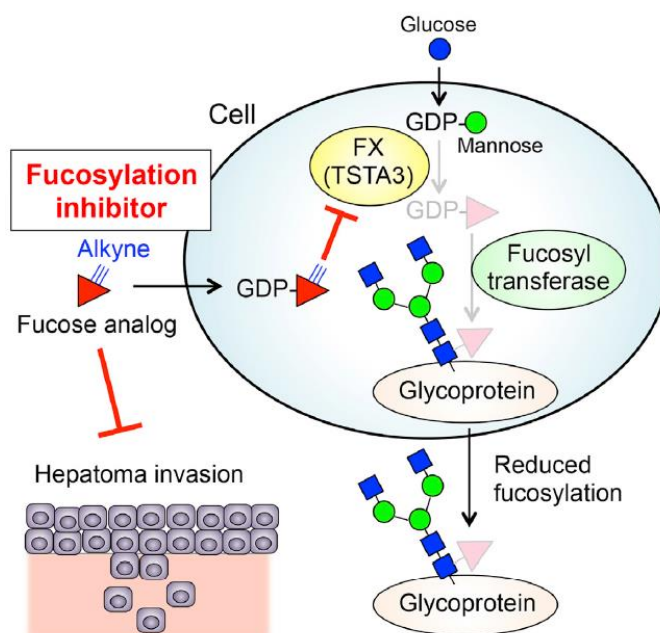
Figure 1.12: Illustration of how fucosylation inhibitors, alkynyl-fucose and 2-fluoro-L-fucose, elicit reduced glycoprotein fucosylation and halts hepatoma invasion. Alkynyl-fucose, a widely used fucosylation probe, strongly inhibits fucosylation by competitively inhibiting the enzyme, GDP-fucose synthase (FX). Reproduced with permission from [162].

1.9 Aims

The aims of this thesis are broad, encompassing the synthesis and characterisation of HMOs as well as biological applications of HMOs, including the development of simple, specific glycan binding assays.

As explained above, there are many challenges that impede the chemical and/or enzymatic synthesis of HMOs and their subsequent characterisation and purification, therefore, to this end this thesis aims to improve on current methods for the characterisation and purification of HMOs, principally LC methods such as HPAEC-PAD

and HILIC HPLC used in tandem with mass spectrometry. The use of HmFucT, a previously unreported enzyme, will be thoroughly explored for the enzymatic synthesis of fucosylated HMO analogues using a range of simple, complex and functionalised acceptor substrates. Functionalised lactosides with varying length linkers and functional groups will be examined as acceptor substrates to enable synthesis of complex functionalised HMOs that can be used for wider purposes, including immobilised carbohydrate assay development or synthesis of glycan-functionalised biosensors, probes, and nanomaterials. In addition, this enzyme will be investigated in one pot synthesis with FKP to catalyse the synthesis of the GDP-L-fucose sugar nucleotide donor for HmFucT. Novel non-natural HMO analogues incorporating, for example, D-Ara or L-Gal in place of L-Fuc, are thought to elicit variable responses in biological activity and could, therefore, be used to examine the molecular nature of HMO interactions with bacteria, including pathogens of interest. To enable access to rare GDP sugar nucleotide analogues and, subsequently, novel non-natural HMO analogues; FKP will also be assessed in terms of its substrate specificity and exploitation for this purpose.

Once structurally characterised HMOs are obtained, they will be used to establish the capabilities of wild-type non-pathogenic *Escherichia coli* as a model organism to study HMO utilisation of fucosylated HMOs for commensals and pathogens of interest. The interactions of glycans with microbes will also be examined following the development of a basic microplate-based glycan assay capable of determining specific binding interactions of glycans with live bacteria. *E. coli* strains will also be used as model organism to test said assay following validation with plant lectins.

In summary, the aim of this research is to:

1. Develop improved liquid chromatography strategies for the characterisation and purification of HMOs.
2. Investigate the synthesis of fucosylated HMOs and novel, non-natural HMO analogues.
3. Assess the ability of HMOs to modulate bacterial growth.
4. Develop a selective and sensitive microplate-based assay to detect glycan-bacterial interactions in live bacteria.

1.10 References

1. Becker, D. J. & Lowe, J. B. (2003) Fucose: biosynthesis and biological function in mammals, *Glycobiology.* **13**, 41r-53r.
2. Blackburn, D. G. (1993) Lactation: historical patterns and potential for manipulation, *J Dairy Sci.* **76**, 3195-212.
3. Kunz, C., Rudloff, S., Baier, W., Klein, N. & Strobel, S. (2000) Oligosaccharides in human milk: structural, functional, and metabolic aspects, *Annu Rev Nutr.* **20**, 699-722.
4. Andreas, N. J., Kampmann, B. & Le-Doare, K. M. (2015) Human breast milk: A review on its composition and bioactivity, *Early Human Development.* **91**, 629-635.
5. Akkerman, R., Faas, M. M. & de Vos, P. (2018) Non-digestible carbohydrates in infant formula as substitution for human milk oligosaccharide functions: Effects on microbiota and gut maturation, *Crit Rev Food Sci Nutr*, 1-12.
6. Sprenger, N., Lee, L. Y., De Castro, C. A., Steenhout, P. & Thakkar, S. K. (2017) Longitudinal change of selected human milk oligosaccharides and association to infants' growth, an observatory, single center, longitudinal cohort study, *PloS one.* **12**, e0171814.
7. Ballard, O. & Morrow, A. L. (2013) Human milk composition: nutrients and bioactive factors, *Pediatr Clin North Am.* **60**, 49-74.
8. Austin, S., De Castro, C. A., Benet, T., Hou, Y., Sun, H., Thakkar, S. K., Vinyes-Pares, G., Zhang, Y. & Wang, P. (2016) Temporal change of the content of 10 oligosaccharides in the milk of chinese urban mothers, *Nutrients.* **8**.
9. Meyer, K. M., Mohammad, M., Bode, L., Chu, D. M., Ma, J., Haymond, M. & Aagaard, K. (2017) Maternal diet structures the breast milk microbiome in association with human milk oligosaccharides and gut-associated bacteria, *American Journal of Obstetrics and Gynecology.* **216**, S15-S15.
10. Bode, L. (2006) Recent advances on structure, metabolism, and function of human milk oligosaccharides, *J Nutr.* **136**, 2127-2130.
11. Bode, L. (2012) Human milk oligosaccharides: Every baby needs a sugar mama, *Glycobiology.* **22**, 1147-1162.
12. Smilowitz, J. T., Lebrilla, C. B., Mills, D. A., German, J. B. & Freeman, S. L. (2014) Breast milk oligosaccharides: structure-function relationships in the neonate, *Annu Rev Nutr.* **34**, 143-169.
13. Grabarics, M., Csernak, O., Balogh, R. & Beni, S. (2017) Analytical characterization of human milk oligosaccharides - potential applications in pharmaceutical analysis, *J Pharm Biomed Anal.* **146**, 168-178.
14. Sprenger, G. A., Baumgärtner, F. & Albermann, C. (2017) Production of human milk oligosaccharides by enzymatic and whole-cell microbial biotransformations, *J Biotechnol.* **258**, 79-91.
15. Tao, N., Ochonicky, K. L., German, J. B., Donovan, S. M. & Lebrilla, C. B. (2010) Structural determination and daily variations of porcine milk oligosaccharides, *J Agric Food Chem.* **58**, 4653-4659.
16. Tao, N. A., Wu, S. A., Kim, J., An, H. J., Hinde, K., Power, M. L., Gagneux, P., German, J. B. & Lebrilla, C. B. (2011) Evolutionary Glycomics: Characterization of Milk Oligosaccharides in Primates, *Journal of Proteome Research.* **10**, 1548-1557.
17. Kumar Singh, A., Ranjan, A., Srivastava, G. & Deepak, D. (2015) Structure elucidation of two novel yak milk oligosaccharides and their DFT Studies, *J Mol Struct.* **7**, 9-13.
18. Gangwar, L., Singh, R. & Deepak, D. (2017) Structure elucidation of a novel oligosaccharide (Medalose) from camel milk, *J Mol Struct.* **1153**.
19. Verma, P., Gangwar, L., Singh, R., Sarkar, J. & Deepak, D. (2017) Original article isolation and structural determination of ramelose (a novel oligosaccharide from camel milk) by 2D-NMR, *International Journal of Carbohydrate Research.* **7**, 14-20.
20. Gangwar, L., Kumar, A. & Deepak, D. (2017) Original article isolation and structure elucidation of biologically active novel pentasaccharide from the milk of *Bubalus bubalis*, *International Journal of Carbohydrate Research.* **7**, 9-13.

21. Tao, N., Depeters, E. J., Freeman, S., German, J. B., Grimm, R. & Lebrilla, C. B. (2008) Bovine milk glycome, *J Dairy Sci.* **91**, 3768-3778.

22. Twigger, A. J., Hepworth, A. R., Lai, C. T., Chetwynd, E., Stuebe, A. M., Blancafort, P., Hartmann, P. E., Geddes, D. T. & Kakulas, F. (2015) Gene expression in breastmilk cells is associated with maternal and infant characteristics, *Sci Rep.* **5**, 12933.

23. Thurl, S., Henker, J., Siegel, M., Tovar, K. & Sawatzki, G. (1997) Detection of four human milk groups with respect to Lewis blood group dependent oligosaccharides, *Glycoconjugate Journal.* **14**, 795-799.

24. Newburg, D., Ruiz-Palacios, G. & Morrow, A. (2005) Human milk glycans protect infants against enteric pathogens, *Annu Rev Nutr.* **25**, 37-58.

25. Tissier, H. (1900) *Recherches sur la flore intestinale des nourrissons : état normal et pathologique*, G. Carré et C. Naud, Paris.

26. Ward, R. E., Niñonuevo, M., Mills, D. A., Lebrilla, C. B. & German, J. B. (2006) In vitro fermentation of breast milk oligosaccharides by *Bifidobacterium infantis* and *Lactobacillus gasseri*, *Appl Environ Microbiol.* **72**, 4497-4499.

27. Chambers, S. A., Moore, R. E., Craft, K. M., Thomas, H. C., Das, R., Manning, S. D., Codreanu, S. G., Sherrod, S. D., Aronoff, D. M., McLean, J. A., Gaddy, J. A. & Townsend, S. D. (2020) A solution to antifolate resistance in Group B *Streptococcus*: Untargeted metabolomics identifies human milk oligosaccharide-induced perturbations that result in potentiation of trimethoprim, *mBio.* **11**, e00076-20.

28. Ackerman, D., Doster, R., Weitkamp, J.-H., Aronoff, D., Gaddy, J. & Townsend, S. (2017) Human milk oligosaccharides exhibit antimicrobial and anti-biofilm properties against Group B Streptococcus, *ACS Infectious Diseases.* **3**.

29. Ray, C., Kerketta, J. A., Rao, S., Patel, S., Dutt, S., Arora, K., Pournami, F. & Bhushan, P. (2019) Human milk oligosaccharides: The journey ahead, *International Journal of Pediatrics.* **2019**.

30. Sela, D. A. & Mills, D. A. (2010) Nursing our microbiota: molecular linkages between bifidobacteria and milk oligosaccharides, *Trends Microbiol.* **18**, 298-307.

31. Bezirtzoglou, E., Tsiotsias, A. & Welling, G. W. (2011) Microbiota profile in feces of breast- and formula-fed newborns by using fluorescence in situ hybridization (FISH), *Anaerobe.* **17**, 478-82.

32. Gibson, G. R. & Roberfroid, M. B. (1995) Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics, *J Nutr.* **125**, 1401-12.

33. Coppa, G. V., Pierani, P., Zampini, L., Bruni, S., Carloni, I. & Gabrielli, O. (2001) Characterization of oligosaccharides in milk and feces of breast-fed infants by high-performance anion-exchange chromatography in *Bioactive Components of Human Milk* (Newburg, D. S., ed) pp. 307-314, Springer US, Boston, MA.

34. Engfer, M. B., Stahl, B., Finke, B., Sawatzki, G. & Daniel, H. (2000) Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract, *Am J Clin Nutr.* **71**, 1589-96.

35. Collado, M. C., Cernada, M., Baüerl, C., Vento, M. & Pérez-Martínez, G. (2012) Microbial ecology and host-microbiota interactions during early life stages, *Gut Microbes.* **3**, 352-65.

36. Gyorgy, P., Norris, R. F. & Rose, C. S. (1954) Bifidus factor. I. A variant of Lactobacillus bifidus requiring a special growth factor, *Arch Biochem Biophys.* **48**, 193-201.

37. Asakuma, S., Hatakeyama, E., Urashima, T., Yoshida, E., Katayama, T., Yamamoto, K., Kumagai, H., Ashida, H., Hirose, J. & Kitaoka, M. (2011) Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria, *J Biol Chem.* **286**, 34583-92.

38. Underwood, M. A., Davis, J. C. C., Kalanetra, K. M., Gehlot, S., Patole, S., Tancredi, D. J., Mills, D. A., Lebrilla, C. B. & Simmer, K. (2017) Digestion of human milk oligosaccharides by *Bifidobacterium breve* in the premature infant, *Journal of pediatric gastroenterology and nutrition.* **65**, 449-455.

39. Karav, S., Le Parc, A., Leite Nobrega de Moura Bell, J. M., Frese, S. A., Kirmiz, N., Block, D. E., Barile, D. & Mills, D. A. (2016) Oligosaccharides released from milk glycoproteins are selective growth substrates for infant-associated Bifidobacteria, *Appl Environ Microbiol.* **82**, 3622-3630.

40. Masi, A. C. & Stewart, C. J. (2022) Untangling human milk oligosaccharides and infant gut microbiome, *iScience.* **25**, 103542.

41. James, K., Motherway, M. O., Bottacini, F. & van Sinderen, D. (2016) *Bifidobacterium breve* UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways, *Sci Rep.* **6**, 38560.

42. Miwa, M., Horimoto, T., Kiyohara, M., Katayama, T., Kitaoka, M., Ashida, H. & Yamamoto, K. (2010) Cooperation of β-galactosidase and β-N-acetylhexosaminidase from bifidobacteria in assimilation of human milk oligosaccharides with type 2 structure, *Glycobiology.* **20**, 1402-9.

43. Yoshida, E., Sakurama, H., Kiyohara, M., Nakajima, M., Kitaoka, M., Ashida, H., Hirose, J., Katayama, T., Yamamoto, K. & Kumagai, H. (2012) *Bifidobacterium longum* subsp. *infantis* uses two different β-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides, *Glycobiology.* **22**, 361-8.

44. Sela, D. A., Chapman, J., Adeuya, A., Kim, J. H., Chen, F., Whitehead, T. R., Lapidus, A., Rokhsar, D. S., Lebrilla, C. B., German, J. B., Price, N. P., Richardson, P. M. & Mills, D. A. (2008) The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome, *Proceedings of the National Academy of Sciences of the United States of America.* **105**, 18964-18969.

45. Stewart, C. J., Ajami, N. J., O'Brien, J. L., Hutchinson, D. S., Smith, D. P., Wong, M. C., Ross, M. C., Lloyd, R. E., Doddapaneni, H., Metcalf, G. A., Muzny, D., Gibbs, R. A., Vatanen, T., Huttenhower, C., Xavier, R. J., Rewers, M., Hagopian, W., Toppari, J., Ziegler, A.-G., She, J.-X., Akolkar, B., Lernmark, A., Hyoty, H., Vehik, K., Krischer, J. P. & Petrosino, J. F. (2018) Temporal development of the gut microbiome in early childhood from the TEDDY study, *Nature.* **562**, 583-588.

46. Shao, Y., Forster, S. C., Tsaliki, E., Vervier, K., Strang, A., Simpson, N., Kumar, N., Stares, M. D., Rodger, A., Brocklehurst, P., Field, N. & Lawley, T. D. (2019) Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth, *Nature.* **574**, 117-121.

47. Yu, Z. T., Chen, C. & Newburg, D. S. (2013) Utilization of major fucosylated and sialylated human milk oligosaccharides by isolated human gut microbes, *Glycobiology.* **23**, 1281-1292.

48. Marcobal, A., Barboza, M., Froehlich, J. W., Block, D. E., German, J. B., Lebrilla, C. B. & Mills, D. A. (2010) Consumption of human milk oligosaccharides by gut-related microbes, *J Agric Food Chem.* **58**, 5334-5340.

49. Thongaram, T., Hoeflinger, J. L., Chow, J. & Miller, M. J. (2017) Human milk oligosaccharide consumption by probiotic and human-associated bifidobacteria and lactobacilli, *J Dairy Sci.* **100**, 7825-7833.

50. Schwab, C. & Gänzle, M. (2011) Lactic acid bacteria fermentation of human milk oligosaccharide components, human milk oligosaccharides and galactooligosaccharides, *FEMS Microbiol Lett.* **315**, 141-148.

51. Salli, K., Hirvonen, J., Siitonen, J., Ahonen, I., Anglenius, H. & Maukonen, J. (2021) Selective utilization of the human milk oligosaccharides 2'-fucosyllactose, 3-fucosyllactose, and difucosyllactose by various probiotic and pathogenic bacteria, *J Agric Food Chem.* **69**, 170-182.

52. Hoeflinger, J. L., Davis, S. R., Chow, J. & Miller, M. J. (2015) In vitro impact of human milk oligosaccharides on Enterobacteriaceae growth, *J Agric Food Chem.* **63**, 3295-302.

53. Zuurveld, M., van Witzenburg, N. P., Garssen, J., Folkerts, G., Stahl, B., Van't Land, B. & Willemsen, L. E. M. (2020) Immunomodulation by human milk oligosaccharides: The potential role in prevention of allergic diseases, *Front Immunol.* **11**, 801.

54. Alessandri, G., Ossiprandi, M. C., MacSharry, J., van Sinderen, D. & Ventura, M. (2019) Bifidobacterial dialogue with its human host and consequent modulation of the immune system, *Front Immunol.* **10**, 2348.

55. Parkkinen, J., Finne, J., Achtman, M., Vaisanen, V. & Korhonen, T. K. (1983) *Escherichia coli* strains binding neuraminyl alpha 2-3 galactosides, *Biochem Biophys Res Commun.* **111**, 456-61.

56. Firon, N., Ofek, I. & Sharon, N. (1983) Carbohydrate specificity of the surface lectins of *Escherichia coli, Klebsiella pneumoniae*, and *Salmonella typhimurium*, *Carbohydr Res.* **120**, 235-49.

57. Ruiz-Palacios, G. M., Cervantes, L. E., Ramos, P., Chavez-Munguia, B. & Newburg, D. S. (2003) *Campylobacter jejuni* binds intestinal H(O) antigen (Fuc alpha 1, 2Gal beta 1, 4GlcNAc), and fucosyloligosaccharides of human milk inhibit its binding and infection, *J Biol Chem.* **278**, 14112-14120.

58. Morrow, A. L., Ruiz-Palacios, G. M., Altaye, M., Jiang, X., Guerrero, M. L., Meinzen-Derr, J. K., Farkas, T., Chaturvedi, P., Pickering, L. K. & Newburg, D. S. (2004) Human milk oligosaccharides are associated with protection against diarrhea in breast-fed infants, *J Pediatr.* **145**, 297-303.

59. Morrow, A. L., Ruiz-Palacios, G. M., Altaye, M., Jiang, X., Guerrero, M. L., Meinzen-Derr, J. K., Farkas, T., Chaturvedi, P., Pickering, L. K. & Newburg, D. S. (2004) Human milk oligosaccharide blood group epitopes and innate immune protection against *Campylobacter* and calicivirus diarrhea in breastfed infants, *Protecting Infants through Human Milk.* **554**, 443-446.

60. Downham, M. A., Scott, R., Sims, D. G., Webb, J. K. & Gardner, P. S. (1976) Breast-feeding protects against respiratory syncytial virus infections, *Br Med J.* **2**, 274-276.

61. Abrahams, S. W. & Labbok, M. H. (2011) Breastfeeding and otitis media: a review of recent evidence, *Curr Allergy Asthma Rep.* **11**, 508-12.

62. Martin-Sosa, S., Martin, M. J. & Hueso, P. (2002) The sialylated fraction of milk oligosaccharides is partially responsible for binding to enterotoxigenic and uropathogenic *Escherichia coli* human strains, *J Nutr.* **132**, 3067-72.

63. Raman, R., Tharakaraman, K., Sasisekharan, V. & Sasisekharan, R. (2016) Glycan-protein interactions in viral pathogenesis, *Curr Opin Struct Biol.* **40**, 153-162.

64. Nauwynck, H. J., Favoreel, H. W., Van Breedam, W., Pöhlmann, S. & de Groot, R. J. (2014) Bitter-sweet symphony: Glycan–lectin interactions in virus biology, *FEMS Microbiol Rev.* **38**, 598-632.

65. Etzold, S. & Bode, L. (2014) Glycan-dependent viral infection in infants and the role of human milk oligosaccharides, *Current opinion in virology.* **7**, 101-107.

66. Moore, R. E., Xu, L. L. & Townsend, S. D. (2021) Prospecting human milk oligosaccharides as a defense against viral infections, *ACS Infectious Diseases.* **7**, 254-263.

67. Gozalbo-Rovira, R., Ciges-Tomas, J. R., Vila-Vicent, S., Buesa, J., Santiso-Bellón, C., Monedero, V., Yebra, M. J., Marina, A. & Rodríguez-Díaz, J. (2019) Unraveling the role of the secretor antigen in human rotavirus attachment to histo-blood group antigens, *PLoS Path.* **15**, e1007865.

68. Weichert, S., Koromyslova, A., Singh, B. K., Hansman, S., Jennewein, S., Schroten, H. & Hansman, G. S. (2016) Structural basis for norovirus inhibition by human milk oligosaccharides, *J Virol.* **90**, 4843-4848.

69. Su, S. V., Gurney, K. B. & Lee, B. (2003) Sugar and spice: Viral envelope-DC-SIGN interactions in HIV pathogenesis, *Curr HIV Res.* **1**, 87-99.

70. van Kooyk, Y. & Geijtenbeek, T. B. (2003) DC-SIGN: Escape mechanism for pathogens, *Nat Rev Immunol.* **3**, 697-709.

71. Wu, L. & KewalRamani, V. N. (2006) Dendritic-cell interactions with HIV: infection and viral dissemination, *Nat Rev Immunol.* **6**, 859-68.

72. Doores, K. J., Bonomelli, C., Harvey, D. J., Vasiljevic, S., Dwek, R. A., Burton, D. R., Crispin, M. & Scanlan, C. N. (2010) Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens, *Proc Natl Acad Sci U S A.* **107**, 13800-5.

73. Hong, P., Ninonuevo, M. R., Lee, B., Lebrilla, C. & Bode, L. (2009) Human milk oligosaccharides reduce HIV-1-gp120 binding to dendritic cell-specific ICAM3-grabbing non-integrin (DC-SIGN), *British Journal of Nutrition.* **101**, 482-486.

74. Chutipongtanate, S., Morrow, A. L. & Newburg, D. S. (2022) Human milk oligosaccharides: Potential applications in COVID-19, *Biomedicines.* **10**, 346.

75. Morozov, V., Hansman, G., Hanisch, F. G., Schroten, H. & Kunz, C. (2018) Human milk oligosaccharides as promising antivirals, *Mol Nutr Food Res.* **62**.

76. Nizet, V., Varki, A., Aebi, M. (2017) Essentials of Glycobiology 3rd edition, Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

77. Backert, S., Boehm, M., Wessler, S. & Tegtmeyer, N. (2013) Transmigration route of Campylobacter jejuni across polarized intestinal epithelial cells: paracellular, transcellular or both?, *Cell communication and signaling : CCS.* **11**, 72-72.

78. Camilleri, M., Madsen, K., Spiller, R., Greenwood-Van Meerveld, B. & Verne, G. N. (2012) Intestinal barrier function in health and gastrointestinal disease, *Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society.* **24**, 503-12.

79. Kuntz, S., Rudloff, S. & Kunz, C. (2008) Oligosaccharides from human milk influence growth-related characteristics of intestinally transformed and non-transformed intestinal cells, *Br J Nutr.* **99**, 462-71.

80. Šuligoj, T., Vigsnæs, L. K., Abbeele, P. V. d., Apostolou, A., Karalis, K., Savva, G. M., McConnell, B. & Juge, N. (2020) Effects of human milk oligosaccharides on the adult gut microbiota and barrier function, *Nutrients.* **12**, 2808.

81. Angeloni, S., Ridet, J. L., Kusy, N., Gao, H., Crevoisier, F., Guinchard, S., Kochhar, S., Sigrist, H. & Sprenger, N. (2005) Glycoprofiling with micro-arrays of glycoconjugates and lectins, *Glycobiology.* **15**, 31-41.

82. Albrecht, S., Schols, H. A., van Zoeren, D., van Lingen, R. A., Groot Jebbink, L. J. M., van den Heuvel, E. G. H. M., Voragen, A. G. J. & Gruppen, H. (2011) Oligosaccharides in feces of breast- and formula-fed babies, *Carbohydrate Research.* **346**, 2173-2181.

83. Triantis, V., Bode, L. & van Neerven, R. J. J. (2018) Immunological effects of human milk oligosaccharides, *Frontiers in Pediatrics.* **6**, 14.

84. Dotz, V., Rudloff, S., Meyer, C., Lochnit, G. & Kunz, C. (2015) Metabolic fate of neutral human milk oligosaccharides in exclusively breast-fed infants, *Mol Nutr Food Res.* **59**, 355-364.

85. Eiwegger, T., Stahl, B., Haidl, P., Schmitt, J., Boehm, G., Dehlink, E., Urbanek, R. & Szepfalusi, Z. (2010) Prebiotic oligosaccharides: *in vitro* evidence for gastrointestinal epithelial transfer and immunomodulatory properties, *Pediatr Allergy Immunol.* **21**, 1179-88.

86. Kim, J., Kim, Y. J. & Kim, J. W. (2018) Bacterial clearance enhanced by α2, 3- and α2, 6-sialyllactose via receptor-mediated endocytosis and phagocytosis, *Infect Immun.* **87**.

87. Springer, T. A. (1994) Traffic signals for lymphocyte recirculation and leukocyte emigration: the multistep paradigm, *Cell.* **76**, 301-14.

88. Bode, L., Kunz, C., Muhly-Reinholz, M., Mayer, K., Seeger, W. & Rudloff, S. (2004) Inhibition of monocyte, lymphocyte, and neutrophil adhesion to endothelial cells by human milk oligosaccharides, *Thromb Haemostasis.* **92**, 1402-1410.

89. Prudden, A. R., Barile, D., Boons, G.-J., Bode, L., Pohl, N., Contractor, N., Jennewein, S. & Jin, Y.-S. (2016) Overcoming the limited availability of human milk oligosaccharides: challenges and opportunities for research and application, *Nutr Rev.* **74**, 635-644.

90. Jantscher-Krenn, E., Zherebtsov, M., Nissan, C., Goth, K., Guner, Y. S., Naidu, N., Choudhury, B., Grishin, A. V., Ford, H. R. & Bode, L. (2012) The human milk oligosaccharide disialyllacto-N-tetraose prevents necrotising enterocolitis in neonatal rats, *Gut.* **61**, 1417-1425.

91. Urashima, T., Katayama, T., Sakanaka, M., Fukuda, K. & Messer, M. (2021) Evolution of milk oligosaccharides: Origin and selectivity of the ratio of milk oligosaccharides to lactose among mammals, *Biochimica et biophysica acta General subjects.* **1866**, 130012.

92. Urashima, T., Saito, T., Nakamura, T. & Messer, M. (2001) Oligosaccharides of milk and colostrum in non-human mammals, *Glycoconjugate Journal.* **18**, 357-371.

93. Sakai, T., Hirata, S., Fuwa, K., Sugama, K., Kusunoki, K., Makishima, H., Eguchi, T., Yamada, S., Ogihara, N. & Takeshita, H. (2012) Fetal brain development in chimpanzees versus humans, *Curr Biol.* **22**, R791-R792.

94. Castanys-Muñoz, E., Martin, M. J. & Prieto, P. A. (2013) 2′-Fucosyllactose: an abundant, genetically determined soluble glycan present in human milk, *Nutr Rev.* **71**, 773-789.

95. Kuhn, N. J., Carrick, D. T. & Wilde, C. J. (1980) Lactose synthesis: The possibilities of regulation, *J Dairy Sci.* **63**, 328-336.

96. Prieto, P., Mukerji, P., Kelder, B., Erney, R., Gonzalez, D., S Yun, J., Smith, D., Moremen, K., Nardelli, C. & Pierce, M. (1996) Remodeling of mouse milk glycoconjugates by transgenic expression of a human glycosyltransferase, *J Biol Chem.* **270**, 29515-29519.

97. O'Sullivan, A., Salcedo, J. & Rubert, J. (2018) Advanced analytical strategies for measuring free bioactive milk sugars: from composition and concentrations to human metabolic response, *Anal Bioanal Chem.* **410**, 3445-3462.

98. Porfirio, S., Archer-Hartmann, S., Moreau, G. B., Ramakrishnan, G., Haque, R., Kirkpatrick, B. D., Petri, W. A., Jr & Azadi, P. (2020) New strategies for profiling and characterization of human milk oligosaccharides, *Glycobiology*.

99. Guo, J. & Ye, X. S. (2010) Protecting groups in carbohydrate chemistry: influence on stereoselectivity of glycosylations, *Molecules.* **15**, 7235-65.

100. Abragam, J. A., Pardo-Vargas, A. & Seeberger, P. H. (2020) Total Synthesis of Polysaccharides by Automated Glycan Assembly, *J Am Chem Soc.* **142**, 8561-8564.

101. Benkoulouche, M., Fauré, R., Remaud-Siméon, M., Moulis, C. & André, I. (2019) Harnessing glycoenzyme engineering for synthesis of bioactive oligosaccharides, *Interface Focus.* **9**.

102. Jung, S. M., Park, Y. C. & Seo, J. H. (2019) Production of 3-fucosyllactose in engineered *Escherichia coli* with alpha-1,3-fucosyltransferase from *Helicobacter pylori*, *Biotechnology Journal.* **14**.

103. Baumgartner, F., Conrad, J., Sprenger, G. A. & Albermann, C. (2014) Synthesis of the human milk oligosaccharide lacto-*N*-tetraose in metabolically engineered, plasmid-free *E. coli*, *ChemBioChem.* **15**, 1896-1900.

104. Baumgärtner, F., Seitz, L., Sprenger, G. A. & Albermann, C. (2013) Construction of *Escherichia coli* strains with chromosomally integrated expression cassettes for the synthesis of 2′-fucosyllactose, *Microbial Cell Factories.* **12**, 40.

105. Yu, S., Liu, J. J., Yun, E. J., Kwak, S., Kim, K. H. & Jin, Y. S. (2018) Production of a human milk oligosaccharide 2'-fucosyllactose by metabolically engineered *Saccharomyces cerevisiae*, *Microbial Cell Factories.* **17**.

106. Zeuner, B., Teze, D., Muschiol, J. & Meyer, A. S. (2019) Synthesis of human milk oligosaccharides: Protein engineering strategies for improved enzymatic transglycosylation, *Molecules.* **24**.

107. Zeuner, B., Vuillemin, M., Holck, J., Muschiol, J. & Meyer, A. S. (2018) Loop engineering of an alpha-1,3/4-L-fucosidase for improved synthesis of human milk oligosaccharides, *Enzyme Microb Technol.* **115**, 37-44.

108. Deng, J., Lv, X., Li, J., Du, G., Chen, J. & Liu, L. (2021) Recent advances and challenges in microbial production of human milk oligosaccharides, *Systems Microbiology and Biomanufacturing.* **1**, 1-14.

109. EFSA Panel on Dietetic Products, N. a. A. N. (2015) Safety of 2'-O-fucosyllactose as a novel food ingredient pursuant to Regulation (EC) No 258/97 in, European Food Standards Agency (EFSA),

110. Vandenplas, Y., Berger, B., Carnielli, V. P., Ksiazyk, J., Lagstrom, H., Sanchez Luna, M., Migacheva, N., Mosselmans, J. M., Picaud, J. C., Possner, M., Singhal, A. & Wabitsch, M. (2018) Human milk oligosaccharides: 2'-fucosyllactose (2'-FL) and lacto-N-neotetraose (LNnT) in infant formula, *Nutrients.* **10**.

111. Meng, J., Zhu, Y., Wang, H., Cao, H. & Mu, W. (2023) Biosynthesis of human milk oligosaccharides: Enzyme cascade and metabolic engineering approaches, *J Agric Food Chem.* **71**, 2234-2243.

112. Prudden, A. R., Liu, L., Capicciotti, C. J., Wolfert, M. A., Wang, S., Gao, Z. W., Meng, L., Moremen, K. W. & Boons, G. J. (2017) Synthesis of asymmetrical multiantennary human milk oligosaccharides, *Proceedings of the National Academy of Sciences of the United States of America.* **114**, 6954-6959.

113. Zhao, C., Wu, Y. J., Yu, H., Shah, I. M., Li, Y. H., Zeng, J., Liu, B., Mills, D. A. & Chen, X. (2016) The one-pot multienzyme (OPME) synthesis of human blood group H antigens and a human milk oligosaccharide (HMOS) with highly active *Thermosynechococcus elongatus* alpha 1-2-fucosyltransferase, *Chemical Communications.* **52**, 3899-3902.

114. Xiao, Z., Guo, Y., Liu, Y., Li, L., Zhang, Q., Wen, L., Wang, X., Kondengaden, S. M., Wu, Z., Zhou, J., Cao, X., Li, X., Ma, C. & Wang, P. G. (2016) Chemoenzymatic synthesis of a library of human milk oligosaccharides, *The Journal of Organic Chemistry.* **81**, 5851-5865.

115. Ye, J., Xia, H., Sun, N., Liu, C.-C., Sheng, A., Chi, L., Liu, X.-W., Gu, G., Wang, S.-Q., Zhao, J., Wang, P., Xiao, M., Wang, F. & Cao, H. (2019) Reprogramming the enzymatic assembly line for site-specific fucosylation, *Nature Catalysis*.

116. Javaud, C., Dupuy, F., Maftah, A., Julien, R. & Petit, J.-M. (2003) The fucosyltransferase gene family: An amazing summary of the underlying mechanisms of gene evolution, *Genetica.* **118**, 157-170.

117. Dumon, C., Priem, B., Martin, S. L., Heyraud, A., Bosso, C. & Samain, E. (2001) *In vivo* fucosylation of lacto-N-neotetraose and lacto-N-neohexaose by heterologous expression of *Helicobacter pylori* alpha-1,3 fucosyltransferase in engineered *Escherichia coli*, *Glycoconjugate Journal.* **18**, 465-474.

118. Martin, S. L., Edbrooke, M. R., Hodgman, T. C., vandenEijnden, D. H. & Bird, M. I. (1997) Lewis X biosynthesis in *Helicobacter pylori* - Molecular cloning of an alpha(1,3)-fucosyltransferase gene, *J Biol Chem.* **272**, 21349-21356.

119. Tsai, T. W., Fang, J. L., Liang, C. Y., Wang, C. J., Huang, Y. T., Wang, Y. J., Li, J. Y. & Yu, C. C. (2019) Exploring the synthetic application of *Helicobacter pylori* alpha 1,3/4-fucosyltransferase FucTIII toward the syntheses of fucosylated human milk glycans and Lewis antigens, *Acs Catalysis.* **9**, 10712-10720.

120. Rasko, D. A., Wang, G., Monteiro, M. A., Palcic, M. M. & Taylor, D. E. (2000) Synthesis of mono- and di-fucosylated type I Lewis blood group antigen by *Helicobacter pylori*, *Eur J Biochem.* **267**, 6059-6066.

121. Bai, J., Wu, Z., Sugiarto, G., Gadi, M. R., Yu, H., Li, Y., Xiao, C., Ngo, A., Zhao, B., Chen, X. & Guan, W. (2019) Biochemical characterization of *Helicobacter pylori* alpha1-3-fucosyltransferase and its application in the synthesis of fucosylated human milk oligosaccharides, *Carbohydrate Research.* **480**, 1-6.

122. Abrantes, J., Posada, D., Guillon, P., Esteves, P. J. & Le Pendu, J. (2009) Widespread gene conversion of alpha-2-fucosyltransferase genes in mammals, *J Mol Evol.* **69**, 22-31.

123. de Vries, T., Knegtel, R. M., Holmes, E. H. & Macher, B. A. (2001) Fucosyltransferases: structure/function studies, *Glycobiology.* **11**, 119r-128r.

124. Blanas, A., Sahasrabudhe, N. M., Rodríguez, E., van Kooyk, Y. & van Vliet, S. J. (2018) Fucosylated antigens in cancer: An alliance toward tumor progression, metastasis, and resistance to chemotherapy, *Frontiers in oncology.* **8**, 39.

125. Tu, C.-F., Wu, M.-Y., Lin, Y.-C., Kannagi, R. & Yang, R.-B. (2017) FUT8 promotes breast cancer cell invasiveness by remodeling TGF-β receptor core fucosylation, *Breast Cancer Research.* **19**, 111.

126. Yang, G., Wang, Q., Chen, L., Betenbaugh, M. J. & Zhang, H. (2021) Glycoproteomic characterization of FUT8 knock-out CHO cells reveals roles of FUT8 in the glycosylation, *Frontiers in Chemistry.* **9**.

127. Mollicone, R., Moore, S. E., Bovin, N., Garcia-Rosasco, M., Candelier, J. J., Martinez-Duncker, I. & Oriol, R. (2009) Activity, splice variants, conserved peptide motifs, and phylogeny of two new alpha1,3-fucosyltransferase families (FUT10 and FUT11), *J Biol Chem.* **284**, 4723-38.

128. Chen, C. I., Keusch, J. J., Klein, D., Hess, D., Hofsteenge, J. & Gut, H. (2012) Structure of human POFUT2: insights into thrombospondin type 1 repeat fold and O-fucosylation, *EMBO J.* **31**, 3183-97.

129. Stahl, M., Uemura, K., Ge, C., Shi, S., Tashima, Y. & Stanley, P. (2008) Roles of Pofut1 and O-fucose in mammalian Notch signaling, *J Biol Chem.* **283**, 13638-51.

130. Wang, Y., Shao, L., Shi, S., Harris, R. J., Spellman, M. W., Stanley, P. & Haltiwanger, R. S. (2001) Modification of epidermal growth factor-like repeats with O-fucose. Molecular cloning and expression of a novel GDP-fucose protein O-fucosyltransferase, *J Biol Chem.* **276**, 40338-45.

131. Ma, B., Simala-Grant, J. L. & Taylor, D. E. (2006) Fucosylation in prokaryotes and eukaryotes, *Glycobiology.* **16**, 158R-184R.

132. Stein, D. B., Lin, Y. N. & Lin, C. H. (2008) Characterization of *Helicobacter pylori* alpha 1,2-fucosyltransferase for enzymatic synthesis of tumor-associated antigens, *Advanced Synthesis & Catalysis.* **350**, 2313-2321.

133. Engels, L. & Elling, L. (2013) WbgL: a novel bacterial α1,2-fucosyltransferase for the synthesis of 2′-fucosyllactose, *Glycobiology.* **24**, 170-178.

134. Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T. & Ben-Tal, N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules, *Nucleic Acids Res.* **44**, W344-W350.

135. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids, *Nucleic Acids Res.* **38**, W529-W533.

136. Celniker, G., Nimrod, G., Ashkenazy, H., Glaser, F., Martz, E., Mayrose, I., Pupko, T. & Ben-Tal, N. (2013) ConSurf: Using evolutionary data to raise testable hypotheses about protein function, *Israel Journal of Chemistry.* **53**, 199-206.

137. Glaser, F., Pupko, T., Paz, I., Bell, R. E., Bechor-Shental, D., Martz, E. & Ben-Tal, N. (2003) ConSurf: Identification of functional regions in proteins by surface-mapping of phylogenetic information, *Bioinformatics.* **19**, 163-164.

138. Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T. & Ben-Tal, N. (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures, *Nucleic Acids Res.* **33**, W299-W302.

139. Yariv, B., Yariv, E., Kessel, A., Masrati, G., Chorin, A. B., Martz, E., Mayrose, I., Pupko, T. & Ben-Tal, N. (2023) Using evolutionary data to make sense of macromolecules with a "face-lifted" ConSurf, *Protein Sci.* **32**, e4582.

140. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. (2021) Highly accurate protein structure prediction with AlphaFold, *Nature.* **596**, 583-589.

141. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P., Kleywegt, G., Birney, E., Hassabis, D. & Velankar, S. (2021) AlphaFold protein structure aatabase: massively expanding the structural coverage of protein-sequence space with high-accuracy models, *Nucleic Acids Res.* **50**, D439-D444.

142. Coyne, M. J., Reinap, B., Lee, M. M. & Comstock, L. E. (2005) Human Symbionts Use a Host-Like Pathway for Surface Fucosylation, *Science.* **307**, 1778-1781.

143. Liu, Y., Hu, H., Wang, J., Zhou, Q., Wu, P., Yan, N., Wang, H.-W., Wu, J.-W. & Sun, L. (2019) Cryo-EM structure of L-fucokinase/GDP-fucose pyrophosphorylase (FKP) in Bacteroides fragilis, *Protein Cell.* **10**, 365-369.

144. Rabuka, D., Hubbard, S. C., Laughlin, S. T., Argade, S. P. & Bertozzi, C. R. (2006) A chemical reporter strategy to probe glycoprotein fucosylation, *J Am Chem Soc.* **128**, 12078-9.

145. Sawa, M., Hsu, T. L., Itoh, T., Sugiyama, M., Hanson, S. R., Vogt, P. K. & Wong, C. H. (2006) Glycoproteomic probes for fluorescent imaging of fucosylated glycans in vivo, *Proc Natl Acad Sci U S A.* **103**, 12371-6.

146. Wang, W., Hu, T. S., Frantom, P. A., Zheng, T. Q., Gerwe, B., del Amo, D. S., Garret, S., Seidel, R. D. & Wu, P. (2009) Chemoenzymatic synthesis of GDP-L-fucose and the Lewis X glycan derivatives, *Proceedings of the National Academy of Sciences of the United States of America.* **106**, 16096-16101.

147. Schneider, M., Al-Shareffi, E. & Haltiwanger, R. S. (2017) Biological functions of fucose in mammals, *Glycobiology.* **27**, 601-618.

148. Lowe, J. B. (1993) The blood group-specific human glycosyltransferases, *Bailliere's clinical haematology.* **6**, 465-92.

149. Pickard, J. M. & Chervonsky, A. V. (2015) Intestinal fucose as a mediator of host-microbe symbiosis, *Journal of immunology (Baltimore, Md : 1950).* **194**, 5588-5593.

150. Wands, A. M., Cervin, J., Huang, H., Zhang, Y., Youn, G., Brautigam, C. A., Matson Dzebo, M., Bjorklund, P., Wallenius, V., Bright, D. K., Bennett, C. S., Wittung-Stafshede, P., Sampson, N. S., Yrlid, U. & Kohler, J. J. (2018) Fucosylated Molecules Competitively Interfere with Cholera Toxin Binding to Host Cells, *ACS Infect Dis.* **4**, 758-770.

151. Barbé, L., Le Moullac-Vaidye, B., Echasserieau, K., Bernardeau, K., Carton, T., Bovin, N., Nordgren, J., Svensson, L., Ruvoën-Clouet, N. & Le Pendu, J. (2018) Histo-blood group antigen-binding specificities of human rotaviruses are associated with gastroenteritis but not with *in vitro* infection, *Scientific Reports.* **8**, 12961.

152. Roy, R., Ghosh, B. & Kar, P. (2020) Investigating conformational dynamics of Lewis Y oligosaccharides and elucidating blood group dependency of Cholera using molecular dynamics, *ACS Omega.* **5**, 3932-3942.

153. Hu, L., Sankaran, B., Laucirica, D. R., Patil, K., Salmen, W., Ferreon, A. C. M., Tsoi, P. S., Lasanajak, Y., Smith, D. F., Ramani, S., Atmar, R. L., Estes, M. K., Ferreon, J. C. & Prasad, B. V. V. (2018) Glycan recognition in globally dominant human rotaviruses, *Nature communications.* **9**, 2631.

154. Kashyap, P. C., Marcobal, A., Ursell, L. K., Smits, S. A., Sonnenburg, E. D., Costello, E. K., Higginbottom, S. K., Domino, S. E., Holmes, S. P., Relman, D. A., Knight, R., Gordon, J. I. & Sonnenburg, J. L. (2013) Genetically dictated change in host mucus carbohydrate landscape exerts a diet-dependent effect on the gut microbiota, *Proc Natl Acad Sci U S A.* **110**, 17059-64.

155. Pickard, J. M., Maurice, C. F., Kinnebrew, M. A., Abt, M. C., Schenten, D., Golovkina, T. V., Bogatyrev, S. R., Ismagilov, R. F., Pamer, E. G., Turnbaugh, P. J. & Chervonsky, A. V. (2014) Rapid fucosylation of intestinal epithelium sustains host–commensal symbiosis in sickness, *Nature.* **514**, 638-641.

156. Pham, T. A., Clare, S., Goulding, D., Arasteh, J. M., Stares, M. D., Browne, H. P., Keane, J. A., Page, A. J., Kumasaka, N., Kane, L., Mottram, L., Harcourt, K., Hale, C., Arends, M. J., Gaffney, D. J., Dougan, G. & Lawley, T. D. (2014) Epithelial IL-22RA1-mediated fucosylation promotes intestinal colonization resistance to an opportunistic pathogen, *Cell host & microbe.* **16**, 504-16.

157. Pacheco, A. R., Curtis, M. M., Ritchie, J. M., Munera, D., Waldor, M. K., Moreira, C. G. & Sperandio, V. (2012) Fucose sensing regulates bacterial intestinal colonization, *Nature.* **492**, 113-117.

158. Garber, J. M., Hennet, T. & Szymanski, C. M. (2021) Significance of fucose in intestinal health and disease, *Mol Microbiol.* **115**, 1086-1093.

159.   Miyoshi, E., Moriwaki, K., Terao, N., Tan, C. C., Terao, M., Nakagawa, T., Matsumoto, H., Shinzaki, S. & Kamada, Y. (2012) Fucosylation is a promising target for cancer diagnosis and therapy, *Biomolecules.* **2**, 34-45.

160. Kannagi, R. (1997) Carbohydrate-mediated cell adhesion involved in hematogenous metastasis of cancer, *Glycoconj J.* **14**, 577-84.

161. Zhang, Z., Sun, P., Liu, J., Fu, L., Yan, J., Liu, Y., Yu, L., Wang, X. & Yan, Q. (2008) Suppression of FUT1/FUT4 expression by siRNA inhibits tumor growth, *Biochim Biophys Acta.* **1783**, 287-96.

162. Kizuka, Y., Nakano, M., Yamaguchi, Y., Nakajima, K., Oka, R., Sato, K., Ren, C. T., Hsu, T. L., Wong, C. H. & Taniguchi, N. (2017) An alkynyl-fucose halts hepatoma cell migration and invasion by inhibiting GDP-fucose-synthesizing enzyme FX, TSTA3, *Cell Chem Biol.* **24**, 1467-1478 e5.

163. Zhou, Y., Fukuda, T., Hang, Q., Hou, S., Isaji, T., Kameyama, A. & Gu, J. (2017) Inhibition of fucosylation by 2-fluorofucose suppresses human liver cancer HepG2 cell proliferation and migration as well as tumor formation, *Sci Rep.* **7**, 11563.

Chapter 2

Chapter 2 – Enzymatic Synthesis and Analysis of HMOs

2.1 Introduction

Fucosylated oligosaccharides expedite the establishment of a healthy microbiota and provide protection from infection. However, some pathogens have developed strategies to exploit these fucosylated oligosaccharides to adhere to the host cell surface and cause infection [158]. Fucosylated HMOs serve as an important prebiotic in breast-fed infants due to their ability to influence gut microbiota composition in early life. They also protect against pathogenic infections by acting as antiadhesive antimicrobials [11, 158, 164]. Due to their excellent physiological functions on breast-fed infants, fucosylated HMOs have gained ever-increasing interest. However, many studies examine the bioactive properties of pooled milk samples. This is due to a number of reasons, principally that many complex fucosylated HMO structures often exist in nature at low concentrations, which are notoriously difficult to isolate and purify from milk samples, and for which there are still limited synthetic routes. Therefore, this chapter aims to develop robust methods for the synthesis and analysis of fucosylated human milk oligosaccharides.

In the first instance, complementary chromatography and mass spectroscopy techniques will be explored and developed for the characterisation and purification of HMOs. Once these key analytical procedures are in place, we will investigate different chemical and enzymatic tools for the synthesis of fucosylated HMO analogues to probe glycan-bacterial interactions. Specifically, α1,2-fucosyltransferase from *Helicobacter mustelae* – a commercially available enzyme from Chemily Glycoscience for which there is no published data. The substrate specificity of this enzyme will be explored with a range of acceptor substrates, including various HMO core structures and functionalised lactose derivatives. In order to achieve scalable enzymatic synthesis of fucosylated HMO analogues, a one pot reaction with the bifunctional enzyme, L-fucokinase/GDP-L-fucose pyrophosphorylase (FKP) will be developed. The substrate specificity of FKP will be explored for the synthesis of sugar nucleotide donors with unique chemistries for incorporation in novel, non-natural HMO analogues.

2.1.2 HMO Characterisation and Purification

Milk is regarded as the best source of nourishment for infant mammals, providing all the nutrients and bioactives necessary for all aspects of infant development [165]. HMOs are one such class of milk bioactives with more than 200 HMO structures on record [166]. Due to their high structural diversity and complexity because of the high number of possible monosaccharide combinations and linkages, their identification and quantification pose a challenge to researchers. However, the correct identification of HMOs is essential as their specific structures lead to different biological effects [97, 98]. Identification and structural assignment have been accomplished by a number of analytical techniques; principally, liquid chromatography (LC) and capillary electrophoresis (CE) methods, which can be coupled to a mass spectrometer (MS).

LC utilises a number of stationary phases including reversed phase, anion-exchange, porous graphitised carbon and hydrophilic interaction modes (HILIC). However, due to the poor UV absorption of HMOs high performance liquid chromatography (HPLC) detection methods can be limited with conventional diode-array detections (DAD), unless derivatised. Fortunately, HPLC can be used with other label-free detection methods such as, refractive index (RI), evaporative light scattering (ELSD), charged aerosol (CAD) and pulsed amperometric detection (PAD) [167, 168]. Given its capability to resolve isomers, high pH anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) is a well-established technique, often used to separate HMOs [169]. Regrettably, the technique suffers from low sensitivity compared with MS modes [170] and the pH can cause epimerisation in some cases. HILIC is another frequently used analytical tool in the field of HMO analysis - originally developed to analyse polar and ionic species – and provides good isomer separation but can struggle to resolve small, similar carbohydrates, e.g. trisaccharides [171]. Coupled with CAD, un-derivatised carbohydrates can be separated using HILIC-CAD, including neutral and acidic HMOs [172]. Nonetheless, LC-based methods require oligosaccharide standards to determine retention time-based structural elucidation, which are frequently not available, or are expensive or impure [167, 170]. Mass-based methods, on the other hand, offer increased sensitivity although isomer separation is not possible without prior column separation, therefore, the most effective methods rely on LC separation followed by MS ionisation and fragmentation to characterise chemical structures [98]. Examples of HMO analysis by LC methods can be seen in Figure 2.1.
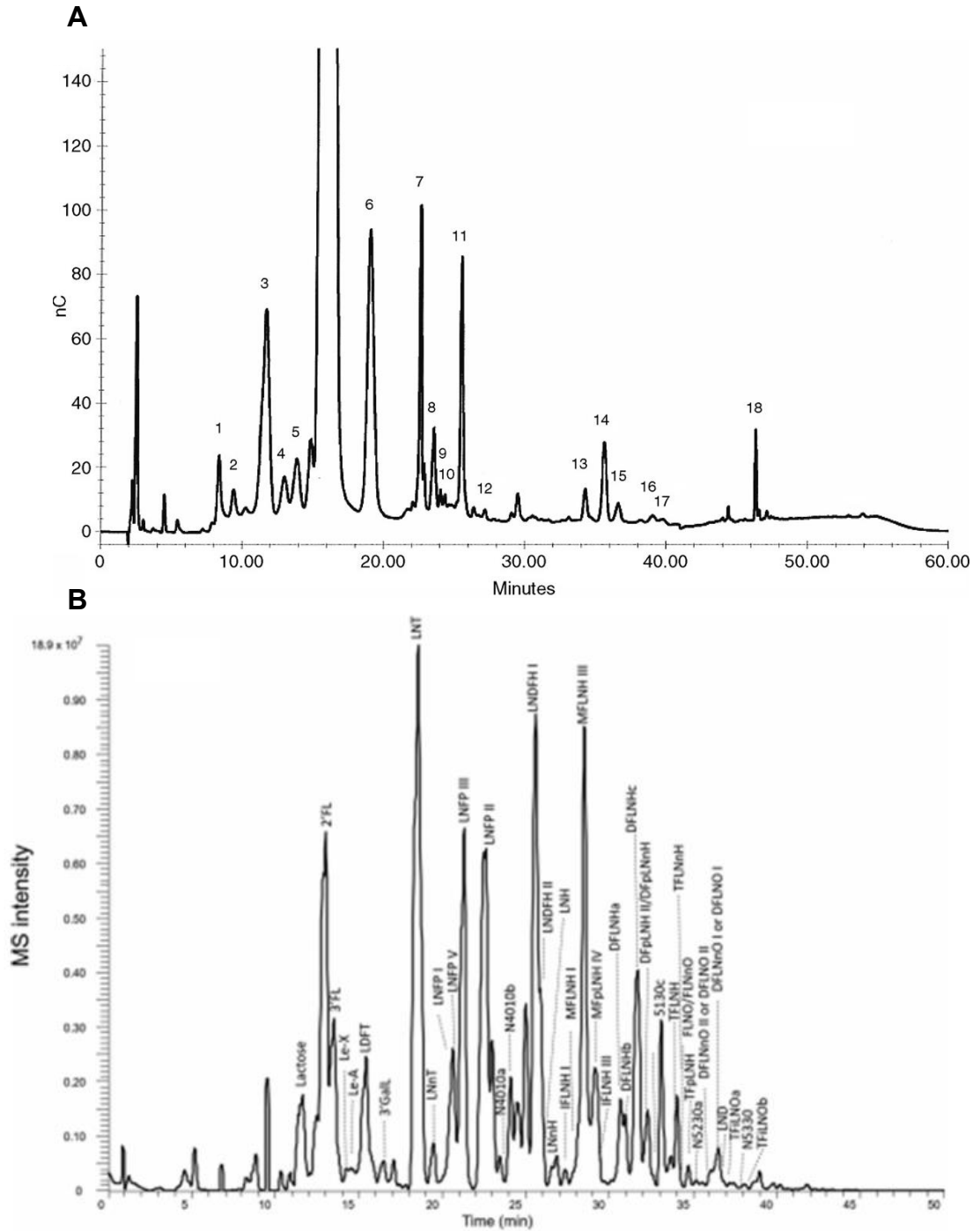
Figure 2.1: Examples of LC methods for HMO analysis. (**A**) High-performance anion-exchange chromatography of oligosaccharides from human milk groups. Peak 1: lacto-*N*-difuco-hexaose II; peak 2: trifucosyllacto-*N*-hexaose; peak 3: difucosyllacto-*N*-hexaose; peak 4: 3-fucosyllactose difucosyllacto-*N*-hexaose; peak 5: lacto-*N*-fucopentaose II; peak 6: 2′-fucosyllactose; peak 7; lacto-*N*-fucopentaose I; peak 8: monofucosyllacto-*N*-hexaose II; peak 9: lacto-*N*-neotetraose; peak 10: lacto-*N*-neohexaose; peak 11: lacto-*N*-tetraose; peak 12: lacto-*N*-hexaose; peak 13: sialylacto-*N*-neotetraose c; peak 14: 6′-sialyllactose; peak 15: sialyllacto-*N*-tetraose a; peak 16: 3′-sialyllactose; peak 17: sialyllacto-*N*-tetraose b; peak 18: disialyllacto-*N*-tetraose. Reproduced with permission from [173]. (**B**) Base peak HILIC-MS analysis of neutral free oligosaccharides in human milk using positive detection mode. Reproduced with permission from [167].

2.2 Results and Discussion

2.2.1 Analysis and Characterisation of HMOs

This section focuses on the development of different analytical approaches for the separation, characterisation and purification of HMOs at analytical and semi-preparative scale. For this work, DSM (formerly Glycom) kindly donated a variety of core and fucosylated HMOs including, but not limited to, 2'-fucosyllactose (2'FL), 3-fucosyllactose (3FL), difucosyllactose (DFL), lacto-N-triose (LNT1), lacto-*N*-neotetraose (LNnT), and lacto-N-fucopentaose-I, -II and -III (LNFPI/II/III). They also provided mixtures of HMOs where the compounds contained within were named, but with no additional information. The aim of this study was to use analytical chemistry techniques to identify and purify the compounds present in the two . These compounds could then either be used as substrates for fucosyltransferase enzymes for enzymatic synthesis of novel HMO analogues or, for the development of other chemoenzymatic approaches towards advanced glycan-based probes. To do this, a range of mass spectrometry and chromatography methods including LC-MS, HPAEC-PAD, HILIC HPLC with CAD were used in the first instance. An alternative, benchtop, approach was also investigated - HILIC cartridges (HILICOM), as a potential larger-scale purification method or pre-purification step.

2.2.1.1 Analysis of HMO Mixtures by HPAEC-PAD

Before any further investigation, the HMO mixtures were first analysed by HILIC LC-MS to confirm the presence of expected masses (see Appendices 2.1). HMO Mix 1 (M1) was reported to contain LNFPIII, 2'FL, lactose (Lac) and LNnT and HMO Mix 2 (M2) was reported to contain LNnT and *para*-lacto-*N*-neohexaose (*para*-LNnH). Analysis of MS spectra for M1 revealed the expected masses for Lac, LNnT and LNFPIII but m/z 512.2 was ambiguous as it could be attributed to either 2'FL, as stated, or its isomer, 3FL. Mass spectra for M2, however, showed clear indications that the mixtures contained more compounds than specified. The masses of Lac, 2'FL/3FL and LNnT were also seen in M2, but the mass for *para*-LNnH could not be seen.

HPAEC-PAD analysis of the HMO mixtures was then carried out and compared to known commercial standards (Lac, 2'FL, 3FL, LNnT and LNFPIII, see Figure 2.2A), which maintain consistent retention times between runs (see Appendices 2.2). HPAEC-PAD of M1 revealed three distinct peaks at 4.54, 6.23 and 9.50 min, respectively. While analysis

of M2 revealed five peaks at, 6.34, 7.06, 7.93, 9.50 and 9.94 min, respectively, despite being reported to only contain LNnT and *para*-LNnH (see Figure 2.2B and C). Therefore, HPAEC-PAD data was found to support the LC-MS findings – M2 contained multiple unknown compounds. Full HPAEC-PAD peak data can be seen in Table 2.1.

Retention times of the commercial standards were compared with sample peaks to assign peak identities, followed by co-injections to confirm assignments. For M1, peak 1 could not be conclusively identified as its retention of 4.54 min was almost identical to both 3FL (4.50 min) and LNFPIII (4.51 min). Therefore, it was thought to be either LNFPIII, 3FL or a mixture of both. Peak 2 eluted at 6.23 min and was tentatively identified as lactose; however, it could also contain 2'FL as commercial standards for lactose and 2'FL were similar, at 6.07 min and 6.53 min, respectively. Peak 3 eluted at 9.50 min and was identified LNnT due to its similarity with commercial LNnT (elution time of 9.46 min).

A series of co-injections was carried out to confirm peak identification, where M1 and M2 was spiked with a known concentration of each commercial standard. For M1 (see Figure 2.3), analysis of HPAEC-PAD data confirmed that peaks 2 and 3 were correctly identified as lactose and LNnT respectively. When the mix was co-injected with LNFPIII the intensity of peak 1 increased from 140 to ~160 nC indicating the presence of LNFPIII. Upon co-injection with 3FL, the intensity of peak 1 remained the same with no additional peaks indicating that peak 1 contained both LNFPIII, with LNFPIII present at a higher concentration than 3FL. Still, this data does not indicate the presence of 3FL, and further analyses were needed. Notably, there was no evidence for the presence of 2'FL, which was stated to be present.

In contrast, M2 was expected to contain two peaks, however, the chromatogram revealed unexpected peaks (see Figure 2.2C). Much like M1, peak 1 (6.34 min) could not be identified as Lac or 2'FL due to similar retention times. Similarly, peaks 2 and 3 (7.06 min and 7.93 min, respectively) did not align with the retention of the commercial standards. Peak 4 (9.50 min) aligned with the retention of LNnT, which was confirmed by co-injection. Peak 5 (9.94 min) was speculated to be *para*-LNnH, however, due to no availability of commercial standards this needed to be confirmed by other analytical methods. Co-injection (see Figure 2.4) revealed that peak 1 consisted of Lac and peak 2, trace amounts of 2'FL. Therefore, as well as LNnT and *para*-LNnH, M2 also contained Lac and 2'FL. Peak 3 could not be identified by HPAEC-PAD.
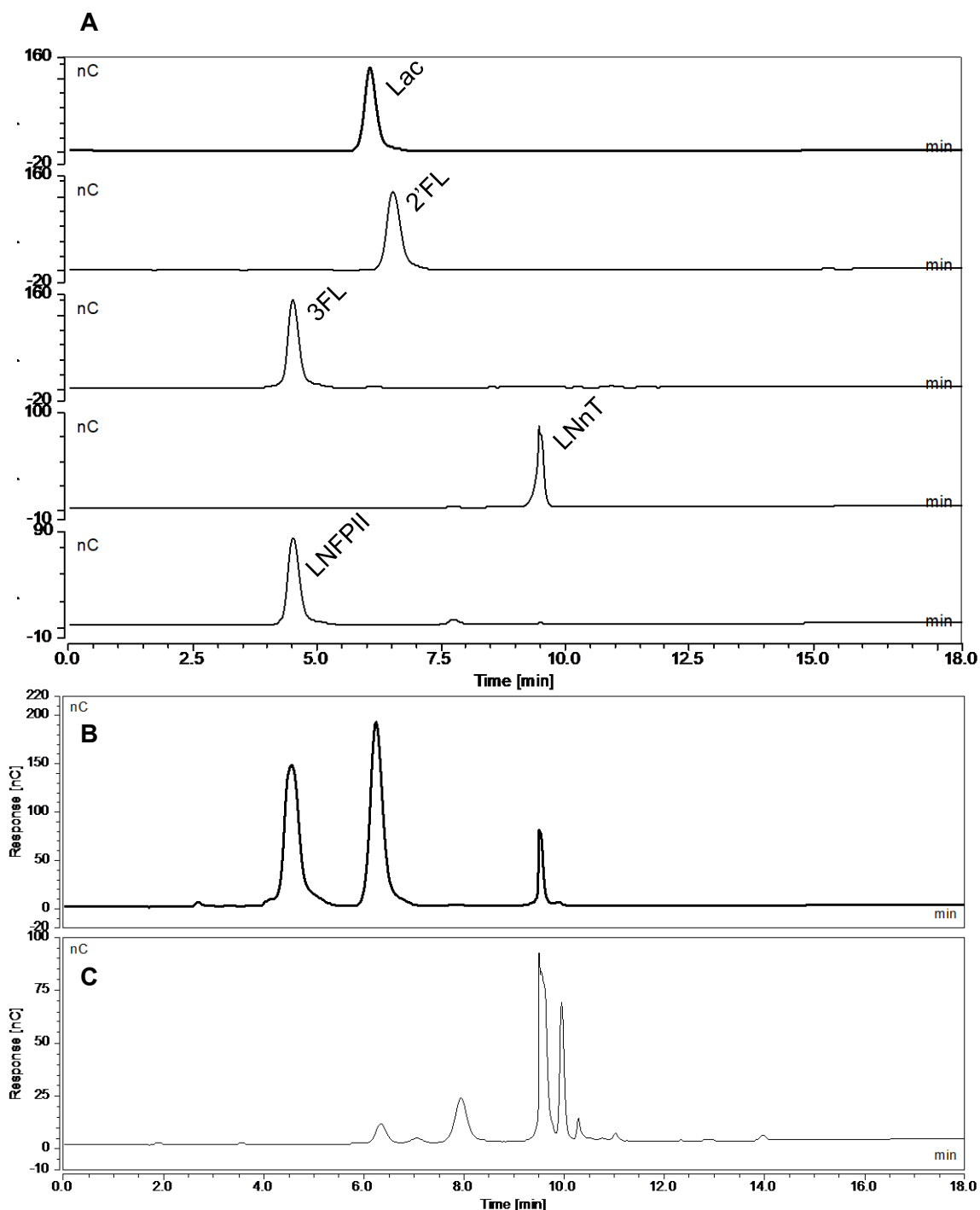
Figure 2.2: Chromatogram of commercial standards (A), M1 (B) and M2 (C) from HPAEC-PAD analysis. (**A**) Commercial standards: Lac, 2'FL, 3FL, LNnT and LNFPIII (top to bottom). (**B**) Peaks 1-3 are provisionally identified as 3FL/LNFPIII, Lac and LNnT (left to right). (**C**) Significant peaks 1-5 are provisionally identified as Lac, 2'FL, unknown, LNnT and *para*-LNnH. Peak identities were assigned by comparing retention times of commercial standards. Samples was injected in water (200 μM and 400 μM, respectively) and eluted with a multistep gradient of NaOH (150mM) and NaOH (150mM) with NAOAc (600mM) over 40 mins (see section 5.2.2).

Table 2.1: HPAEC-PAD data for HMO mixtures, M1 and M2.

Peak No.	Ret. Time (min)	Rel. Area (%)	Area (nC*min)	Height (nC)
M1				
1	4.54	62	47.58	140.27
2	6.23	27	21.09	74.70
3	9.50	11	8.56	77.19
Maximum		62	47.58	140.27
Minimum		11	8.56	74.70
Sum		100	77.231	292.16
M2				
1	6.34	8	2.66	9.72
2	7.06	3	0.86	2.61
3	7.93	22	7.17	21.24
4	9.50	45	14.56	89.54
5	9.94	22	7.36	66.38
Maximum		45	14.56	89.54
Minimum		3	0.864	2.61
Sum		100	32.61	189.49
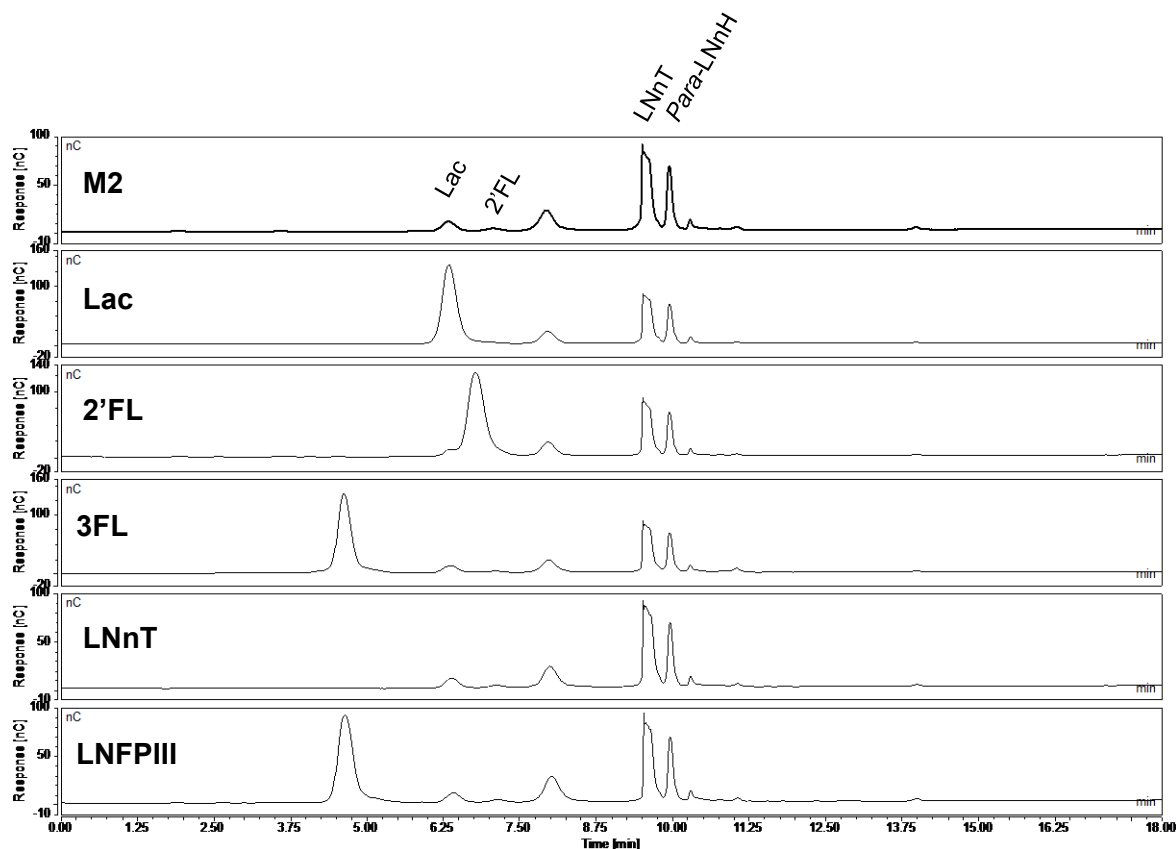
Figure 2.3: HPAEC-PAD chromatograms of M1 co-injections confirming the identity of peaks 1 to 3: LNFPIII/3FL, Lac and LNnT, respectively. Top to bottom: M1 alone followed by co-injections of M1 with Lac, 2'FL, 3FL, LNFPIII and LNnT, respectively. Peaks were confirmed as LNFPIII/3FL, Lac and LNnT, left to right. Samples spiked with 100 µM commercial standards were injected in water and eluted with a multistep gradient of NaOH (150mM) and NaOH (150mM) with NAOAc (600mM) over 40 mins (see section 5.2.2).
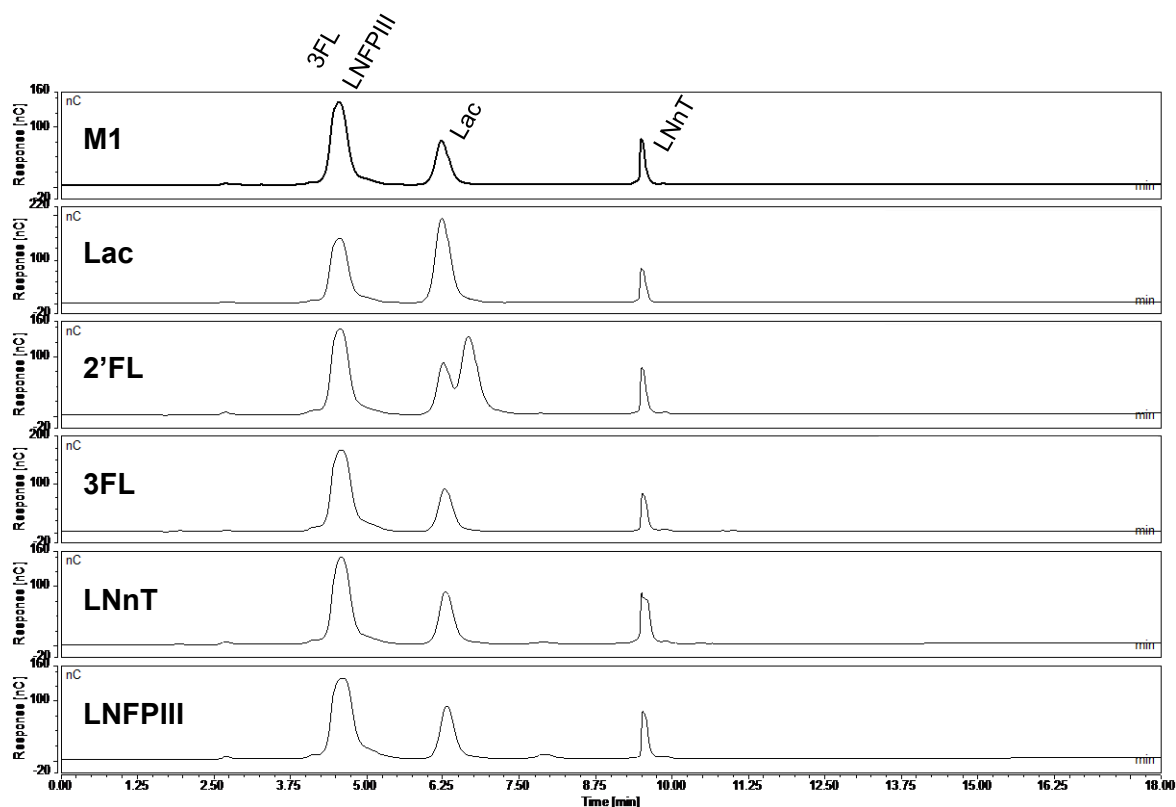
Figure 2.4: HPAEC-PAD chromatograms of M2 co-injections confirming the identity of peaks 1 to 5. Top to bottom: M2 alone followed by co-injections of M1 with Lac, 2'FL, 3FL, LNFPIII and LNnT, respectively. Peaks were confirmed as Lac, 2'FL, LNnT (left to right) but peak 3 could not be confirmed with the above standards and there was no standard for para-LNnH (predicted peak 6). Samples spiked with 100 µM commercial standards were injected in water and eluted with a multistep gradient of NaOH (150mM) and NaOH (150mM) with NAOAc (600mM) over 40 mins (see section 5.2.2).

2.2.1.2 Method Development and Analysis of HMO Mixtures by HILIC HPLC

Due to the high electrolyte content of the eluents and minuscule injection concentrations and volumes for HPAEC, no fractions could practically be collected for further analysis to confirm peak assignments. Therefore, to consolidate the previous findings for M1 and M2, HPLC analysis was carried out using a semi-preparative HILIC column with CAD detection.

Before analysing the HMO mixtures, the HILIC HPLC method was first optimised by elongating the gradient of the organic to the aqueous solvent to obtain baseline separation of compounds. This was done using M1 in the first instance (see Figure 2.5). Method development started from a 12 min method where acetonitrile/water was reduced from 95% to 10% across a linear gradient. The same gradient was then increased to a 40 min method, which improved the separation. The gradient was then further extended by reducing the linear gradient from 95% to 50% acetonitrile/water and baseline resolution

was achieved. As fractions were to be collected manually, the linear gradient was further extended from 95% to 60% acetonitrile/water to achieve better baseline resolution and allow more time for fraction collection.
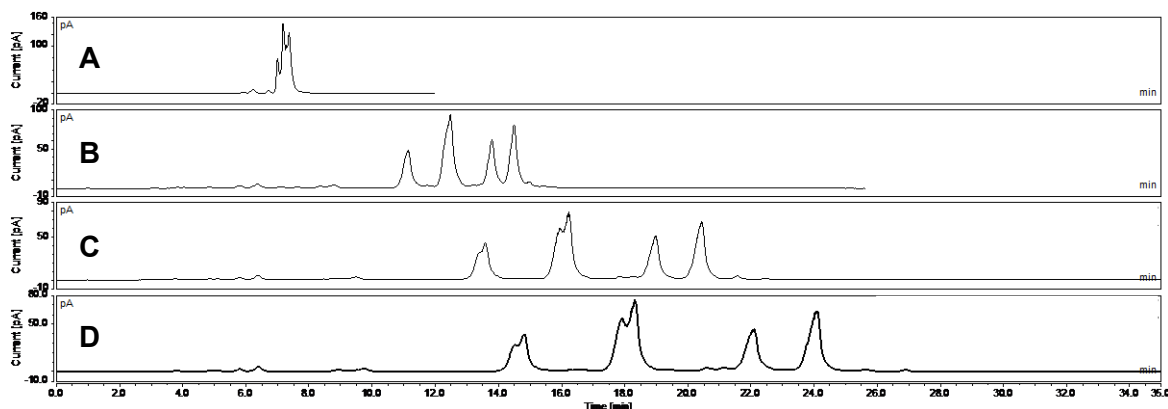
Figure 2.5: HILIC HPLC method optimisation towards baseline resolution of compounds for M1. Sample was injected in water (50 µL, 20 mg/mL) and method development started from a 12 min method where acetonitrile/water was reduced from 95% to 10% across a linear gradient (**A**). The same gradient was then extended over 40 mins (**B**). The gradient was further extended by reducing the linear gradient from 95% to 50% acetonitrile water (**C**) and again, by reducing the linear gradient further from 95% to 60% acetonitrile water to achieve baseline resolution (**D**). See section 5.2.3.

Up on application of HMO mixtures, this method showed four peaks for both M1 and M2, which could be assigned by comparison with commercial standards (see Figure 2.6). Fortunately, unlike in HPAEC, HILIC chromatography was able to resolve all compounds with distinct retention times. When compared against commercial standards, peaks 1-4 were assigned as Lac, 3FL, LNnT and LNFPIII, respectively. This confirms that peak 1 of the HPAEC chromatogram was a mix of 3FL and LNFPIII as predicted from the inconclusive co-injection data. For M2, peaks 1-4 were assigned as Lac, 2'FL, LNnT and para-LNnH, respectively. Furthermore, each peak was collected and further analysed by low-resolution MS to confirm their assignment (corresponding spectra can be found in the appendix, Appendices 2.3 and Appendices 2.4). This was particularly important for para-LNnH for which no commercial standard was available. Peak splitting was seen for both commercial standards and samples, but this is likely due to a difference between the injection solvent (water) and mobile phase (95% acetonitrile).
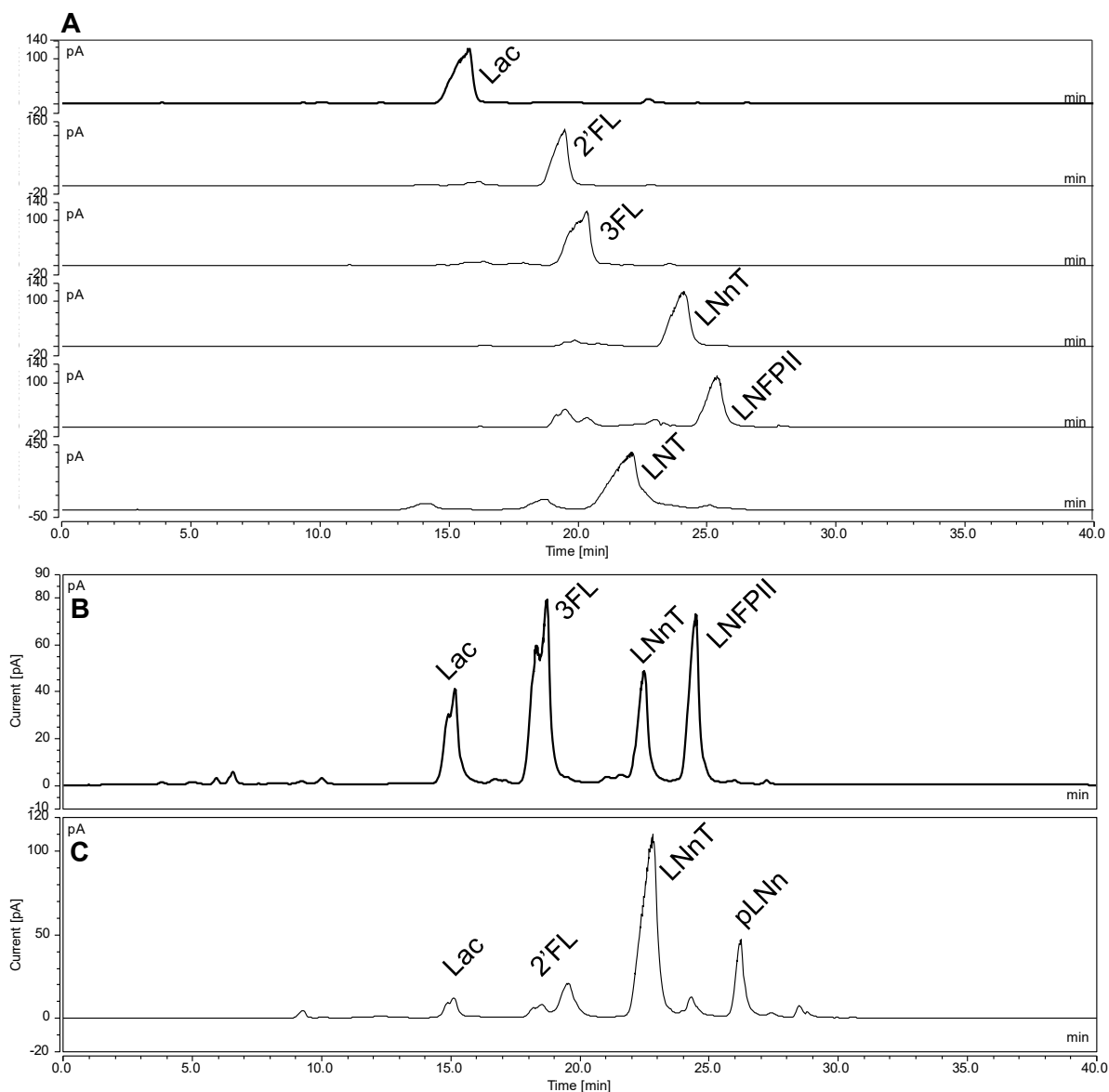
Figure 2.6: Chromatogram of commercial standards (A), M1 (B) and M2 (C) from semi-preparative HILIC HPLC with CAD detection. (A) Commercial standards: Lac, 2'FL, 3FL, LNnT, LNFPIII and LNT (top to bottom). (B) Significant peaks 1-4 were identified as Lac, 3FL, LNnT and LNFPIII (left to right). (C) Significant peaks 1-4 were identified as Lac, 2'FL, LNnT and para-LNnH (left to right). Peak identities were assigned by comparison of retention with commercial standards and confirmed by MS. Samples were injected in water (50 µL, 20 mg/mL) and eluted in a linear gradient of acetonitrile and water from 95% (v/v) to 60% (v/v) over 40 mins (see section 5.2.3).

To conclude this section, using a range of analytical chromatography and mass spectrometry techniques, we are able to successfully characterise and purify mixtures of HMOs. Furthermore, it was found that HMO mixtures, M1 and M2, both contained different HMOs than were quoted by the company. Where M1 was thought to contain LNFPIII, 2'FL, Lac and LNnT, M2 was reported to contain LNnT and *para*-LNnH. From our analysis, M1 contained Lac, 3FL, LNnT and LNFPIII and M2: Lac, 2'FL, LNnT and *para*-LNnH (see Table 2.2). The misidentification of components may be due to the method of synthesis

used by DSM. It is reported that their HMOs are enzymatically synthesised by α-L-fucosidases, which catalyse either the hydrolysis of fucosides or the transfer of fucosyl residues, i.e., transfucosylation [164]. Fucosidase-catalysed synthesis of fucosylated HMOs is often limited by reversible equilibrium and relaxed regioselectivity, which may explain the confusion between 2'FL and 3FL in M1 and the presence of other "impurities" in M2.

Table 2.2: Contents of HMO mixtures as reported by DSM and following HPAEC-PAD and HILIC HPLC-CAD analysis.

HMO Mix 1 (M1)		HMO Mix 2 (M2)	
Before	*After*	*Before*	*After*
Lac	Lac	LNnT	Lac
2'FL	3FL	*para*-LNnH	2'FL
LNnT	LNnT		LNnT
LNFPIII	LNFPIII		*para*-LNnH

2.2.1.3 Purification of HMO Mixtures with HILIC Cartridges

We have previously shown that HILIC HPLC is a lucrative tool for the purification of HMOs from a mixture (see 2.2.1.1). However, there may be more efficient ways for large-scale, benchtop purification of HMOs without the need for a HPLC system. In this study, we developed a method for the use of HILIC cartridges to purify HMOs from a mixture.

To test the method, M1 (now known to contain Lac, 3'FL, LNnT and LNFPIII) and M2 (Lac, 2'FL, LNnT and *para*-LNnH) were used as well as a mixture of commercial standards (Lac, 2'FL, 3FL, LNnT and LNFPIII, 1:1:1:1:1). TLC analysis was used to detect elution of compounds after testing detection of commercial standards. By eluting the mixtures using a gradient of acetonitrile (90 – 72%), the HILIC column was able to effectively separate HMOs across all samples when total carbohydrate loading was 10 mg in 50 µL (200 mg/mL) (see Figure 2.7). However, the cartridge was unable to resolve isomers, 2'FL and 3FL, which mirrors the results from the semi-preparative HILIC HPLC. Moreover, the method does not give complete baseline separation, which is possible with HILIC HPLC methods.

To further evaluate the performance of the cartridge, we increased the total carbohydrate loading of the commercial standards sample from 10 mg to 100 mg in 50 µL. At 10 times the loading concentration, the HILIC cartridge was still able to separate compounds but at

a lower resolution (see Figure 2.8). However, these cartridges are available in larger with up to 2 g of stationary phase. Therefore, it would be expected that these cartridges would be capable of successfully separating HMOs at concentrations exceeding 500 mg in one application. Furthermore, they are reusable, following washing and re-equilibration, and enable easy benchtop separation HMOs from di- to hexasaccharides without the expense and time needed for HPLC applications. Additionally, the cartridge can be used as a pre-purification step before semi-preparative or preparative HILIC HPLC to enable faster, more efficient purification of HMO mixtures and other HILIC compatible compounds. However, a disadvantage of this method is the need for a TLC detection system, which is tedious and labour intensive compared with HPLC in-line detection.

To conclude this section, HPAEC-PAD and HILIC HPLC methods could be optimised for the analysis of HMO mixtures. In combination, and with sufficient commercial standards for peak identification, these methods could be used to characterise mixtures of HMOs, including structural isomers, for which the reported contents were misidentified. These methods could also be used for the semi-preparative purification of such mixtures.

The use of HILIC cartridges was also explored. For the HMO mixtures mentioned above, 500 mg HILIC cartridges were also able to resolve distinct HMOs at 20 mg/mL and 200 mg/mL loading. As such, they enable simple and effective benchtop purification of HMOs or, alternatively, could offer a robust method of pre-purification if selecting for specific compounds, before subsequent HPLC purification.
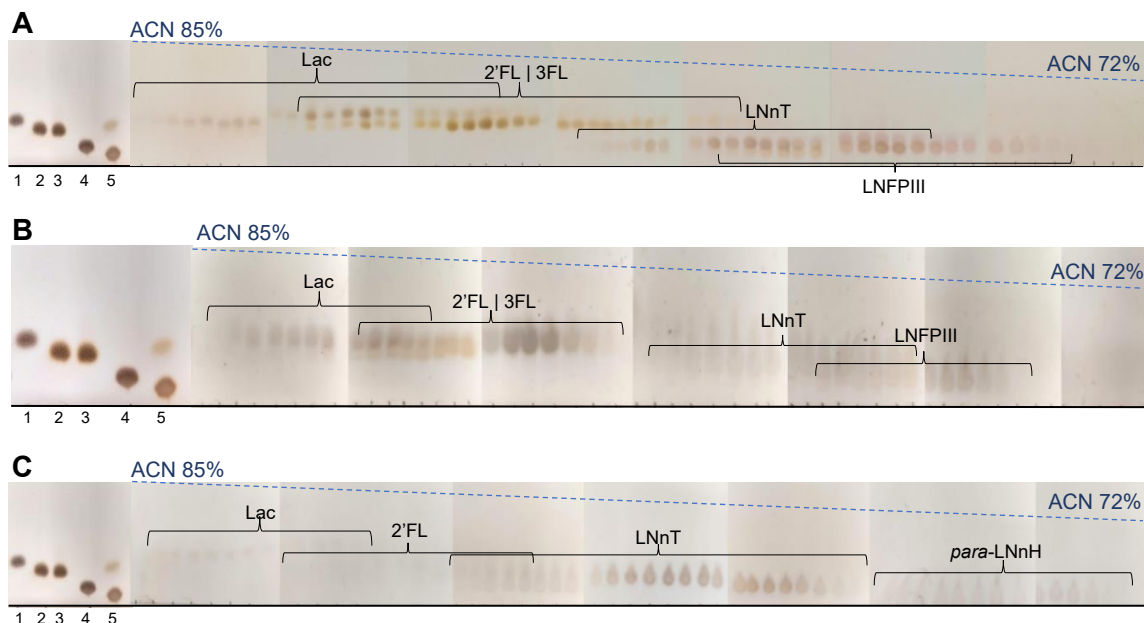
Figure 2.7: TLC analysis of fractions following elution from HILIC cartridge with acetonitrile (90 – 72%). (**A**) Commercial HMO standards: Lac, 2'FL, 3FL, LNnT, LNFPIII (1 – 5), (**B**) M1 and (**C**) M2. A 500 mg HILIC cartridge was equilibrated with 90% acetonitrile (4 mL) after which the sample was applied to the column in water (50 µL, 20 mg/mL). Products were gradually eluted with repeated applications of acetonitrile (4 mL) of decreasing concentration (85%, 82%, 80%, 78%, 76%, 74%, 72%) and 0.5 mL fraction collected. Compounds are eluted as follows: Lac, 2'FL, 3FL, LNnT, LNFPIII, *para*-LNnH.
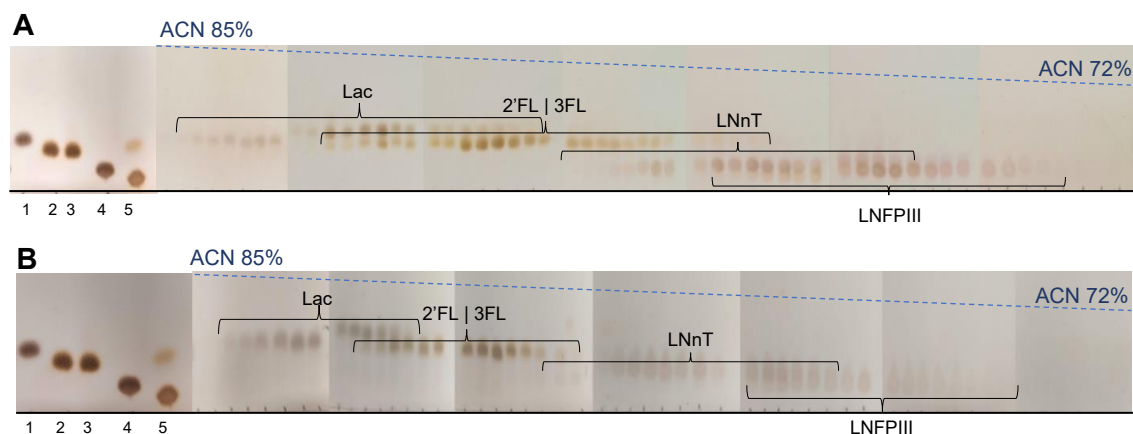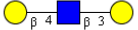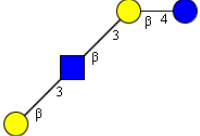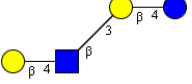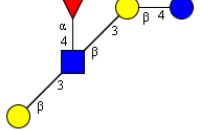
Figure 2.8: TLC analysis of fractions following elution from HILIC cartridge with acetonitrile (90 – 72%). TLC of commercial HMO standards, loading (**A**) 1mg and (**B**) 10mg. A 500 mg HILIC cartridge was equilibrated with 90% acetonitrile (4 mL) after which the sample was applied to the column in water (50 µL). Products were gradually eluted with repeated applications of acetonitrile (4 mL) of decreasing concentration (85%, 82%, 80%, 78%, 76%, 74%, 72%) and 0.5 mL fraction collected. Compounds are eluted as follows: Lac, 2'FL, 3FL, LNnT, LNFPIII.

2.2.2 Investigation of α1,2-Fucosyltransferase from *Helicobacter mustelae* for the Synthesis of HMO Analogues

This study evaluates the ability of HmFucT, a previously unreported enzyme, to function in the enzymatic synthesis of fucosylated HMOs, including their action on natural and chemically modified glycans (Table 2.3). The recombinant protein was obtained from Chemily and reactions were carried out according to the optimal conditions for HpFucT concluded by Stein [132]. Due to the cost of the sugar nucleotide donor, GDP-L-fucose, the bi-functional enzyme FKP will also be examined for two uses: 1) its use in combination with HmFucT for cost effective, scaled up synthesis of fucosylated HMOs; and 2) as a means to access non-natural or modified sugar nucleotide donors and therefore, novel HMO analogues in a two-step one pot strategy.

Table 2.3: List of acceptor substrates for HmFucT.

Compound	Substrate Name	Abbreviation	Structure
1	D-Galactose	Gal	
2	Lactose	Lac	
3	Type I (Lacto-*N*-biose)	LacNB	
4	Type II (*N*-Acetyl-*D*-Lactosamine)	LacNAc	
5	3-Fucosyllactose	3FL	
6	Lacto-*N*-triose I	LNT1	
7	Lacto-*N*-triaose II	LNT2	
8	Lacto-*N*-tetraose	LNT	
9	Lacto-*N*-neotetraose	LNnT	
10	Lacto-N-fucopentaose II	LNFPII	

11	Lacto-N-fucopentaose III	LNFPIII	
12	*1-Azido lactose*	Lac-N$_3$	
13	*3-Aminopropyl lactose*	Lac-AP	
14	*3-Trifluoroacetamidopropyl lactoside*	Lac-triF	
15	*1-Azido-PEG$_4$-3-aminopropyl lactoside*	Lac-PEG$_4$-N$_3$	

2.2.2.1 Synthesis of Functionalised Lactosides

To test the promiscuity of the HmFucT with different functional groups at the reducing end of the sugar with view to access to enzymatically synthesised libraries of functionalised HMOs, lactose was chemically functionalised with either an anomeric azide moiety (**12**), a trifluoracetmidopropyl group (**14**) and a longer linker consisting of azido-PEG$_4$-aminopropyl (**15**) (see Table 2.3).

Introduction of an azide at the anomeric position of lactose would allow copper-catalysed click chemistry to be utilised for conjugation to many different linkers or carrier proteins, such as BSA via a propargyl group. 1-Azido lactose (**12**) was accessed from lactose following Shoda's methodology [174] and functionalisation of the anomeric position of lactose with an azide was achieved as outlined in Figure 2.9.
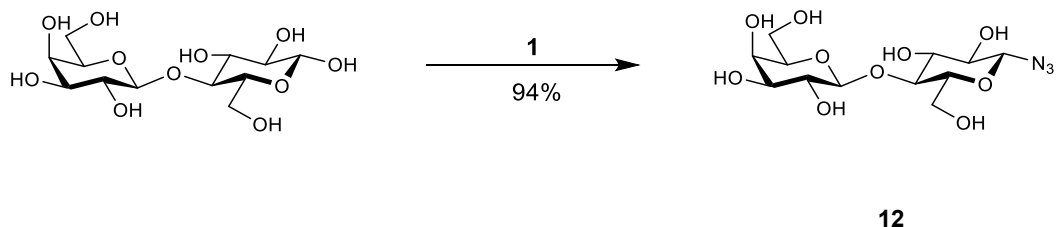
Figure 2.9: Chemical scheme for the synthesis of 1-azido lactose (12). Reagents and conditions: 1. NaN$_3$ (15 eq.), DMC (5 eq.), DIPEA (22.5 eq.), in H$_2$O (0.5 mL), 36 h, RT.

The use of this sterically bulky group trifluoroacetate group offers further insight into the promiscuity of the HmFucT. It also offers purification advantages as the amine group of 3-aminopropyl lactose is protected. In initial fucosylation experiments with 3-aminopropyl lactose, we experienced difficulties purifying the result trisaccharide from unreacted starting material by gel permeation chromatography. As such, the trifluoroacetate group was installed to aid subsequent purification, which could then be removed at a later stage for subsequent conjugation. 3-Trifluoacetamidopropyl lactoside (**14**), was accessed by protecting the amine group of commercial 3-aminopropyl lactose (**13**) by acylation with ethyl trifluoroacetate in sodium carbonate under nitrogen (see scheme outlined in Figure 2.10).
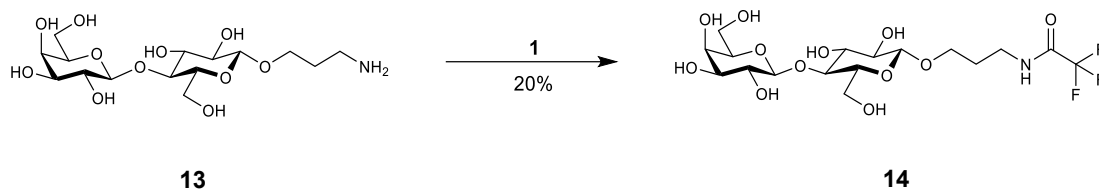
Figure 2.10: Chemical scheme for the synthesis of *3-trifluoracetamidopropyl lactoside* (14). Reagents and conditions: 1. Ethyl trifluoroacetate (2 eq.), sodium carbonate (2 eq.), dry MeOH (1 mL), 18 h, RT.

To access functional groups with a longer, flexible linker region, 3-aminopropyl lactose was coupled with a linker consisting of azido-PEG$_4$-N-hydroxysuccinimide (NHS) ester. The dual functionality of the azido-PEG$_4$-NHS ester allows for specific reactivity of the NHS ester to modify exposed primary amine groups with an azido group via a stable amide bond, while the hydrophilic polyethylene glycol (PEG) spacer increases solubility in aqueous solutions. The azido group could then be further modified by Staudinger reduction to an amine for subsequent functionalisation. 1-azido-PEG$_4$-3-aminopropyl lactoside (**15**) was accessed using a "click" reaction to couple the azido-PEG$_4$-NHS ester linker to 3-aminopropyl lactose (**13**) (see scheme outlined in Figure 2.11).
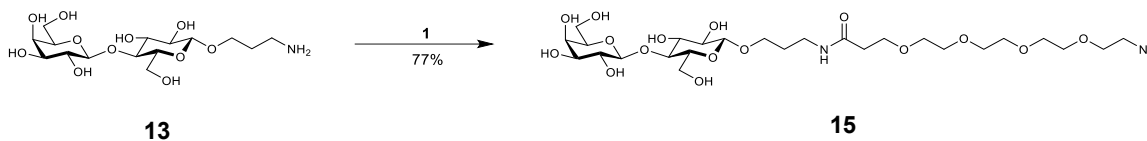
Figure 2.11: Chemical scheme for the synthesis of *1-azido-PEG$_4$-3-aminopropyl lactoside* (15). Reagent and conditions: 1. Azido-PEG$_4$-NHS ester (100 mg/mL in DMF, 1 eq.), in 0.2 mL H$_2$O, 18 h, RT.

2.2.2.2 Acceptor Substrate Specificity of *H. mustelae* α1,2-FucT

A variety of mono-, di-, tri- and oligosaccharides (see Table 2.3) were utilised to determine acceptor specificity of HmFucT. All commercial standards were first purified by HPLC before use due to the presence of impurities (see Appendices 2.5) as discussed in section 2.2.1, from which compound **7** was also obtained. TLC analysis (see Figure 2.12) showed little to no turnover of the substrates (compounds 1 – 6) due to similar retention factors (Rf) despite solvent system optimisation. However, HPAEC-PAD analysis proved a reliable method to quantify the conversion of substrate to product(s) following method optimisation.

1   2   3 4   5   6   R1 R2 R3 R4 R5 R6
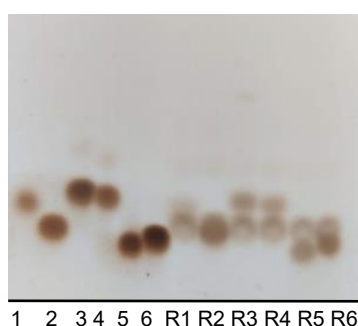
Figure 2.12: TLC analysis of reactions by HmFucT on substrates 1 – 6, including starting materials. Starting materials lanes 1 – 6 (Gal, Lac, LacNB, LacNAc, 3FL, LNT1, respectively). Lanes R1 – R6, α1,2-FucT reactions for compounds 1 – 6).

First, a HPAEC-PAD method was developed for resolution of acceptor and product compounds to enable quantified conversion of substrates by peak integration (see Figure 2.13). To this a CarboPac PA100 was used with a multistep gradient of 150 mM NaOH (buffer A) to 100% 150 mM NaOH with 600 mM NaOAc (buffer B) over 40 mins. This gradient was then extended by increasing buffer B to only 30% over 40 mins but this made little difference to retention of standards Lac and 2'FL. An isocratic elution of compounds in 50 mM NaOH was also tried to encourage longer retention, followed by use of CarboPac PA1 and CarboPac PA20 columns with the same isocratic elution. The CarboPac PA20 column with isocratic elution in 50 mM NaOH gave the best resolution so was used for all future analyses.
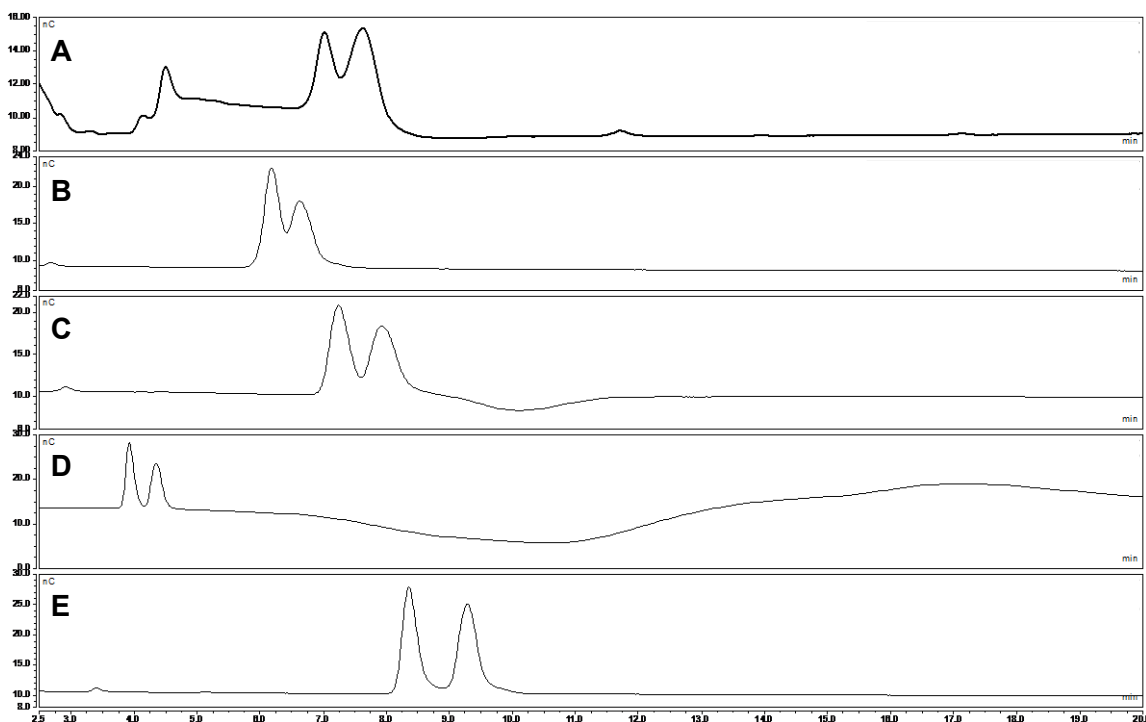
Figure 2.13: HPAEC-PAD method optimisation for monitoring fucosylation reactions. Lactose and 2'FL were used as test samples and the flow rate remained at 0.25 mL/min for each method. (**A**) PA100 column with a multi-step gradient of buffer A (150 mM NaOH) into buffer B (150 mM NaOH with 600mM NaOAc); T_0 = 0% B; T_1 = 0% B; T_{40} = 100% B; T_{45} = 100% B; T_{50} = 0% B; T_{60} = 0% B; (**B**) PA100 with an extended multi-step gradient of T_0 = 0% B; T_1 = 0% B; T_{40} = 30% B; $T_{40.5}$ = 100% B; T_{45} = 100% B; $T_{45.5}$ = 0% B; T_{55} = 0% B; (**C**) PA100 with isocratic 50 mM NaOH; (**D**) PA1 with isocratic 50 mM NaOH and (**E**) PA20 with isocratic 50 mM NaOH. D showed best resolution and was therefore, used for all future analyses. See section 5.2.8.

The reaction for compounds 1 to 11 were analysed by HPAEC-PAD and co-injections were used to identify the starting material from potential product peaks as shown in Figure 2.14 by comparing to commercial standards (see Appendices 2.6) and retention times for all compounds can been seen in Table 2.4. All peaks before retention of 10 min were regarded as non-glycan reaction components due to their commonality in all reactions and the peak present at 14.65 min for compounds **8** to **11** was excluded as a contaminant.

For compound **1**, only one product peak was seen at 17.43 min, with no evidence of the starting material indicating complete conversion, which was confirmed by co-injection. Multiple peaks were seen for compound **2**. Co-injection identified compound **2** as the peak at 18.73 min and the product peak was determined to retain at 20.50 min as is consistent with retention of Lac and 2'FL during method development. For compound **3**, the type 1 disaccharide, the major peak in Figure 2.14C proved to the be the product (16.21 min), with starting material eluting at 20.58 min. While for compound **4**, the type 2 disaccharide, the reverse was true. Where, first major peak at 13.96 min was the starting material with

68

the product retaining at 15.84 min. Similarly, compound **5**, eluted at 11.46 min while its corresponding product eluted at 15.48 min. Compound **6**, was found to retain until 22.37 min, while its product eluted as the first major peak at 15.49 min. No product was seen for compound **7** (20.14 min). Conversely, compounds **8** and **9**, where not visible (see Figure 2.14H and Figure 2.14I), while their products were seen at 22.83 min and 37.03 min, respectively, indicating complete conversion by the enzyme. Compound **10** was identified as the major peak at 12.72 min, and its product at 11.25 min. Finally, for compound **11**, the starting material was seen at 11.86 min, and traces product at 18.13 min. A single fucosylated product was detected for all samples.
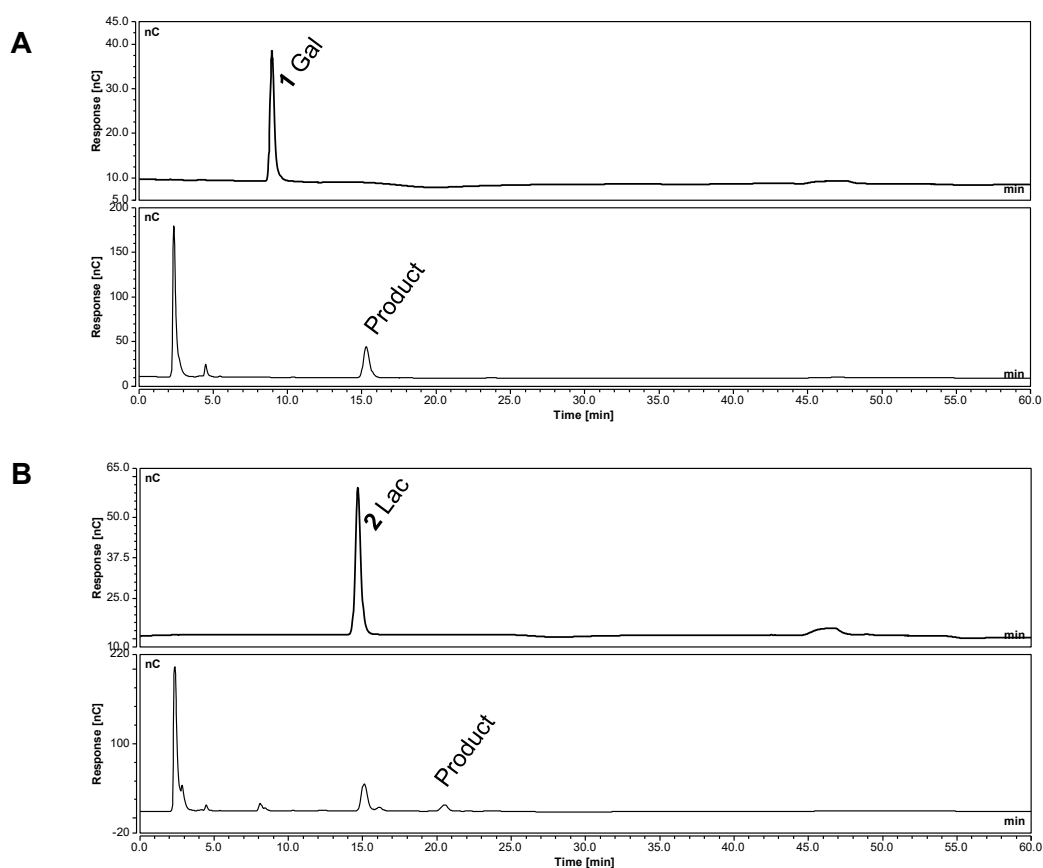
Figure 2.14: HPAEC-PAD analysis of reactions by HmFucT on substrates 1 – 11 (A to K). Starting material (50 µM) is shown above, with the reaction containing the respective compound below, and eluted with isocratic 50 mM NaOH at 0.25 mL/min. Full integration analysis can be seen in Table 2.3.

C

D

E

F

Figure 2.14: continued.

G

H

I

J
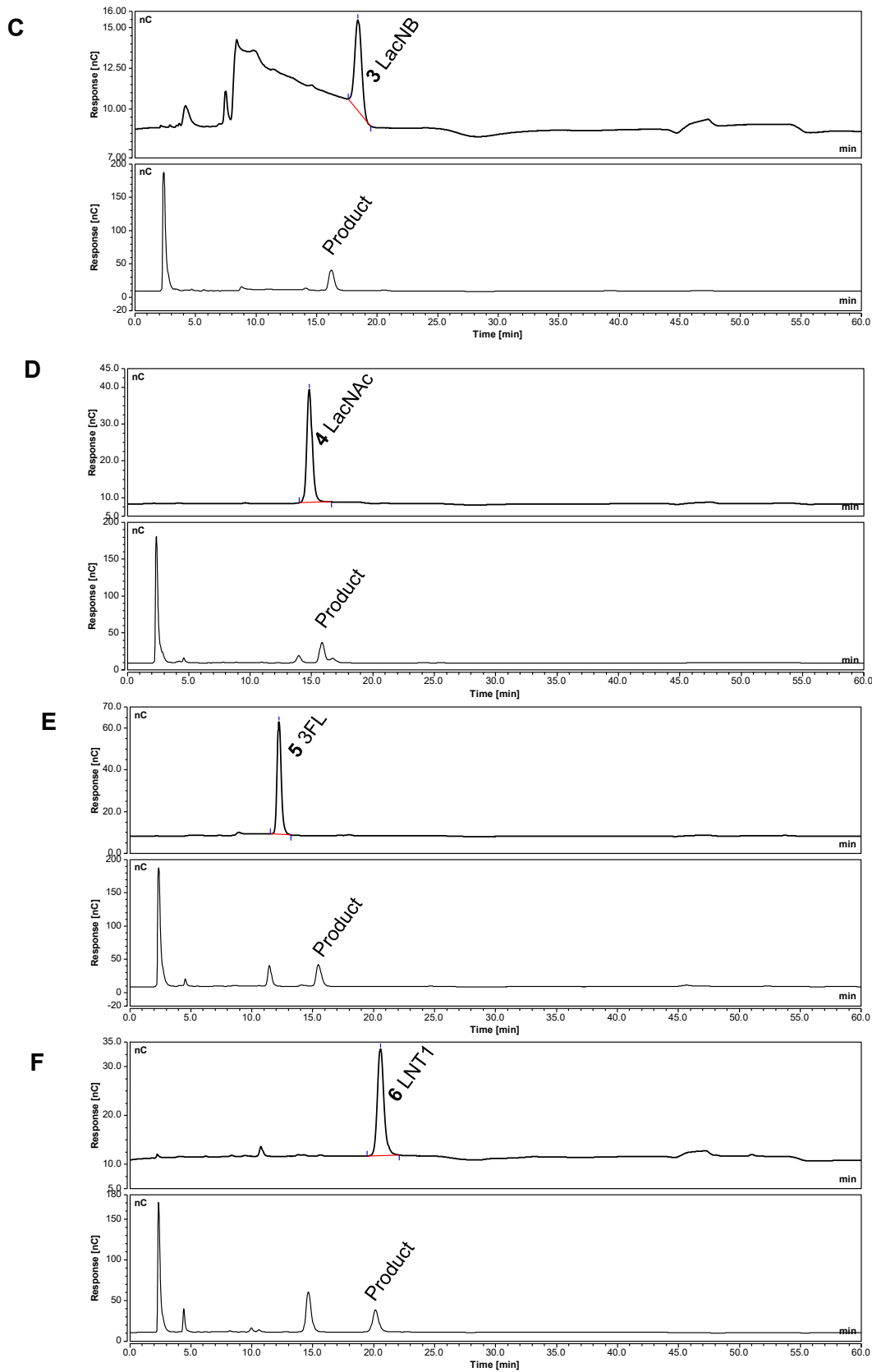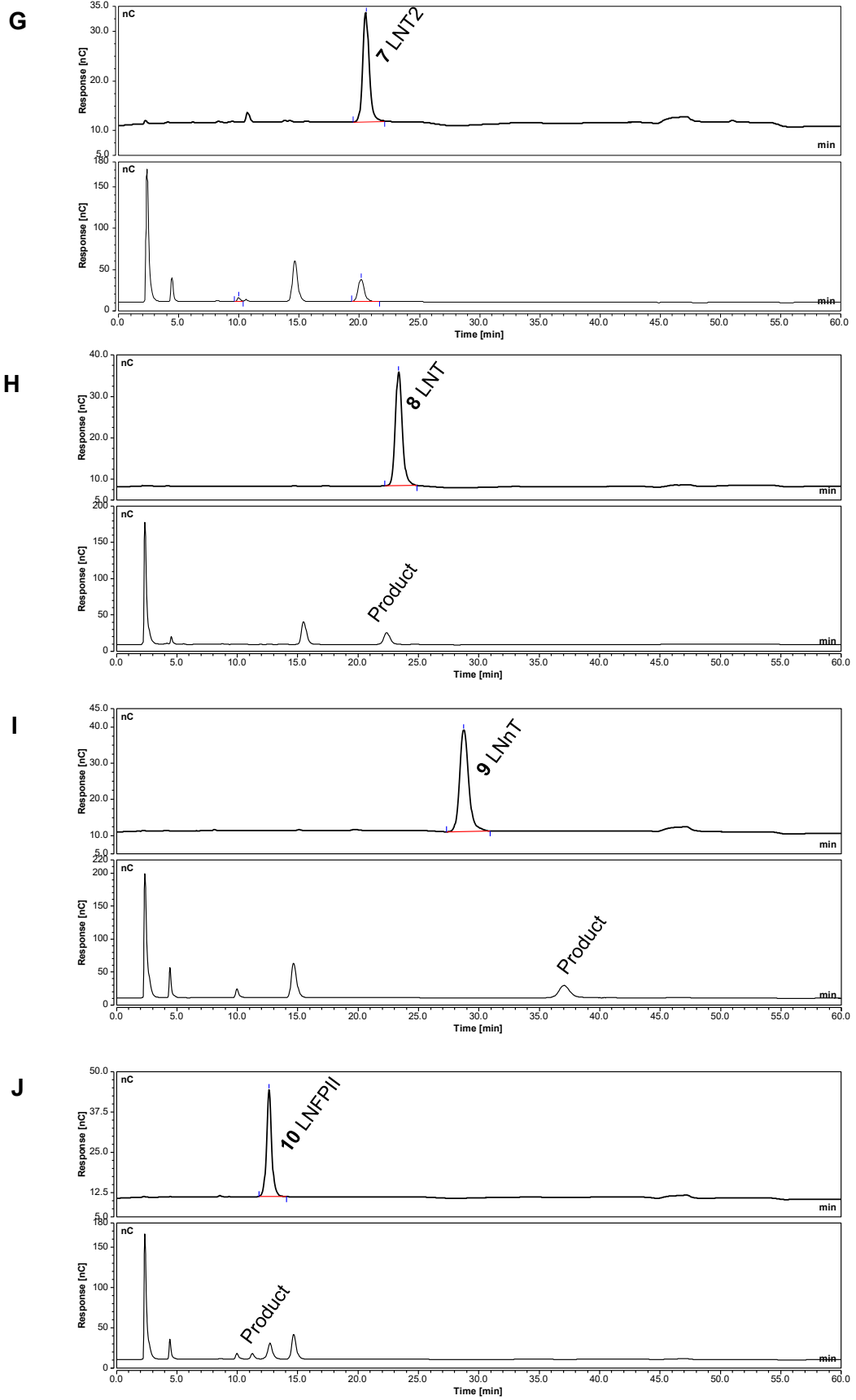
Figure 2.14: continued.
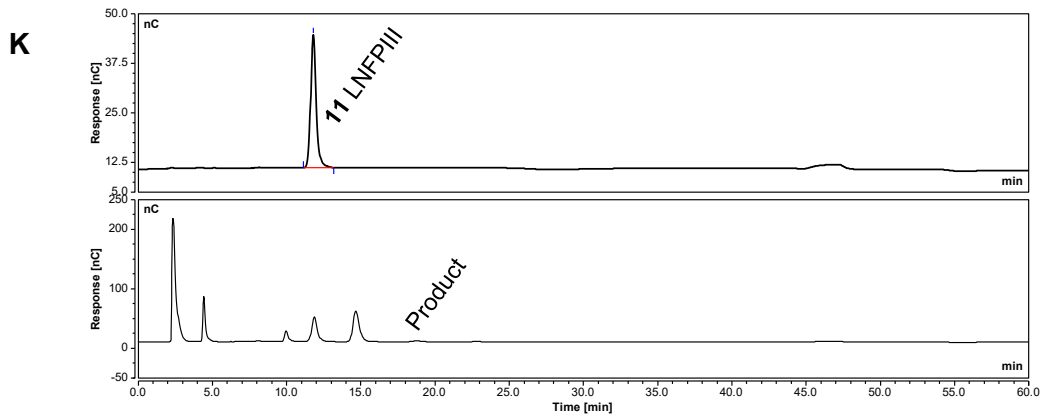
Figure 2.14: continued.

Table 2.4: Summary of HPAEC-PAD analysis for compounds 1 – 11 from HmFucT reactions (see Table 2.3). Peaks are identified as either starting material (SM) or products.

Compound	Peak Identity	Observed Ret. Time (min)	Rel. Area (%)	Area (nC*min)	Height (nC)	Conversion SM (%)
1	Product	17.43	100	12.87	22.28	100
2	SM	16.12	24	1.35	3.75	76
	Product	20.50	76	4.36	8.22	
3	Product	16.21	95	16.13	30.53	95
	SM	20.58	5	0.78	1.17	
4	SM	13.96	29	4.57	9.84	71
	Product	15.84	71	11.10	25.07	
5	SM	11.46	40	11.19	30.98	60
	Product	15.48	60	16.60	32.46	
6	Product	15.49	61	15.92	31.28	61
	SM	22.37	39	10.33	16.18	
7	SM	20.14	100	15.65	26.95	0
8	Product	22.83	100	25.01	34.73	100
9	Product	37.03	100	18.67	18.19	100
10	Product	11.25	21	2.42	6.78	21
	SM	12.72	79	9.34	19.55	
11	SM	11.86	95	18.13	41.24	5
	Product	18.13	5	0.99	1.63	

72
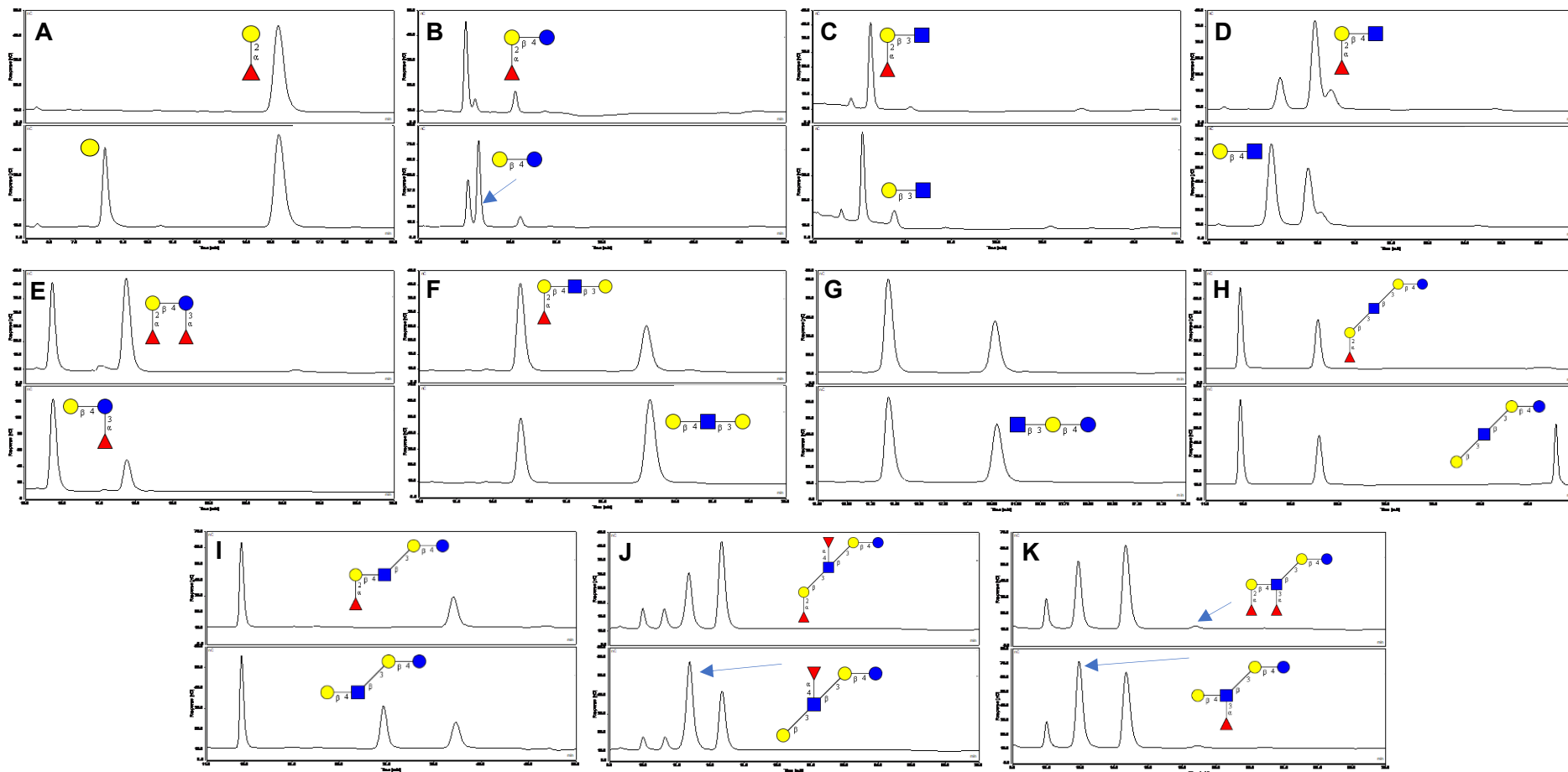
Figure 2.15: HPAEC-PAD chromatograms of the HmFucT reactions (top) and when co-injected with starting material (bottom) for compounds 1 – 11 (labelled A – K, respectively). Samples and co-injections (sample spiked with corresponding commercial standard, 50 µM) were injected in water and eluted isocratically in NaOH (50 mM) over 45 mins before re-equilibration (see section 5.2.8).

The relative activities calculated from HPAEC data are illustrated in Figure 2.16. The data shows that there was a preference toward type 2 glycans, over type I disaccharides. The values obtained for the reaction with type 1 vs. type 2 disaccharides (compounds **3** and **4,** LacNB and LacNAc) were 95% and 71%, respectively. Increased rates of fucosylation were detected when substrates LNT1 (compound **6**) (61%) and LNT (compound **8**) (100%) were tested, indicating that elongation of the type 1 chain did not significantly negatively affect the ability of α1,2-fucosyl transfer and, in fact, appears to improve enzyme activity. Moreover, the monosaccharide, Gal (compound **1**), was determined to be a surprisingly good substrate of HmFucT (100%).

Fucosylation of the elongated type 2-containing glycan, LNnT (compound **9**), was found to be no different to its type I counterpart, LNT (both 100%). However, when the GlcNAc of the pentasaccharide, type 1 (compound **10**, LNFPII) or type 2 (compound **11**, LNFPIII), is fucosylated, a significant decrease in fucosyl transfer activity was observed compared to its non-fucosylated counterpart, LNT or LNnT, e.g., 100% to 21% and 100% to 5%. The same pattern was also seen between compound **2** (Lac) and compound **5** (3FL), but to a lesser extent, where the fucosylated substrate showed slightly decreased activity, 76% and 60%, respectively.

No activity was also observed for compound **7** (LNT2), where there is only an internal Gal present. Therefore, it can be presumed that the resultant products where isomer synthesis was possible (i.e., for compounds **8** to **11**) were exclusively fucosylated at the terminal Gal residue. Equally, LNT1 can be predicted to be fucosylated at either the reducing or non-reducing end Gal. Further studies will need to be done to conclusively confirm which isomer is present.

Different functional groups were introduced to the reducing end of lactose to aid development of further applications and these compounds were also tested for their activity with HmFucT. In this case, the enzyme activity was assessed by processing LC-MS data using an R script (TidyMass) [175] to quantify relative area for each m/z at specified retention times (see Figure 2.17). For all compounds (**12**-**15**), HmFucT was able to successfully generate its corresponding trisaccharide. Introduction of an azido group at the reducing end of lactose (**12,** 1-azido lactose, Lac-N_3) showed conversion of 98.4%; an increase in activity compared to Lac (69 %) of the α1,2-fucosylation. The incorporation of an amine group via a propyl linker at the reducing end of lactose (**13**, 3-aminopropyl lactose, Lac-AP) all but abolished the activity of the HmFucT (2.8%), while acylation of the amine group to a trifluoroacetate (**14**, 3-trifluoacetamidopropyl lactoside, Lac-triF) and of the longer, flexible azido linker (**15**, 1-azido-PEG$_4$-3-aminopropyl lactoside, Lac-PEG$_4$-N_3)

showed similar minimal turnover of 6.6% and 8.5% relative activity by the HmFucT, respectively. A table summary table of conversion for all compounds can be seen in Appendices 2.7. Mass spectra for each compound can be seen in Figure 2.18. Therefore, the reducing end azide group was the most well-tolerated, causing only almost no difference in the relative activity of the enzyme. To understand the reasons for the lack of catalysis of some compounds, i.e., compounds **7**, **10**, **11**, **13**, **14**, **15**, future studies could include saturation transfer difference NMR to determine transient glycan-protein binding.
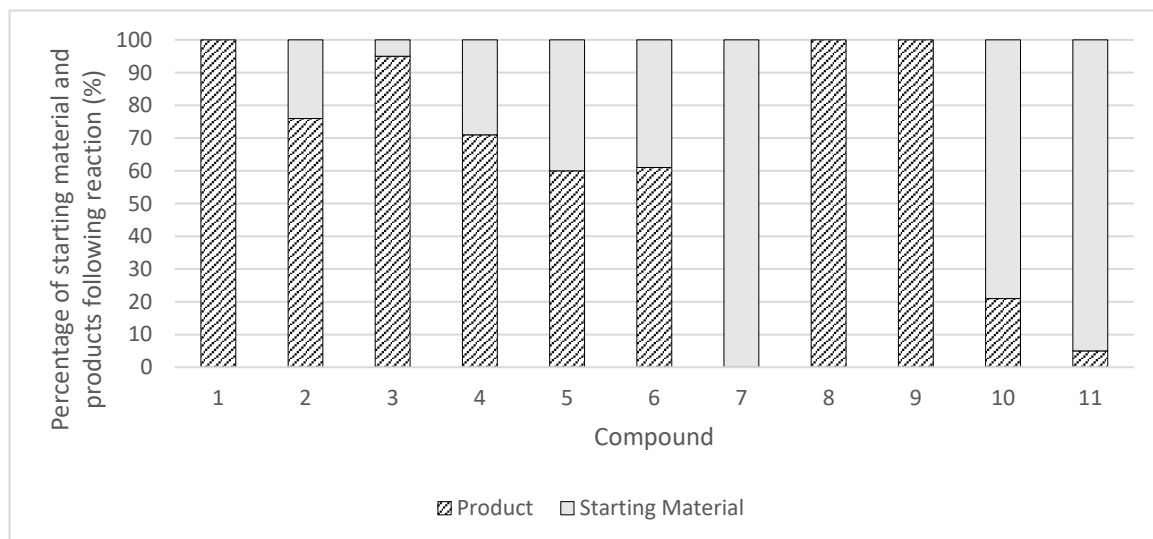
Figure 2.16: Substrate specificity of HmFucT (Compounds 1 to 11). Shown as percentage conversion from starting material to total products as determined by HPAEC-PAD. Relative values were determined with the indicated acceptor saccharides and illustrated in a vertical bar chart. For comprehensibility, the applied substrates are labelled by compound number. The respective structures are assigned in Table 2.3.
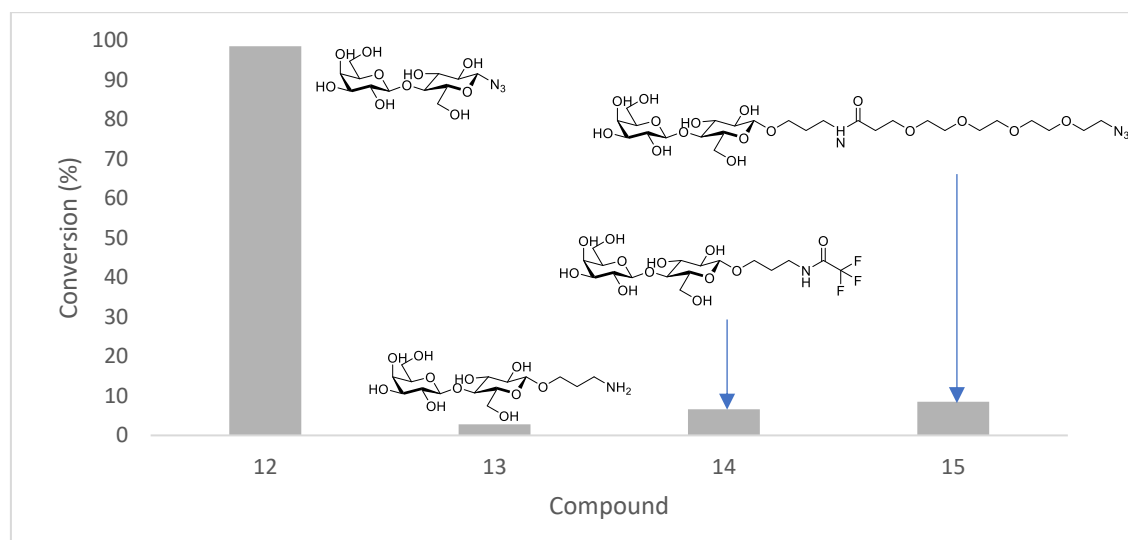
Figure 2.17: Substrate specificity of the HmFucT for compounds 12 to 15. Shown as percentage conversion from starting material to total products as determined by HILIC LC-MS. Relative values were determined with the indicated acceptor saccharides and illustrated in a vertical bar chart. For comprehensibility, the applied substrates are labelled by compound number. The respective structures are assigned in Table 2.3.

75
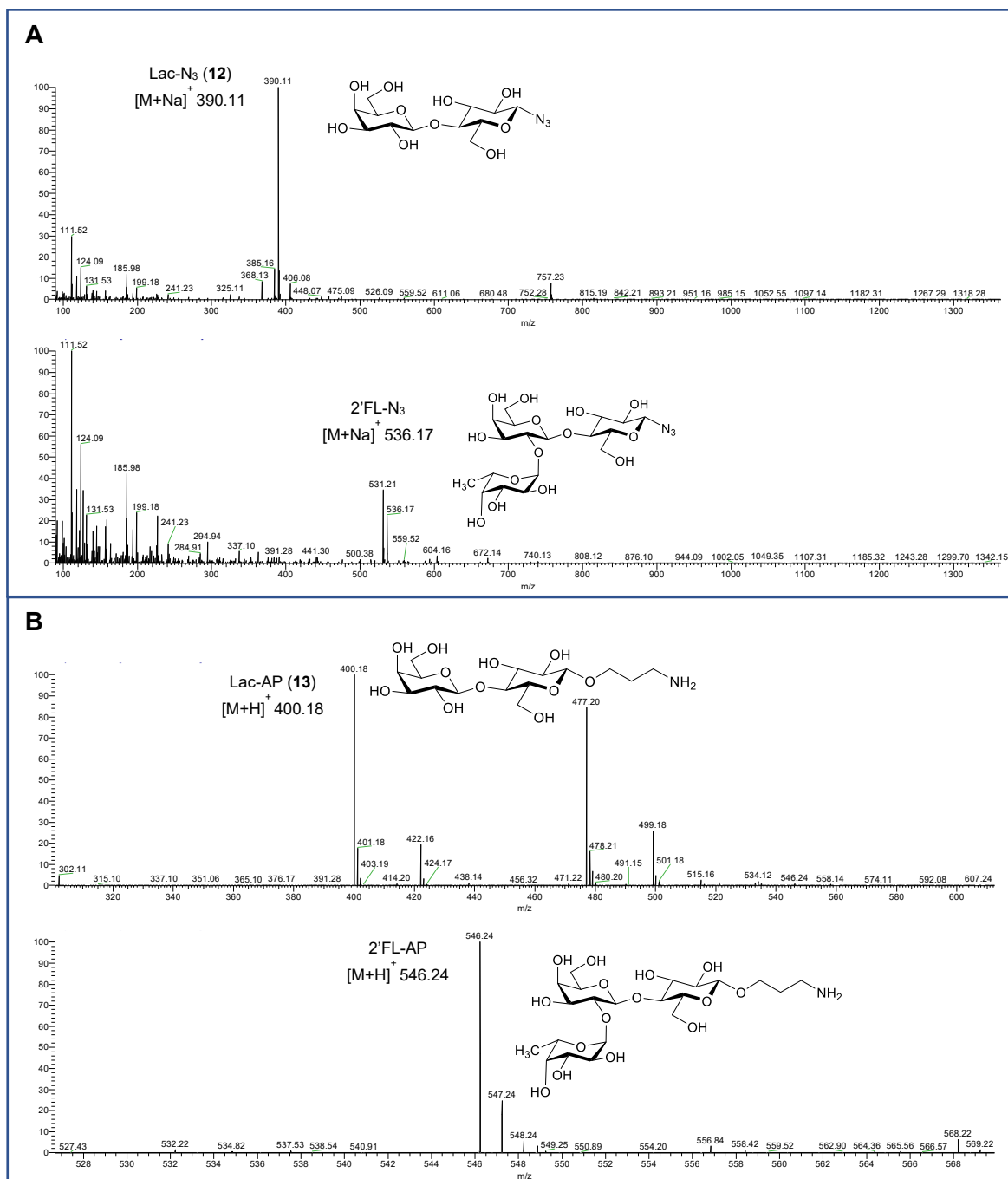
Figure 2.18: Mass spectra of HmFucT reaction with compounds 12 to 15. (**A**) Top: compound **12** (Lac-N3); bottom: its respective fucosylated product, both seen as sodium adducts. (**B**) Top: compound **13** (Lac-AP); bottom: its respective fucosylated product, both seen as protonated adducts. (**C**) Top: compound 14 (Lac-triF); bottom: its respective fucosylated product, both seen as proton and sodium adducts. (**D**) Top: compound **15** (Lac-PEG4-N3); bottom: its respective fucosylated product, both seen as proton and sodium adducts.
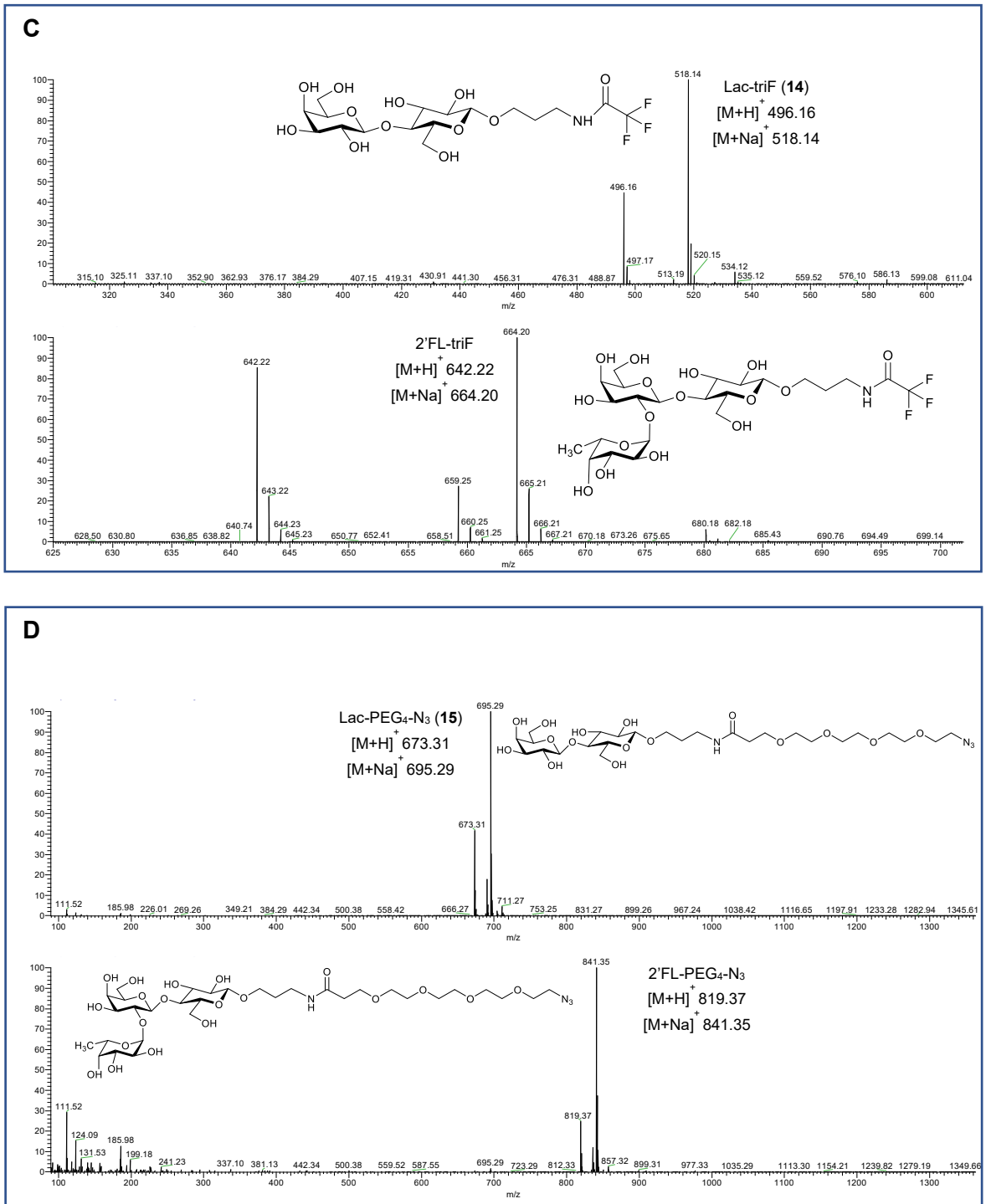
Figure 2.18: continued.

77

Bacterial fucosyltranferases have proved to be useful tools for the synthesis of biorelevant and disease-associated oligosaccharides. Although, putative genes encoding FucTs have been identified in various bacteria, *H. pylori* is the main organism from which the FucT enzymes used in research originate. However, it should be noted that there has also been documented use of α1,2-FucT from *Thermosynechococcus elongatus* in a one pot multienzyme reaction preparative scale synthesis of LNFPI [113].

Of the papers that have investigated the substrate specificity of HpFucT against a range of acceptors [132], it was found to show distinctly different relative activity to the HmFucT. The HpFucT showed a substantial preference towards the type 1 disaccharide over type 2 (relative activity 100% vs. 20%, respectively), while the conversion of both compounds by H. mustelae α1,2-FucT was 95% to 71%, respectively. Similarly, while the conversion of LNT and LNFPIII were comparable in both studies, the activity towards LNFPII was significantly lower in the case of the HpFucT (59% compared to ~10% activity). Lactose and L-Gal also showed strikingly different conversions of 76% and 100%, respectively by HmFucT compared to ~30% and ~20% by HpFucT. Furthermore, to the best of my knowledge, LNT1, LNT2 and LNnT have not previously been tested with any α1,2-FucT, nor has the ability of HmFucT to catalyse the fucosylation of functionalised lactose acceptors (compounds 12 -15). However, it should be noted that there were differences in both the analysis method and reaction composition, i.e., HPAEC-PAD integration vs. kinetic assay using GDP-[^{14}C]-fucose. Though, HPAEC-PAD has been known the suffer from lack of sensitivity and lower than expected sample abundance [167], this would be not explain the consistently increased conversion of acceptors seen with HmFucT. It is hoped that in the future, synthesis of HMOs will become increasingly commonplace as enzymatic routes and purification techniques, and as multiomics approaches discover more genes involved in HMO biosynthesis [176].

2.2.2.3 Isomer Determination of Fucosylated HMO Analogues

To confirm the identity of fucosylated isomers for the products of compounds **6**, **8**, **9**, **10**, **11** (LNT1, LNT, LNnT, LNFPII and LNFPIII), all reactions were first purified by HILIC HPLC to obtain the fucosylated product, and fractions analysed by MS (see Figure 2.19 and Appendices 2.8). Of the fractions collected, a fucosylated product could not be identified for LNT1 and all other products collected were mono-fucosylated, which aligns with the HPAEC-PAD data (see Figure 2.15).
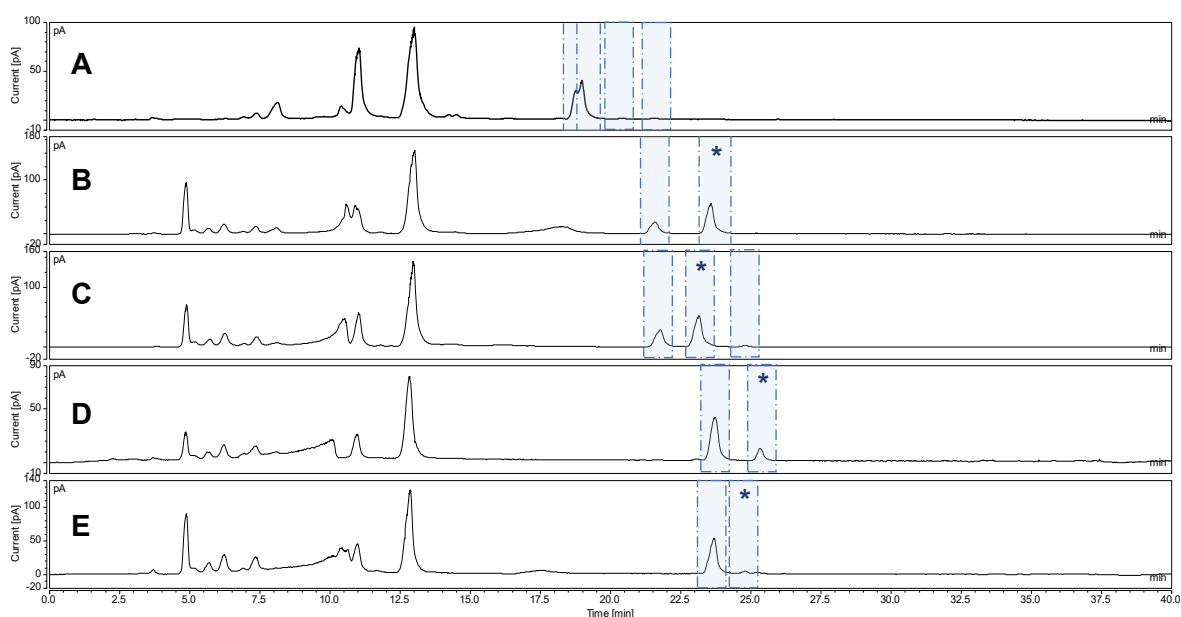
Figure 2.19: HILIC HPLC chromatogram showing purification of α1,2-FucT reaction products from LNT1, LNT, LNnT, LNFPII and LNFPIII (A-E). Fractions collected are marked with a blue box and * marks fractions containing the fucosylated product, as confirmed by LR-MS. All other fractions collected were identified as the corresponding starting materials. Remaining reaction samples were injected in water and eluted in a gradient of acetonitrile and water (95% to 60%) over 34 mins before re-equilibration (see section 5.2.9).

Although data from compound **7** indicated the fucosylation of the internal Gal residue was not possible by HmFucT, a digestion assay using β-1,3/4-galactosidase was used. Figure 2.20 shows the HPAEC-PAD chromatogram for the β-1,3/4-galactosidase digestion of mono-fucosylated LNT, LNnT, LNFPII and LNFPIII. The negative control (A, 2'FL) and positive control (B, Lac) are clearly visible and confirm the expected activity of the enzyme, i.e., the release of Gal/Glc from Lac (red). Unfortunately, despite method optimisation, there are no visible peaks for either the starting material or released Gal. As an alternative method of analysis, the reactions were also analysed by LC-MS but again, neither the fucosylated starting material nor D-Gal could be detected. It should be noted that there was also presence of a peak of 180.97 m/z in all samples, which may have obscured a D-Gal peak (predicted $[M+H]^+$ 181.2). Further investigations are needed to establish the identity of the isomers synthesised by the HmFucT, however, the lack of fucosylation observed for compound **7** indicates that only the terminally fucosylated isomer was present in all reactions. It follows that all other compounds are likely fucosylated at the terminal Gal and not at other positions, which would also explain the catalysis of only one fucosyl transfer.
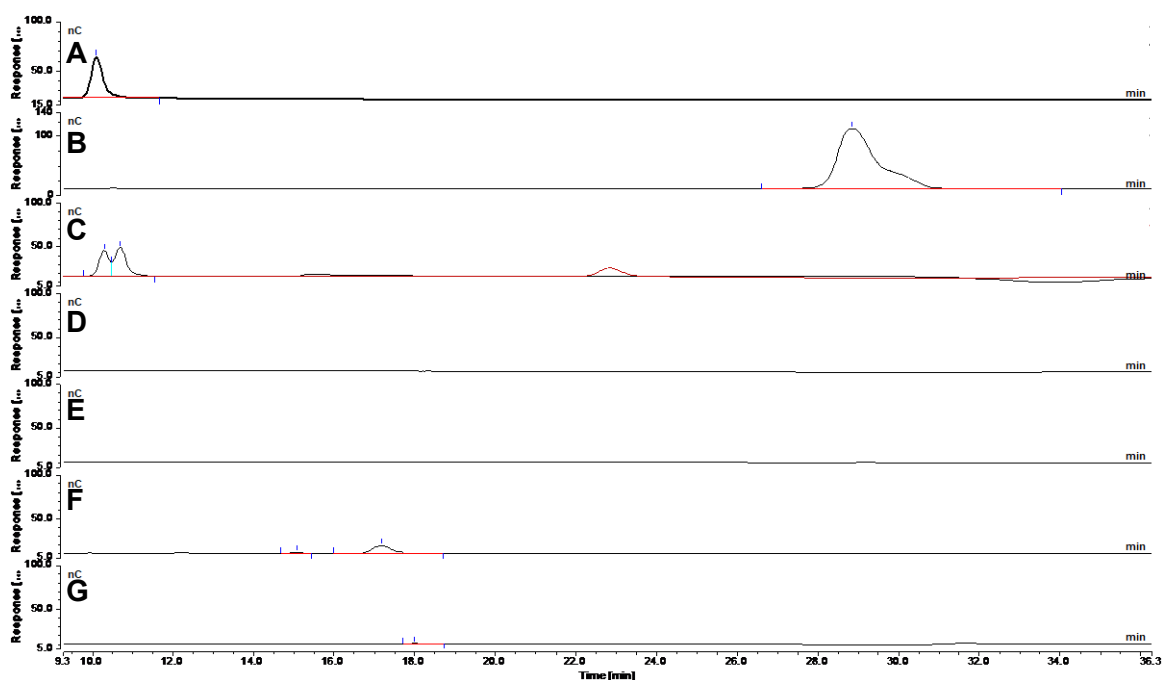
Figure 2.20: HPAEC-PAD chromatogram of β1,3/4-galactsidase digestions of fucosylated products from LNT, LNnT, LNFPII, LNFPIII (D – G, respectively). D-Gal was used as a standard (**A**), and 2'FL (**B**) and Lac (**C**) were used as negative and positive controls, respectively. 2'FL showed no release of D-Gal (B), while Lac (red standard) was hydrolysed to D-Gal and D-Glc (**C**). Neither the starting material nor D-Gal can be seen in the subsequent traces (**D-G**). Samples were analysed by HPAEC-PAD from 50 µM injections and eluted isocratically in 50 mM NaOH over 40 mins (see section 5.2.9).

2.2.2.4 Synthesis of HMO analogues using one-pot reactions with FKP

GDP-fucose derivatives are key intermediates for the synthesis of fucosylated oligosaccharides and glycoconjugates. L-fucokinase/GDP-L-fucose phosphorylase (FKP) has already been used as part of one pot enzymatic routes towards fucosylated oligosaccharides, such as Lewis X antigens and some HMOs [113, 146, 177] and with some azido and alkynyl non-natural substrates [144, 145]. Here, we investigate the efficiency of FKP from *Bacteroides fragilis* to convert L-fucose to GDP-L-fucose independently and in one-pot reactions with HmFucT. First, the promiscuity of the FKP enzyme with different non-natural substrates was examined as a potential method toward new routes for the synthesis of non-natural HMO analogues, incorporating D-arabinose (D-Ara) or L-galactose (L-Gal), which all form six-membered rings with identical stereoconfigurations but differ in their functionality at C-5 (see Figure 2.21). These D-arabinosylated and L-galactosylated HMO analogues would provide unique opportunities to probe the molecular interactions of different monosaccharides and their chemical

groups. Importantly, the monosaccharides used were commercially available. Following the confirmation of GDP-sugar nucleotide formation, HmFucT will be added to the crude reaction to form a two-step one pot reaction for the scalable synthesise of non-natural HMO analogues (see Figure 2.22).
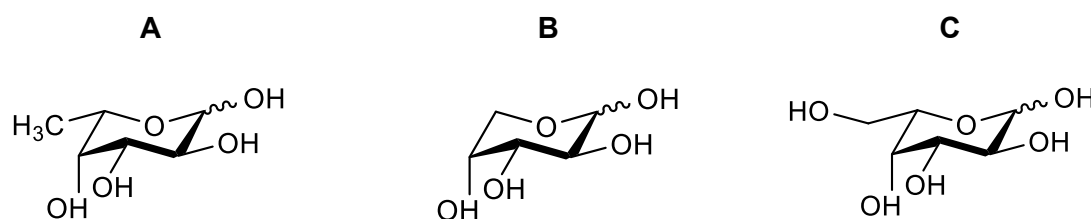
Figure 2.21: Structure of L-Fuc (A), D-Ara (B) and L-Gal (C), which all form six-membered rings with identical stereoconfigurations.
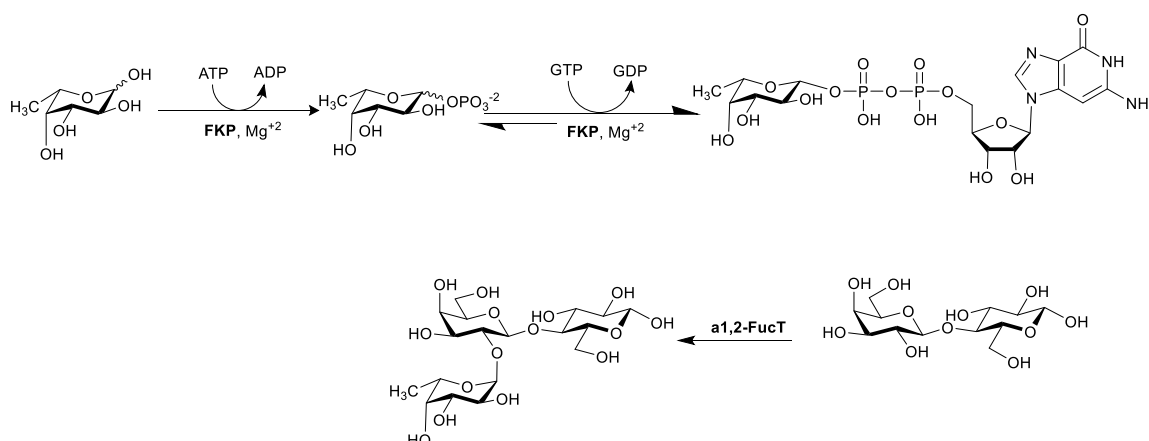
Figure 2.22: Onepot enzymatic scheme for the formation of fucosylated HMOs using FKP from *B. fragilis* and α1,2-FucT from *H. mustelae*.

FKP was expressed in-house and the SDS-PAGE gel (Figure 5.1). First, the reaction conditions of substrate concentration (L-Fuc) and metal ion (Mg^{2+} or Mn^{2+}) were optimised based on the methods of Wang, W., *et al.,* 2009 [146]. By examining differences in approximate substrate conversion by TLC, there was a slightly improved activity of FKP when Mn^{2+} was used in place of Mg^{2+}. Yet, due to the use of Mg^{2+} as the metal ion for hmFucT, Mg^{2+} was chosen to take forward for subsequent reactions. However, there was a marked difference between the FKP activity at 5 mM and 10 mM substrate, where FKP

activity decreased as substrate concentration increased (see Figure 2.23). The enzyme appeared to become self-inhibited by the substrate at higher concentrations, which does not appear to be noted in the literature. Yet, it should be noted that in nature, the cascade from GDP-D-Man to GDP-L-Fuc in the de novo pathways is inhibited by the GDP-L-Fuc, as it is a potent competitive inhibitor of GDP-D-Man-4,6-dehydratase (GMD) [178, 179].
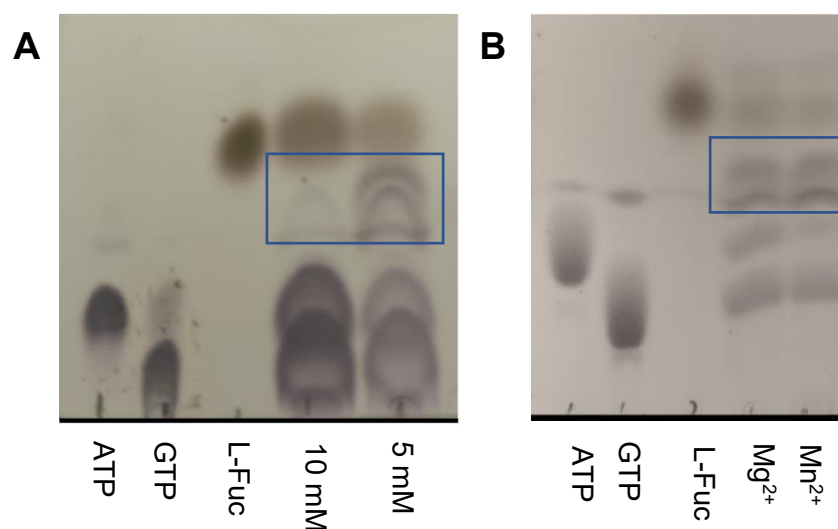
Figure 2.23: TLC analysis of reaction optimisation for FKP with L-Fuc substrate. (**A**) Optimising substrate concentration (5mM or 10mM) where 5 mM shows improved activity as evidence by an increase in GDP-Fuc production (blue square); and (**B**) metal ion concentration (Mg^{2+} or Mn^{2+}), where activity is similar with both metal ions.

In a typical enzymatic reaction, a fucose analogue (5 mM) was incubated with the recombinant FKP in the presence of one equivalent each of ATP and GTP. An inorganic pyrophosphatase from *Saccharomyces cerevisiae*, containing $MgCl_2$ to maintain its maximal activity, was included to hydrolyse pyrophosphate generated in the reaction and to drive the reaction to completion. The reaction progress was followed by TLC analysis and strong anion-exchange (SAX) chromatography to quantitatively monitor product formation. Figure 2.24 shows the SAX analysis of the reactions. Peaks were identified by comparison of retention times with those of commercial nucleotide standards: adenosine monophosphate (AMP), adenosine 5'-diphosphate (ADP), adenosine 5'-triphosphate (ATP), guanosine monophosphate (GMP), guanosine 5'-diphosphate (GDP) and guanosine 5'-triphosphate (GTP) (for retention times, see Appendices 2.9). FKP catalysed the conversion of all substrates tested as shown by the presence of peak 4 of the SAX chromatograms, as indicated by the blue arrow. Integration of ATP, GTP and peak 4 revealed that FKP catalysis of L-Fuc to its sugar nucleotide product, GDP-L-fucose (GDP-

Fuc), was most efficient (36% relative area) followed by L-Gal to GDP-L-galactose (GDP-Gal) and D-Ara to GDP-D-arabinose (GDP-Ara), and at 31% and 19% relative area, respectively. High-resolution mass spectrometry confirmed the presence of GDP-sugar nucleotide products (see Figure 2.25). In all cases, GTP was almost completely depleted from the reaction mixture despite being present in equal ratio in the reaction, indicating that it may be the limiting reagent to further GDP-sugar nucleotide formation. Alternatively, the availability of monosaccharide may be limiting, although this is not visible by SAX or clear by TLC.
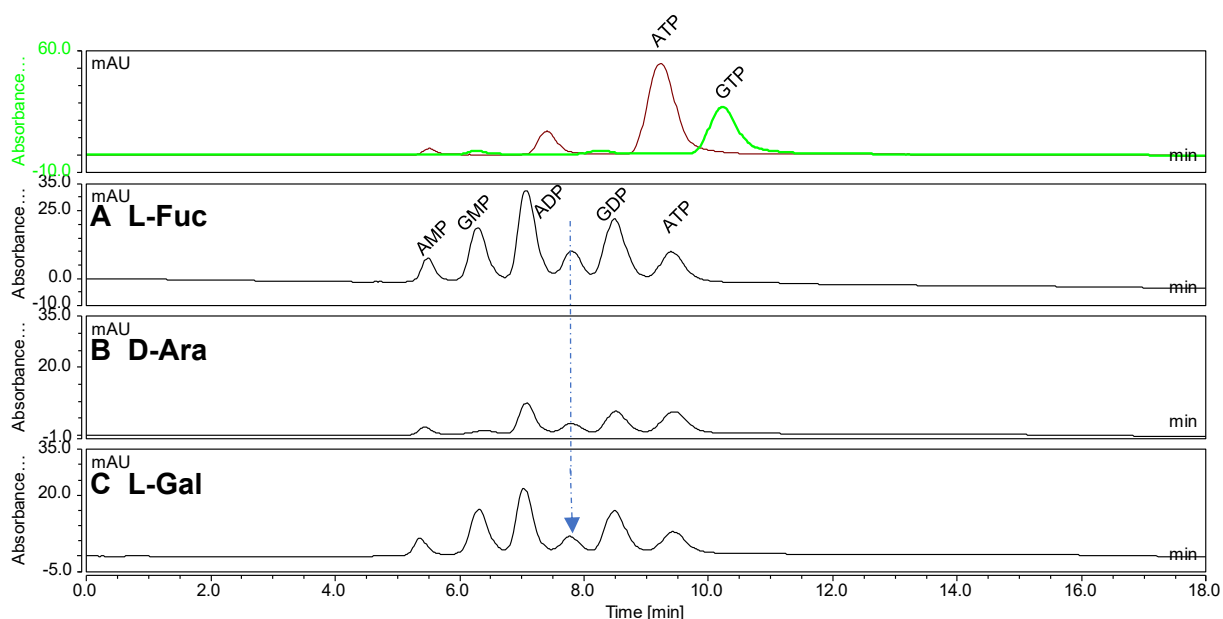
Figure 2.24: SAX chromatography of FKP reactions with L-Fuc, D-Ara and L-Gal, A-C, respectively. Reaction mixture contained 50 mM Tris-HCl (pH 7.5), 5 mM acceptor, 5 mM ATP, 5 mM GTP, 5 mM MnCl$_2$, inorganic pyrophosphatase (0.25 units) and 200 µL FKP (1 mg/ml) to give a total volume of 1 mL. Formation of the GDP-sugar nucleotide can be indicated by the red arrow and its retention is consistent across all reactions. Other reaction side products were identified by comparison with commercial standards. Top stacked traces are ATP (red) and GTP (green). From left to right: AMP, GMP, ADP, GDP-sugar nucleotide, GDP, ATP and traces of GTP.
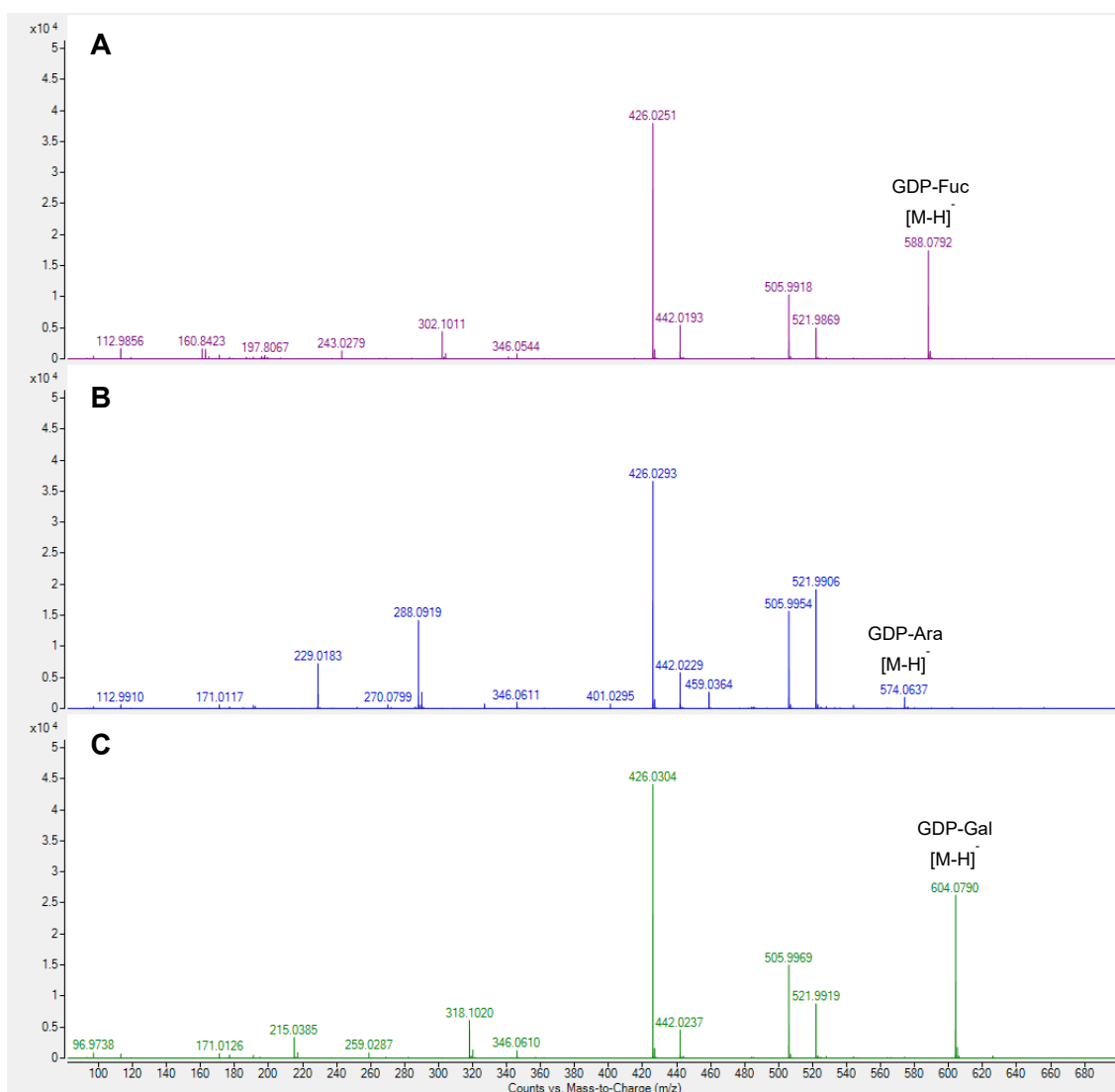
Figure 2.25: High resolution mass spectrometry spectra of FKP reaction products from (A) L-Fuc, (B) D-Ara and (C) L-Gal.

Having characterised the activity and promiscuity of FKP and synthesised of a panel of GDP-sugar nucleotide derivatives, including GDP-L-fucose, GDP-D-arabinose and GDP-L-galactose, we then continued to establish if HmFucT could utilise these new GDP-sugar nucleotides as donors. Figure 2.26 shows the formation of trisaccharides with a lactose core, following incubation of the crude FKP reaction with HmFucT and lactose. In all cases, it appears that the respective "fucosylated" trisaccharides product is formed, therefore, the HmFucT is able to use these non-natural GDP-sugar nucleotide donors for "fucosyl" transfer. For GDP-L-fucose, 33% of the lactose was catalysed to 2'FL. Likewise, for GDP-D-arabinose and GDP-L-galactose, the lactose acceptor was catalysed but to a lesser extent; 12% to 2'-arabinosyllactose and 9% to 2'-L-galactosyllactose, respectively. HR-MS confirmed the presence of said trisaccharides in the reaction mixture (see Figure 2.27). And, considering the ability of both FKP and HmFucT to accept numerous substrates, it

can be concluded, therefore, that it is highly likely that this two-step one pot approach could be extended to use with functionalised lactose acceptors as we have shown in section 2.2.2.2, for the synthesis of novel HMO analogues.
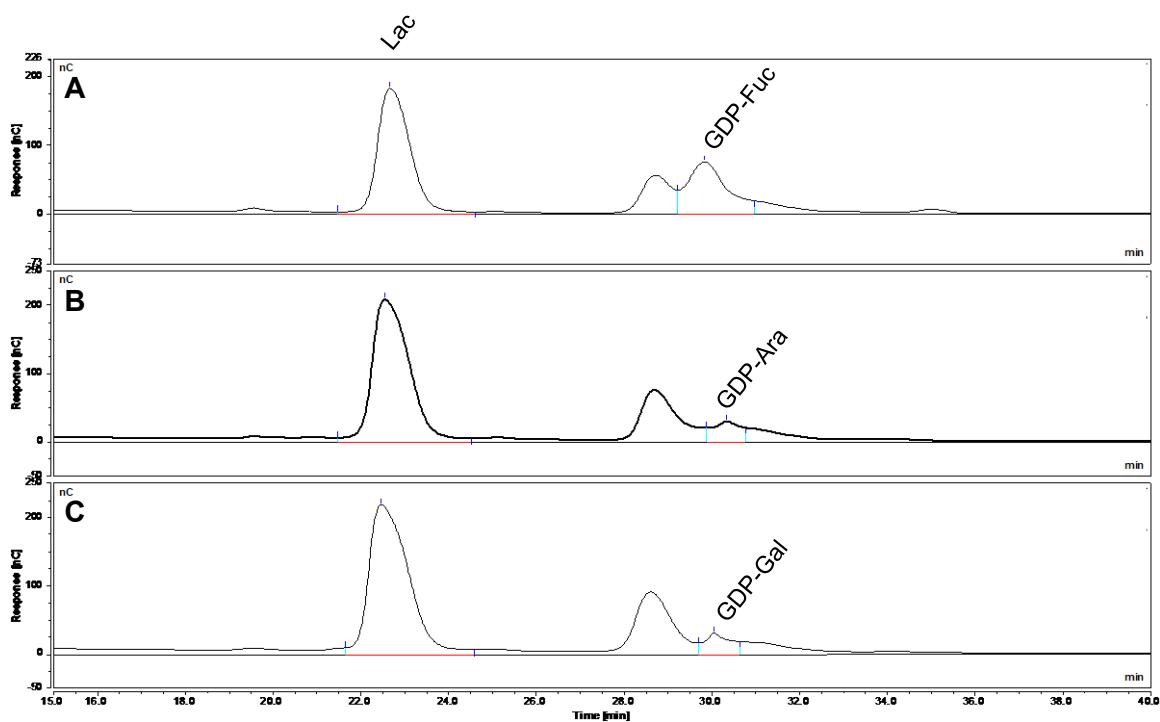
Figure 2.26: HPAEC-PAD chromatograms of two-step one pot reactions of FKP and α1,2-FucT for each of the GDP-sugar nucleotides; L-Fuc (A), D-Ara (B) and L-Gal (C) and lactose as an acceptor. Excess lactose acceptor can be seen at 22.74 min and potential trisaccharide products between at 29.79, 29.62 and 29.63 mins. The peak at 28.71 min is regarded as a reaction component due its presence in all samples.
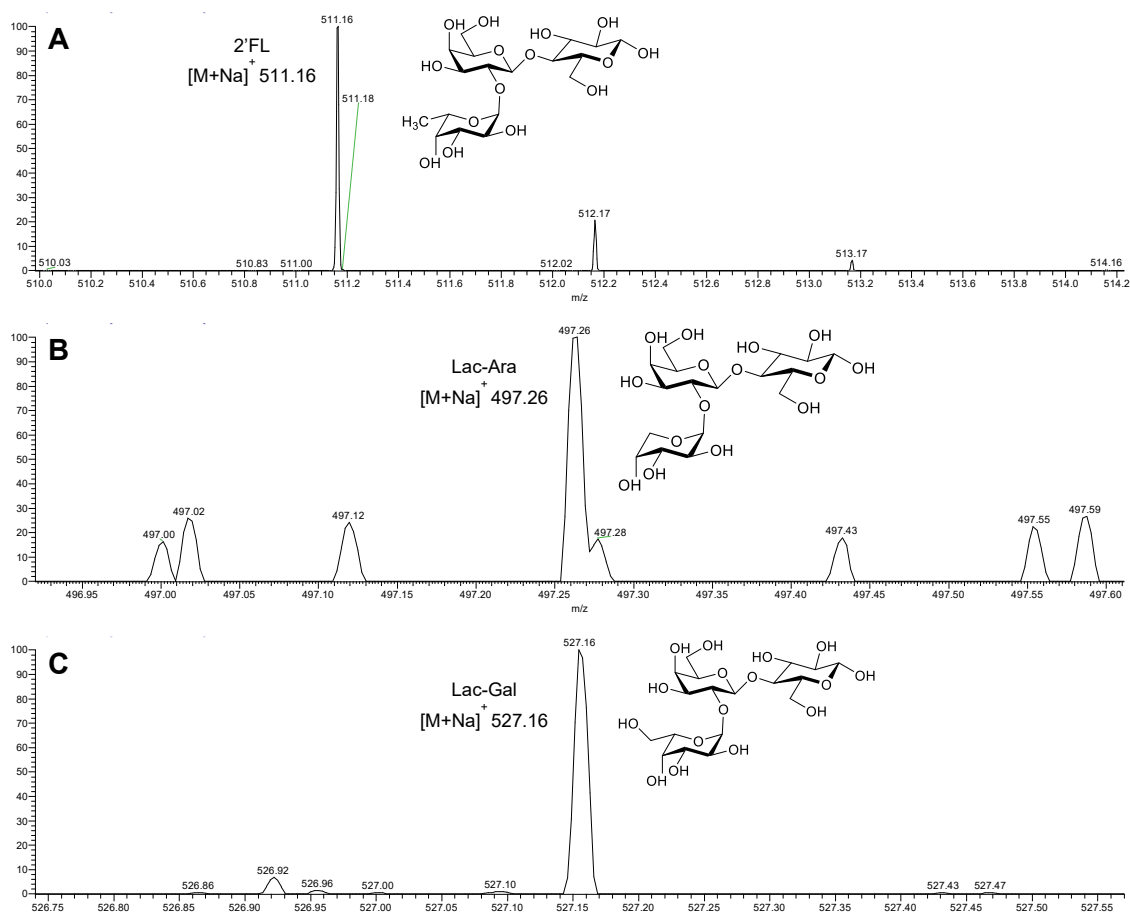
Figure 2.27: High resolution mass spectrometry spectra of trisaccharide products of two-step one pot reaction products with FKP and α1,2-FucT with lactose. (A) L-Fuc, (B) D-Ara and (C) L-Gal.

From past studies, it was expected that FKP was capable of using different fucosyl donor substrates as Wang *et al.* showed when they synthesised GDP-L-fucose analogues and consequently, their respective Lewis x derivatives in one pot reactions with α1,3-FucT. Briefly, seven monosaccharides were chosen with different unnatural substituents, each with varying stereoelectronics and hydrophobicity. In all cases, they achieved yields of greater than 80% using the same enzyme and conditions [146]. Although, we tried to replicate these results with different fucose surrogates, the obtained yields were significantly lower, including for the natural substrate despite thorough troubleshooting. Furthermore, one pot reactions for the synthesis of HMOs have exclusively been done with α1,3-FucT from *H. pylori*, with no evidence of α1,2-FucT unless as part of a whole-cell microbial biotransformation or cell factory approach towards 2'FL synthesis in *E. coli* or *Saccharomyces cerevisiae* [105] or a binding-elution strategy [180].

However, there are opportunities for improving yields of GDP-L-fucose derivatives by exploring the use of FKPs from different species, i.e., *Arabidopsis* [181]; carrying out site-

directed mutagenesis; or pursuing expression of FKP in different recombinant expression systems. This has previously been shown to be successful in *Saccharomyces cerevisiae* to synthesise nucleotide sugars from corresponding external monosaccharides in a mechanism similar to that of GDP-L-Fuc via the salvage pathway [182]. And further to the synthesis of HMOs, there are opportunities to explore the promiscuity of inter species FKPs for the synthesis of GDP-L-fucose derivatives as selective fucosyltransferase inhibitors [183]. Alternatively, there are further opportunities for mutagenesis of both FucT and FKP to increase efficiency of the enzymes and/or the production of specific products. Although, preferably structural information, i.e., cryo-EM or crystallographic structures, would be used to shape rationale for said mutations.

It should also be noted that for a successful scale up of the one-pot FucT/FKP system, a nucleotide triphosphate regeneration system would be needed, such as incorporation of bi-functional polyphosphate kinase, or phosphocreatine and three coupled kinase reactions [184-187]. Such systems also provide increased yields when used in multienzyme cascade reactions by shifting the reaction equilibrium toward the desired sugar-nucleotide donor and, therefore, consequent fucosylated products.

2.3 Conclusions

By optimising known techniques for the characterisation of glycans (HPAEC-PAD and HILIC HPLC with CAD) and using these techniques complementary chromatography techniques in combination with mass spectroscopic methods, the components of HMO mixtures were successfully identified, even if in the case of isomeric compounds. Furthermore, semi-preparative HILIC chromatography was an effective way of purifying said mixtures to a 10s mg scale, achieving purity greater than commercial methods as evidenced by the high contamination standards and the misidentification of HMOs present in mixtures from commercial sources. To enable benchtop analysis and purification of HMOs, a method for the use of HILIC cartridges was developed. Here, HMO mixtures were purified with adequate resolution at 10 mg loading capacity or used as a pre-purification step preceding HPLC steps. However, if columns with higher solid phase capacity were used, it is hypothesised that as much as 500 mg of HMOs could be purified in one bench top application.

Regarding the *in vitro* investigations with HmFucT, we can conclude that it is a promiscuous enzyme that can accept and convert a variety of di-, tri-, tetra- and pent-saccharides, including core HMO structures and fucosylated glycans. However, this fucosylation can occur at both the terminal non-reducing end but does so with greater than

60% efficiency in most cases. Therefore, this enzyme is capable of synthesising HMO analogues and this procedure can be scaled up when used in combination with FKP in a two-step, one pot reaction. Furthermore, FKP can convert D-Ara and L-Gal into their respective GDP-sugar nucleotides, and these can be processed as donors by HmFucT. This enables access to a range of novel non-natural HMO analogues with controllable chemistries at the C-5 position of the carbon ring. Moreover, the ability of HmFucT to catalyse the fucosylation of functionalised lactose derivatives could unlock the potential for using functionalised building blocks in a multi-enzyme cascade to chemoenzymatically synthesise novel HMO analogues with functionality for a multitude of uses, including glycan biosensors, arrays and selective inhibitors. It should be noted that this apparent enzymatic promiscuity may more accurately be explained by a lack of intracellular compartmentalisation that would be exhibited in *in vivo* systems.

2.4 References

11. Bode, L. (2012) Human milk oligosaccharides: Every baby needs a sugar mama, *Glycobiology.* **22**, 1147-1162.

97. O'Sullivan, A., Salcedo, J. & Rubert, J. (2018) Advanced analytical strategies for measuring free bioactive milk sugars: from composition and concentrations to human metabolic response, *Anal Bioanal Chem.* **410**, 3445-3462.

98. Porfirio, S., Archer-Hartmann, S., Moreau, G. B., Ramakrishnan, G., Haque, R., Kirkpatrick, B. D., Petri, W. A., Jr & Azadi, P. (2020) New strategies for profiling and characterization of human milk oligosaccharides, *Glycobiology*.

105. Yu, S., Liu, J. J., Yun, E. J., Kwak, S., Kim, K. H. & Jin, Y. S. (2018) Production of a human milk oligosaccharide 2'-fucosyllactose by metabolically engineered *Saccharomyces cerevisiae*, *Microbial Cell Factories.* **17**.

113. Zhao, C., Wu, Y. J., Yu, H., Shah, I. M., Li, Y. H., Zeng, J., Liu, B., Mills, D. A. & Chen, X. (2016) The one-pot multienzyme (OPME) synthesis of human blood group H antigens and a human milk oligosaccharide (HMOS) with highly active *Thermosynechococcus elongatus* alpha 1-2-fucosyltransferase, *Chemical Communications.* **52**, 3899-3902.

132. Stein, D. B., Lin, Y. N. & Lin, C. H. (2008) Characterization of *Helicobacter pylori* alpha 1,2-fucosyltransferase for enzymatic synthesis of tumor-associated antigens, *Advanced Synthesis & Catalysis.* **350**, 2313-2321.

144. Rabuka, D., Hubbard, S. C., Laughlin, S. T., Argade, S. P. & Bertozzi, C. R. (2006) A chemical reporter strategy to probe glycoprotein fucosylation, *J Am Chem Soc.* **128**, 12078-9.

145. Sawa, M., Hsu, T. L., Itoh, T., Sugiyama, M., Hanson, S. R., Vogt, P. K. & Wong, C. H. (2006) Glycoproteomic probes for fluorescent imaging of fucosylated glycans in vivo, *Proc Natl Acad Sci U S A.* **103**, 12371-6.

146. Wang, W., Hu, T. S., Frantom, P. A., Zheng, T. Q., Gerwe, B., del Amo, D. S., Garret, S., Seidel, R. D. & Wu, P. (2009) Chemoenzymatic synthesis of GDP-L-fucose and the Lewis X glycan derivatives, *Proceedings of the National Academy of Sciences of the United States of America.* **106**, 16096-16101.

158. Garber, J. M., Hennet, T. & Szymanski, C. M. (2021) Significance of fucose in intestinal health and disease, *Mol Microbiol.* **115**, 1086-1093.

164. Wan, L., Zhu, Y., Zhang, W. & Mu, W. (2020) α-l-Fucosidases and their applications for the production of fucosylated human milk oligosaccharides, *Appl Microbiol Biotechnol.* **104**, 5619-5631.

165. Le Huërou-Luron, I., Blat, S. & Boudry, G. (2010) Breast- vs. formula-feeding: impacts on the digestive tract and immediate and long-term health effects, *Nutrition research reviews.* **23**, 23-36.

166. Ninonuevo, M. R., Park, Y., Yin, H. F., Zhang, J. H., Ward, R. E., Clowers, B. H., German, J. B., Freeman, S. L., Killeen, K., Grimm, R. & Lebrilla, C. B. (2006) A strategy for annotating the human milk glycome, *J Agric Food Chem.* **54**, 7471-7480.

167. Auer, F., Jarvas, G. & Guttman, A. (2021) Recent advances in the analysis of human milk oligosaccharides by liquid phase separation methods, *Journal of Chromatography B.* **1162**, 122497.

168. van Leeuwen, S. S. (2019) Challenges and pitfalls in human milk oligosaccharide analysis, *Nutrients.* **11**, 2684.

169. Ruhaak, L. R. & Lebrilla, C. B. (2012) Advances in analysis of human milk oligosaccharides, *Advances in nutrition (Bethesda, Md).* **3**, 406s-14s.

170. Austin, S., Cuany, D., Michaud, J., Diehl, B. & Casado, B. (2018) Determination of 2′-fucosyllactose and lacto-*N*-neotetraose in infant formula, *Molecules.* **23**, 2650.

171. Balogh, R., Jankovics, P. & Beni, S. (2015) Qualitative and quantitative analysis of N-acetyllactosamine and lacto-N-biose, the two major building blocks of human milk oligosaccharides in human milk samples by high-performance liquid chromatography-tandem mass spectrometry using a porous graphitic carbon column, *J Chromatogr.* **1422**, 140-146.

172. Ghosh, R. & Kline, P. (2019) HPLC with charged aerosol detector (CAD) as a quality control platform for analysis of carbohydrate polymers, *BMC Research Notes.* **12**, 268.

173. Coppa, G. V., Gabrielli, O., Zampini, L., Galeazzi, T., Ficcadenti, A., Padella, L., Santoro, L., Soldi, S., Carlucci, A., Bertino, E. & Morelli, L. (2011) Oligosaccharides in 4 different milk groups, *Bifidobacteria*, and *Ruminococcus obeum*, *Journal of Pediatric Gastroenterology and Nutrition.* **53**.

174. Noguchi, M., Tanaka, T., Gyakushi, H., Kobayashi, A. & Shoda, S. (2009) Efficient synthesis of sugar oxazolines from unprotected N-acetyl-2-amino sugars by using chloroformamidinium reagent in water, *J Org Chem.* **74**, 2210-2.

175. Shen, X., Yan, H., Wang, C., Gao, P., Johnson, C. H. & Snyder, M. P. (2022) TidyMass an object-oriented reproducible analysis framework for LC–MS data, *Nature communications.* **13**, 4365.

176. Kellman, B. P., Richelle, A., Yang, J.-Y., Chapla, D., Chiang, A. W. T., Najera, J. A., Liang, C., Fürst, A., Bao, B., Koga, N., Mohammad, M. A., Bruntse, A. B., Haymond, M. W., Moremen, K. W., Bode, L. & Lewis, N. E. (2022) Elucidating human milk oligosaccharide biosynthetic genes through network-based multi-omics integration, *Nature communications.* **13**, 2455.

177. Yu, H., Lau, K., Li, Y., Sugiarto, G. & Chen, X. (2012) One-pot multienzyme synthesis of Lewis x and sialyl Lewis x antigens, *Current protocols in chemical biology.* **4**, 233-247.

178. Bisso, A., Sturla, L., Zanardi, D., De Flora, A. & Tonetti, M. (1999) Structural and enzymatic characterization of human recombinant GDP-D-mannose-4,6-dehydratase, *FEBS Lett.* **456**, 370-4.

179. Sturla, L., Bisso, A., Zanardi, D., Benatti, U., De Flora, A. & Tonetti, M. (1997) Expression, purification and characterization of GDP-D-mannose 4,6-dehydratase from *Escherichia coli*, *FEBS Lett.* **412**, 126-30.

180. Zhu, H., Wu, Z., Gadi, M. R., Wang, S., Guo, Y., Edmunds, G., Guan, W. & Fang, J. (2017) Cation exchange assisted binding-elution strategy for enzymatic synthesis of human milk oligosaccharides (HMOs), *Bioorganic & Medicinal Chemistry Letters.* **27**, 4285-4287.

181. Kotake, T., Hojo, S., Tajima, N., Matsuoka, K., Koyama, T. & Tsumuraya, Y. (2008) A bifunctional enzyme with L-fucokinase and GDP-L-fucose pyrophosphorylase activities salvages free L-fucose in *Arabidopsis*, *J Biol Chem.* **283**, 8125-35.

182. Liu, T.-W., Ito, H., Chiba, Y., Kubota, T., Sato, T. & Narimatsu, H. (2011) Functional expression of l-fucokinase/guanosine 5′-diphosphate-l-fucose pyrophosphorylase from *Bacteroides fragilis* in *Saccharomyces cerevisiae* for the production of nucleotide sugars from exogenous monosaccharides, *Glycobiology.* **21**, 1228-1236.

183. Lin, Y. N., Stein, D., Lin, S. W., Chang, S. M., Lin, T. C., Chuang, Y. R., Gervay-Hague, J., Narimatsu, H. & Lin, C. H. (2012) Chemoenzymatic synthesis of GDP-L-fucose derivatives as potent and selective alpha-1,3-fucosyltransferase inhibitors, *Advanced Synthesis & Catalysis.* **354**, 1750-1758.

184. Kameda, A., Shiba, T., Kawazoe, Y., Satoh, Y., Ihara, Y., Munekata, M., Ishige, K. & Noguchi, T. (2001) A novel ATP regeneration system using polyphosphate-AMP phosphotransferase and polyphosphate kinase, *J Biosci Bioeng.* **91**, 557-563.

185. Fehlau, M., Kaspar, F., Hellendahl, K. F., Schollmeyer, J., Neubauer, P. & Wagner, A. (2020) Modular enzymatic cascade synthesis of nucleotides using a (d)ATP regeneration system, *Frontiers in Bioengineering and Biotechnology.* **8**.

186. Tsai, T.-I., Lee, H.-Y., Chang, S.-H., Wang, C.-H., Tu, Y.-C., Lin, Y.-C., Hwang, D.-R., Wu, C.-Y. & Wong, C.-H. (2013) Effective sugar nucleotide regeneration for the large-scale enzymatic synthesis of Globo H and SSEA4, *J Am Chem Soc.* **135**, 14831-14839.

187. Wang, P.-H., Fujishima, K., Berhanu, S., Kuruma, Y., Jia, T. Z., Khusnutdinova, A. N., Yakunin, A. F. & McGlynn, S. E. (2020) A bifunctional polyphosphate kinase driving the regeneration of nucleoside triphosphate and reconstituted cell-free protein synthesis, *ACS Synthetic Biology.* **9**, 36-42.

Chapter 3

Chapter 3 – Probing Bacterial-Glycan Interactions

3.1 General Introduction

Although glycan arrays have been used extensively to monitor protein-glycan interactions, they have been utilised less successfully to identify specific microbe-glycan interactions using live bacteria. Historically, it has only been achieved in a handful of studies where Escherichia *coli* was predominantly used to model mannose-bacteria binding interactions [188-190].

This chapter aims to examine the interactions of microbes with glycans regarding the ability of human milk oligosaccharides to modulate bacterial growth when supplemented using *in vitro* cultures. To investigate the unique binding specificities of different bacterial strains, ELISA-like glycan immobilisation assays will be developed to probe the effect of glycan valency, linker flexibility and chain length on the selectivity of bacterial binding. Further flow cytometry studies will be used to understand the aggregation properties of high binding glycans.

3.2 Part I: Microbial HMO Utilisation

3.2.1 Introduction

The gut microbiota has been established as an important player influencing many aspects of human physiology. HMOs are known to act as prebiotics and most studies on HMO utilisation by microbes focus on the interaction of naturally occurring mixtures of HMOs with gut bacteria. However, more than 200 HMO structures are present in human milk [10, 191]. There are many species of commensal bacteria that are known to utilise HMOs to exert a prebiotic effect. As such, HMOs are of intense interest, not only to the infant formula industry, but also for use in adult health food products where they can be used to restore a beneficial microbiota in dysbiotic adults and provide health benefits. *In vivo* studies show that 2'FL and LNnT are specific modulators of the adult microbiota with a very specific increase in bifidobacteria, where a mix of 2'FL and LNnT was better tolerated by bifidobacteria than the individual HMOs [192].

This study aims to establish the capabilities of wild-type non-pathogenic *Escherichia coli* K12 MG1655 as a model organism to study HMO utilisation of fucosylated HMOs for pathogens of interest. *E. coli* was chosen as a model organism for this study due to its well-characterised catabolic pathways. Work focuses on whether *Escherichia coli* K12 MG1655 can utilise L-fucose and specified fucosylated HMOs as a carbon source to facilitate growth and evaluates the use of *E. coli* as a model organism for similar pathogens of interest. L-fucose, a known chemotactic agent for *E. coli* [157], 2'FL or a combination of both, were added to minimal culture media as the sole energy source, and the growth *in vitro* by *E. coli* K12 was measured over time.

3.2.1.1 *Escherichia coli* Catabolism

Escherichia coli has been a key model organism from the very earliest work on molecular genetics and continues to play an important role to this day. Much of our understanding of the fundamental concepts of molecular biology such as replication, gene expression and protein synthesis have all been achieved through studies of *E. coli.* As such, L-fucose metabolism has been extensively characterised in *E. coli. E. coli* utilises L-fucose and L-rhamnose as carbon and energy sources through parallel pathways. The L-fucose pathway is encoded by the *fuc* operon [193-195]. The *fuc* operon is comprised of *fucP* (L-fucose permease or L-fucose/H$^+$ symporter), *fucI* (L-fucose isomerase), *fucK* (L-fuculose kinase), *fucA* (L-fuculose 1-phosphate aldolase), *fucO* (L-1,2-propanediol oxidoreductase) (see Figure 3.1A) [193]. It should be noted that the L-fucose/H$^+$ symporter is also able to transport L-galactose and D-arabinose but at reduced rates compared with L-fucose, and not the similar monosaccharides, L-rhamnose and L-arabinose [196]. The positive regulator encoded by *fucR* responds to L-fuculose-1-phosphate as the effector [197]. All of the *fuc* genes, including *fucO*, are induced both aerobically and anaerobically and isomerase, kinase, and aldolase were found to be functional in both aerobic and anaerobic environments, which means that the formation of neither DHAP nor L-lactaldehyde was influenced by the environment [198]. Aerobically, L-lactaldehyde is converted by a partially inducible NAD-linked dehydrogenase to L-lactate, which is oxidized to pyruvate for further metabolism [199, 200]. Anaerobically, however, L-lactaldehyde is reduced to L-1,2-propanediol by propanediol oxidoreductase in a typical fermentative mechanism in which NADH is oxidized to NAD and L-1,2-propanediol is excreted as a fermentation product (see Figure 3.1B) [195, 200].

Figure 3.1: Summary of *fuc* operon and L-fucose metabolism pathways in *Escherichia coli*. (**A**) A schematic representation of the different gene clusters of the *fuc* operon. (**B**) Metabolic pathways for the metabolism of L-fucose, including the L-fucose degradation pathway (green), the dihydroxyacetone phosphate degradation pathway (red), the 1,2-propanediol synthesis pathway/anaerobic fermentation of L-lactaldehyde (yellow) and the aerobic conversion of L-lactaldehyde to pyruvate (blue).

3.2.2 Results and Discussion

3.2.2.1 Effect of HMOs on *E. coli* Growth

First, *E. coli* K12 was grown aerobically in M9 minimal media supplemented with either 2% (w/v) glucose or fucose (L-Fuc) over 16 h (see Figure 3.2). This strain grew reasonably well on glucose, but not on L-Fuc (final OD_{600nm} of 0.342 ± 0.002 and 0.093 ± 0.009 of glucose and fucose, respectively). However, when the *E. coli* K12 was grown aerobically in M9 minimal media supplemented with 0 to 10% (w/v) L-Fuc over 36 h, the culture displayed a prolonged lag-phase (10 h) from which growth proceeded normally over the following hours to reach the stationary phase after 20 h (see Figure 3.3) and a lower final cell density. At 1.25%, 2.5% and 5% L-Fuc (w/v), the culture growth is very consistent (final OD_{600nm} of 0.567 ± 0.03, 0.52 ± 0.07, 0.547 ± 0.05, for 1.25%, 2.5% and 5% (w/v) L-Fuc, respectively) however, at 10% L-Fuc no growth is detected indicating a possible inhibitory effect at high concentrations of L-Fuc (final OD_{600nm} of -0.193 ± 0.004).

These findings are consistent with previous research, which demonstrated that L-fucose can act as the sole carbon and energy source for wild-type *E. coli* K-12, while glucose remains the preferred substrate [201]. Additionally, the prolonged lag-phase may in part, be explained by the inducible nature of the fucose catalysis pathway, where *fucR* acts as the transcriptional activator of the *fuc* operon. Recent research has shown that the metabolic activity of *E. coli* responds differently when grown on glucose or fucose. *E. coli* grown on fucose exhibited inefficient carbon metabolism, which resulted in L-fuculose accumulation and slow growth. This inefficient carbon metabolism was postulated to have been caused by ATP deficiency arising from upregulation of gluconeogenesis and quorum sensing and downregulation of glycolysis, following analysis of metabolomic and transcriptomic data [202]. Additionally, due to the ATP deficiency, the metabolic activity of *fucK,* which phosphorylates L-fuculose to L-fuculose-1-phosphate using ATP, may be insufficient to phosphorylate all available L-fuculose, causing accumulation of L-fuculose and inefficient carbon flux. This data may explain why the supplementation of medium with high concentrations (> 5%, w/v) of L-Fuc seen in Figure 3.3 appears to have an inhibitory effect to on cell growth.

Co-feeding experiments were carried out where *E. coli* K12 was cultured aerobically in minimal medium supplemented with 2% (w/v) L-Fuc and increasing concentrations of 2'FL (0 – 5%, w/v) over 48 h. The data, as seen in Figure 3.4, confirms that 2'FL cannot be used as the sole energy source by *E. coli* K12. It follows that *E. coli* K12 in minimal

supplemented with L-Fuc and increasing concentrations of 2'FL does not affect cell growth in treatment groups supplemented with 0%, 1.25% and 2.5% (w/v) L-Fuc. However, when the culture is supplemented with 5% L-Fuc (w/v), concentrations of 2'FL above 0.5% (w/v) appear to increase growth by half. Negative values are presumed to be due to condensation problems. This could potentially be rectified by using deep-well plates to improve aeration, achieve more efficient mixing with minimal evaporation. To confirm these findings, it might be beneficial to repeat the experiment and include treatment groups with 2'FL supplementation above 5% (w/v).
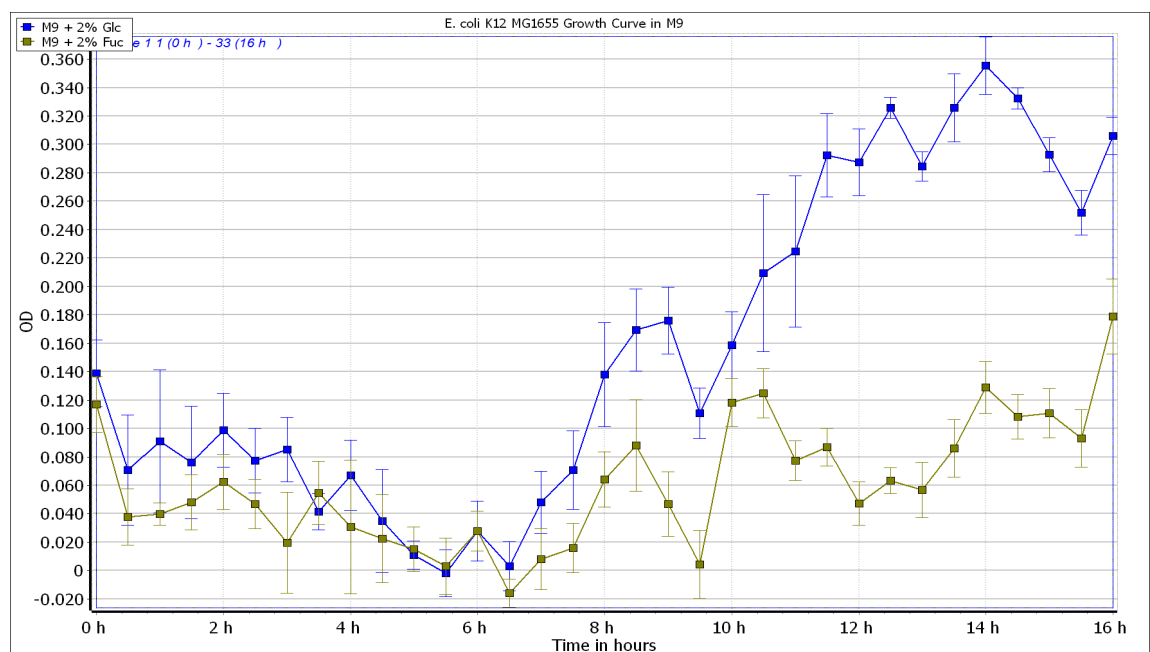
Figure 3.2: Growth of *E. coli* K12 on different carbon sources in aerobic cultures. The M9 minimal medium was supplemented with 2% glucose (w/v) (blue) or 2 % L-Fuc (w/v) (green). Optical density measurements were taken from 0 h onwards every 30 min using spiral average (2 mm) for 16 h. Data points are shown as averages of three replicates with error bars representing the range of OD values for each time point.
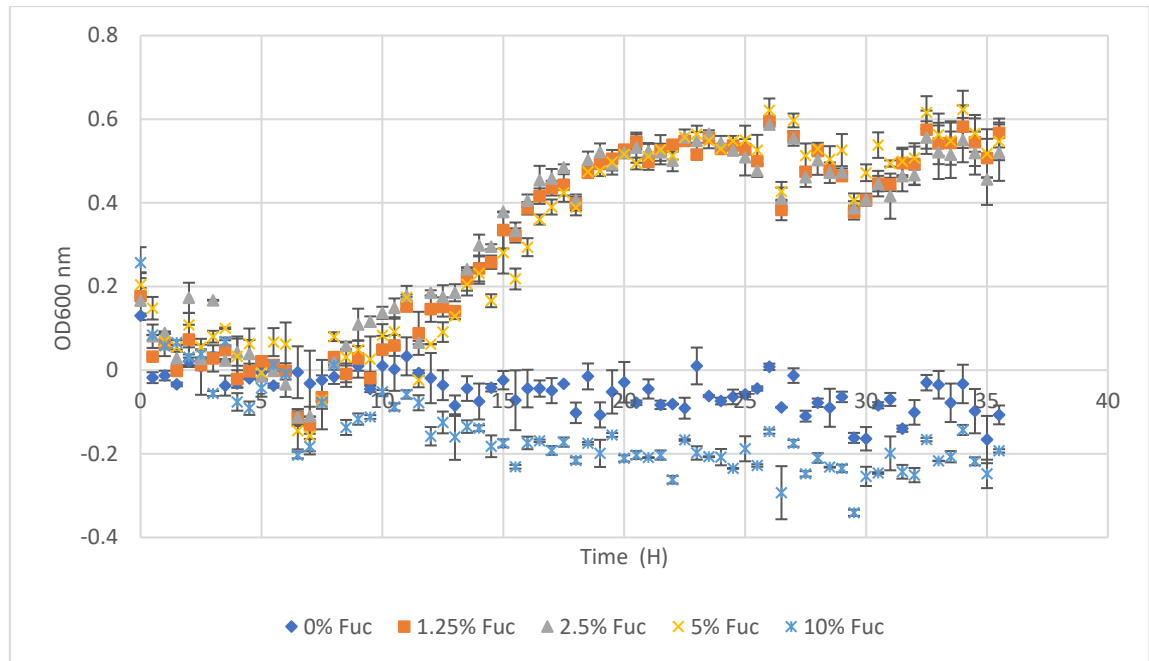
Figure 3.3: Growth of *E. coli* K12 on increasing concentrations of L-fucose in aerobic cultures. The M9 minimal medium was supplemented with 0% (blue diamonds), 1.25% (orange squares), 2.5% (grey triangles), 5% (yellow crosses) and 10% L-Fuc (w/v) (light blue asterisks). Above 1.25% (w/v) L-Fuc supplementation, no different can be seen in microbial growth. However, 10% (w/v) L-Fuc appears to inhibit cell growth. Optical density measurements were taken from 0 h onwards every 30 min using spiral average (2 mm) for 36 h. Data points are shown as averages across three replicates with error bars representing the standard deviation for data points.

The inability of 2'FL to effect *E. coli* growth is likely due to *FucP,* a L-fucose/H+ symporter, which mediates the uptake of L-Fuc across the cell membrane with the concomitant transport of protons into the cell. Unsurprisingly, the L-fucose/H+ symporter is not able to transport 2'FL across the membrane but this data also indicates that 2'FL is not able to inhibit growth by non-specific binding/blockage of the symporter. Typically, L-Fuc utilisation pathways often include fucosidases to cleave L-Fuc from the oligosaccharide backbone, for example in species from the phyla *Actinobacteria, Bacteroidetes, Firmicutes*, and *Proteobacteria* [158, 203]. Some commensal species secrete fucosidases, e.g., *Bacteroides fragilis,* which aid the growth and invasion of pathogens by releasing L-Fuc from intestinal mucins [204]. It is perhaps unsurprising that *E. coli* grown in monoculture, which lack an endogenous fucosidase, are unable to grow on fucosylated HMOs. Notably, though, the presence of genes for sugar degradation alone does not guarantee the ability to grow in the presence of HMOs. For instance, *Lactobacillus casei* BL23 is equipped with an α-fucosidase which can act on 2',FL but is unable to use it as an energy source [205]. The internal localization of the enzyme and lack of a trans-membrane transporter for 2'FL is thought to cause this lack of 2'FL metabolism [40].
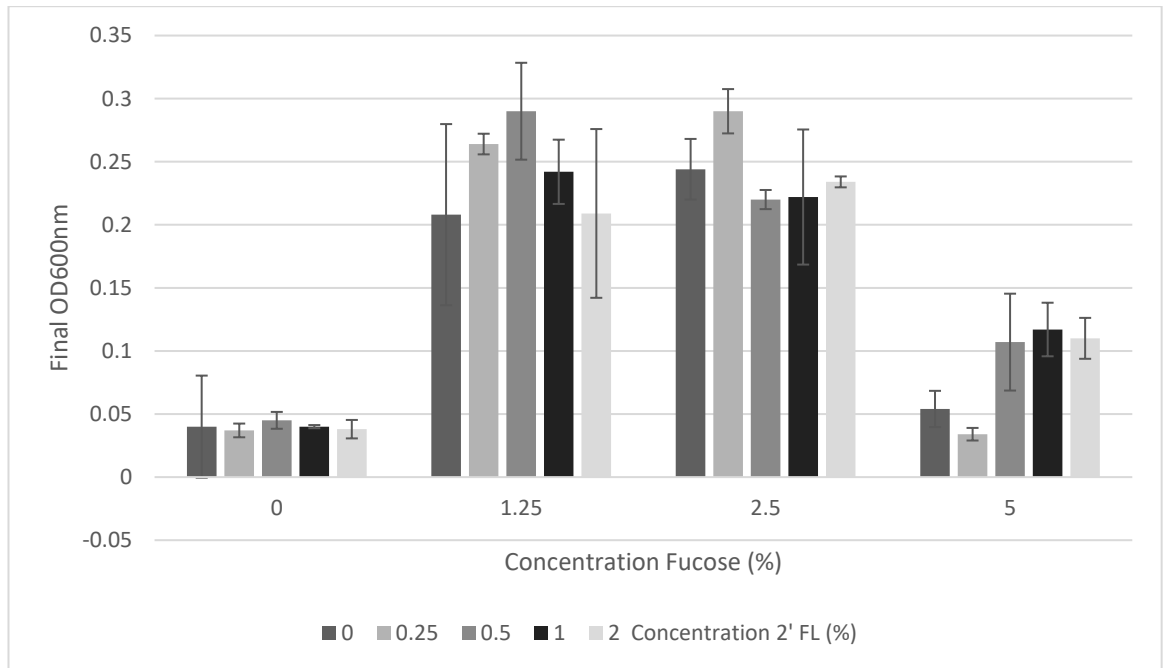
Figure 3.4: Final OD$_{600nm}$ of E. coli K-12 cultures supplemented with 0%, 1.25%, 2.5% and 5% (w/v) L-Fuc and co-fed with increasing concentrations of 2'FL. Optical density measurements were taken from 0 h onwards every 30 min using spiral average (2 mm) for 36 h. Data points are final optical density measurements shown as averages across three replicates with error bars representing the standard deviation for data points.

3.2.3 Conclusions

This section has examined the interactions of HMOs on the growth of *E. coli*. It was found that *E. coli* is able to utilise L-Fuc as a sole energy source. Interestingly, it was also found that high concentrations of L-Fuc (> 5%, w/v) inhibit *E. coli* growth which is postulated to be due inefficient carbon metabolism resulting in the accumulation of L-fuculose from ATP deficiency from upregulation of gluconeogenesis and quorum sensing and downregulation of glycolysis. However, HMOs cannot be metabolised as a substitute to L-Fuc, nor do they effect growth when co-fed with L-Fuc unlike with commensal bacteria. This is hypothesised to be due to the lack of fucosidase expression by *E. coli* and the inability of the *FucP* H$^+$/L-fucose transporter to transport 2'FL. To conclude, *E. coli* can be used as a reliable organism for modelling the metabolism of HMOs, but only for microbes that also possess the *fuc* operon and lack fucosidase expression, such as *Campylobacter jejuni,* unless co-cultured with microbes that have endogenous fucosidase expression allowing the scavenging of released L-Fuc.

3.3 Part II: *E. coli* Binding Assay

3.3.1 Introduction

Carbohydrates are the dominant biopolymer on earth and play important roles ranging from building material for plants to function in many biological systems [206]. Because glycan motifs have a strong and well-characterised interaction, a relatively high surface coverage and a known functioning role used by multiple different pathogenic species, they serve as a promising target for detection of infection and targeted drug therapies. As such, this section aims to develop a basic assay format, like the principles of an enzyme-linked immunosorbent assay (ELISA), but capable of giving specific binding interactions of glycans with live bacteria to understand the specific glycan-bacteria binding interactions that occur with *Escherichia coli* as a model. The experimental workflow can be seen in Figure 3.5.
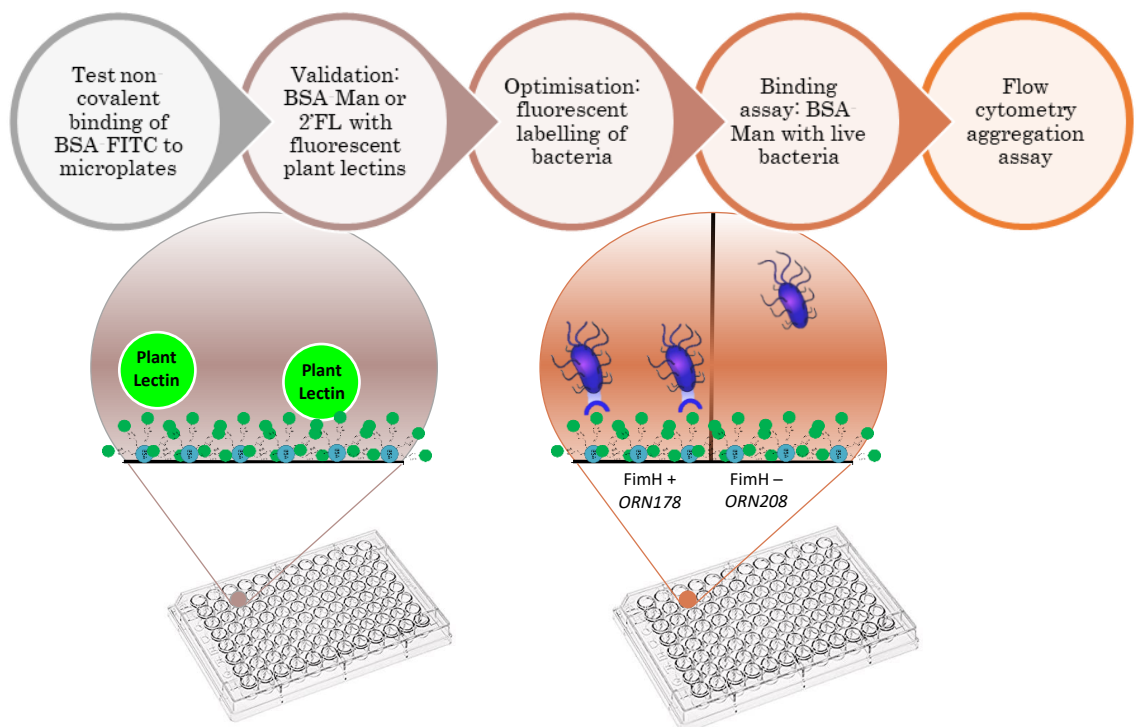
Figure 3.5: Schematic of workflow for development of bacterial binding assay. BSA-mannose conjugates are non-covalently bound to the microplate surface and the assay was validated with fluorescent plant lectins, followed by proof of concept with mannose-binding (ORN178, FimH+) and non-mannose-binding (ORN208, FimH-) strains of *E. coli*.

Glycans containing chains of mannose with different degrees of polymerisation (DP) will be immobilised to a microplate surface via a protein linker. The highly valent polysaccharide, mannan from *Saccharomyces cerevisiae* – a highly mannosylated glycoprotein displayed on the cell wall will also be immobilised to evaluate differences in binding affinity between mono-, di- and trisaccharides compared to polysaccharides. Figure 3.6A shows yeast mannan as a densely branched α-linked mannose polymer with a molecular weight ranging from 20,000 to 200,000 Da [207]. It includes a linear α-1,6-mannoside backbone branched with mono-, di-, tri- and tetramer α-1,2- or α-1,3-mannosides [208, 209].

Fluorescent intensity of bound proteins or cells will be used to indicate binding. To demonstrate the specificity of binding reactions facilitated by the assay format, fluorescent plant lectins, will be used to validate the assay, followed by fluorescently labelled *Escherichia coli* strains, ORN178 (wild-type) and ORN208 (mutant). ORN178 is a positive phenotype for the type-1 pilli expression (FimH+), and the ORN208 that does not express type-1 pilli, which were used as a negative control (FimH-) [210]. Figure 3.6B shows the structure of type-1 pilli in *E. coli*, which is assembled by the Fim system and comprises of a FimC chaperone, a FimD usher, a flexible FimA region consisting of many subunits, and a rigid tip region of FimF, FimG and capped with a FimH lectin-binding domain – a mannose specific lectin [211].
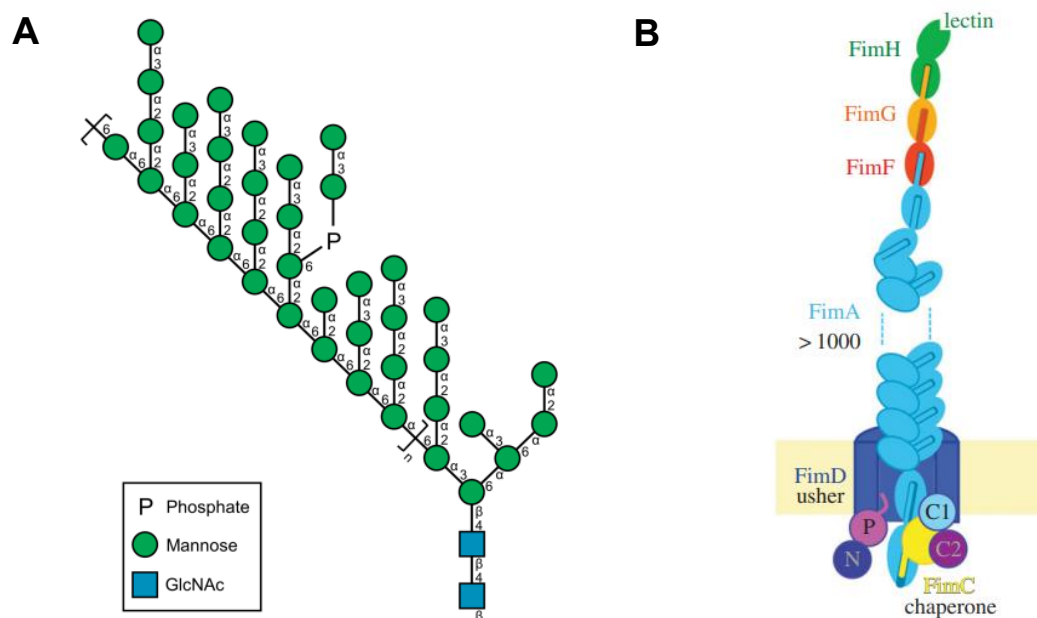
Figure 3.6: The structure of mannan from *Saccharomyces cerevisiae* (A) and an *Escherichia coli* type-1 pillus (B). (A) Mannan includes a linear α1, 6-mannoside backbone branched with α1, 2-mannoside and α1, 3-mannoside bonds in the form of mono-, di-, tri-, and tetramers. **(B)** The type-1 pilus from *Escherichia coli* is assembled by the Fim system and comprises of a FimC chaperone, a FimD usher (transmembrane pore), a flexible FimA region consisting of many subunits, and a rigid tip region of FimF, FimG and capped with a FimH lectin-binding domain. Adapted with permission from [211, 212].

3.3.1.1 *E. coli*-Glycan Binding Studies

Only a handful of prior studies have shown specific and selective binding of live bacteria to a glycan arrays or glycan-functionalised polymers or nanomaterials and these were mainly published by the Seeberger and Field groups over a decade ago. In 2004, Seeberger reported a whole cell application of the glycan microarray [188]. In this work, first five different monosaccharides equipped with an ethanolamine linker at their reducing ends were used to construct microarrays on glass slides coated with amine-reactive homobifunctional disuccinimidyl carbonate linker. *E. coli* cells (ORN178) that had been stained with nucleic acid staining dye were bound to the slide at only mannose residues as a proof of concept. The *E. coli* only bound to the mannose residues and not to other carbohydrates. Two *E. coli* strains were then tested (ORN178 and ORN208) and the mutant strain exhibited a reduced affinity to mannose as expected. Minimal binding of ORN208 to mannose was still detected, which one can assume is due to non-specific binding by other receptors on the cell surface. The format was then used as a biosensor platform to determine the carbohydrate binding "fingerprint" of the bacteria present in a complex mixture, i.e., sheep erythrocytes and serum. The results can been in Figure 3.7. Before this, only intact mammalian cells had been seen to demonstrate carbohydrate-specific binding in a microarray format [213].
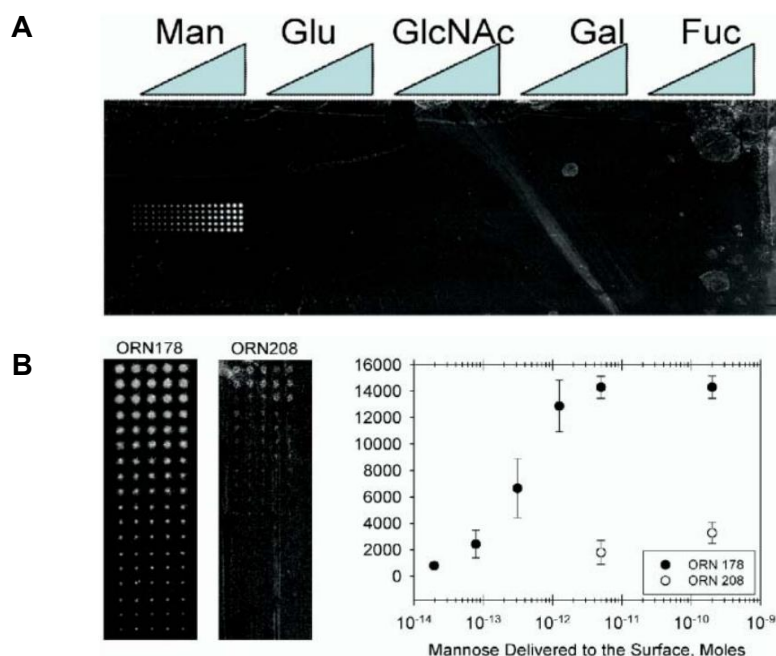
Figure 3.7: Summary of data from Disney *et al.* (A) An image of a carbohydrate array containing monosaccharides with a reducing end ethanolamine linker after incubation with ORN178 that were stained with SYTO 83 cell-permeable nucleic acid staining dye. **(B)** Left: adhesion of with mannose binding *E. coli*, ORN178, or mutant *E. coli*, ORN208 to mannose microarrays where ORN208 has greatly diminished mannose binding affinity. Right: plots of the experimental data from these two slides. Reproduced with permission from [188].

In the same year, the Seeberger group also published on the detection of bacteria with carbohydrate-functionalised fluorescent polymers [214] the results of which can be seen in Figure 3.8A. Carbohydrate-functionalised poly(p-phenylene ethynylene) (PPE) was used for the detection of *E. coli* by multivalent interactions. The mannose-functionalised PPE (**2a**) was able to fluorescently stain E. coli ORN178, while the mutant ORN208 strain gave no polymer fluorescence after incubation. Additionally, following incubation, neither the non-functionalised polymer (1) nor the galactose-functionalised polymer displayed fluorescent staining with either *E. coli* strain. A few years later, the Field group published a similar methodology but employing carbohydrate-functionalise quantum dots for bacterial detection [189]. Mannose-coated cadmium sulphide (Cds) quantum dots (Man-QDs) were able to induce luminescent aggregates of *E. coli* which were, therefore, able to detect bacteria in cell suspensions containing as few as 10^4 cells per mL. The recognition and subsequent detection of *E. coli* using the Man-QD was also shown to the specific as aggregation did not occur with FimH defective E. coli ORN208, or with galactose-coated QDs (see Figure 3.8B).

Techniques other than glycan arrays and glycopolymer aggregation have been used for the detection of bacteria. The most recent instance of supramolecular recognition of *E. coli* was by Arias *et al.,* in 2017 [215], who used fluorescent oligo(phenyleneethynylene)s with mannopyranoside termini groups coupled with surface plasmon resonance (SPR) to specifically recognise ORN178, and not ORN208, to sensitivity of 10^4 colony forming units (CFU)/mL at flow rate of 50 µL/min. Before then, a technique for the discrimination of *E. coli* strains using glycan cantilever array sensors was published [216]. The authors claimed that the method exhibited specific and reproducible deflection with a sensitivity range over four orders of magnitude. The procedure was tested with ORN178 and galactose, trimannose and nonamannose where galactose showed minimal deflection. The binding properties of *E. coli* strains with distinct mannose capabilities was also tested; ORN178, ORN208 and ORN206 – a strain that does not produce pili. However, the largest signal was seen for ORN206, which the authors explained as the missing pili facilitated access to increased non-specific binding of mannose transporters.
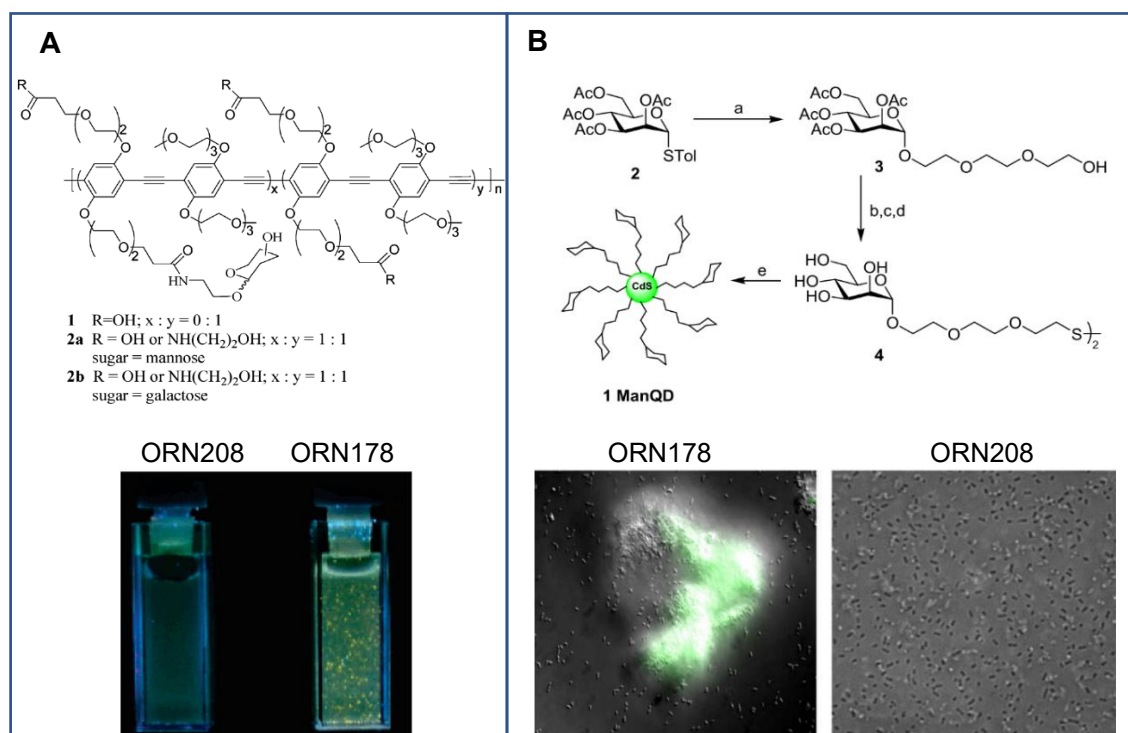
Figure 3.8: Schematic of the mannose probes and resultant observed E. coli aggregation following incubation from the literature. (**A**) Top: structure of the mannosylated polymer synthesised by Disney *et al.*, and bottom: visualization of ORN208 (mutant), left, and ORN178 (mannose-binding), right, E. coli strains after incubation with mannosylated polymer 2a. (**B**) Top: schematic showing the synthesis of synthesis of Man-QD by Mukhopadhyay *et al.*, and bottom: confocal fluorescence microscopy images of bacterial aggregates due to Man-QD-mediated aggregation of ORN178, left, and ORN208, right, showing no aggregation. Reproduced with permission from Copyright 2023 American Chemical Society and Creative Commons Public Licenses [189, 214].

3.3.1.2 Glycan Microplate Studies

Over the past two decades, glycan microarrays have become powerful tools for exploring bacterial surface glycans and their recognition by proteins have been developed [217, 218]. The first publications reporting the development of glycan arrays came in 2002 from several independent research groups. Two reports demonstrated robotically printed arrays of diverse glycan libraries [219, 220], and many other laboratories demonstrated various approaches for immobilisation of glycans in printed slide or multiwell plate formats [221-223].

Glycans have been non-covalently and covalently immobilised to a variety of functionalised surfaces, including glass slides and microplates. For example, lipid tethers and 1, 3-dipolar cycloadditions between azides and alkynes were exploited by Fazio *et al.* to attach oligosaccharides to a C_{14} hydrocarbon chain that noncovalently binds to the microtiter well surface [221, 224]. They used this approach to successfully screen complex sugars for their specific interactions with fluorescein-labelled plant lectins. They then build

on those findings to develop a microplate array that incorporates the covalent array for improved ease of use and reusability [225]. However, for covalent binding NHS-or amine-coated surfaces were needed for alkyne-functionalisation and subsequent triazole formation for glycan attachment.

There appears to be only one entry in the literature for a microplate assay that examines glycan-bacterial binding and that is from the Gibson group in 2016. They used hydrazide-functionalised (carbo-BIND) microplates to glycosylated reducing carbohydrates to the surface with an aniline catalyst. Visualisation of bacterial binding was enabled with a two-step labelling procedure where bacteria were first biotinylated with biotin-NHS and then reacted with FITC-labelled streptavidin. This method produced a heterogenous surface of ring-closed and ring-opened sugars (60:40) and was claimed to give distinct binding profiles for Gram-positive and Gram-negative bacterial strains to carbohydrate functionalised surfaces. However, it could be argued that there are no glycan specific binding interactions for any of the bacterial strains probed (see Figure 3.9C). A summary of the data can be seen in Figure 3.9.
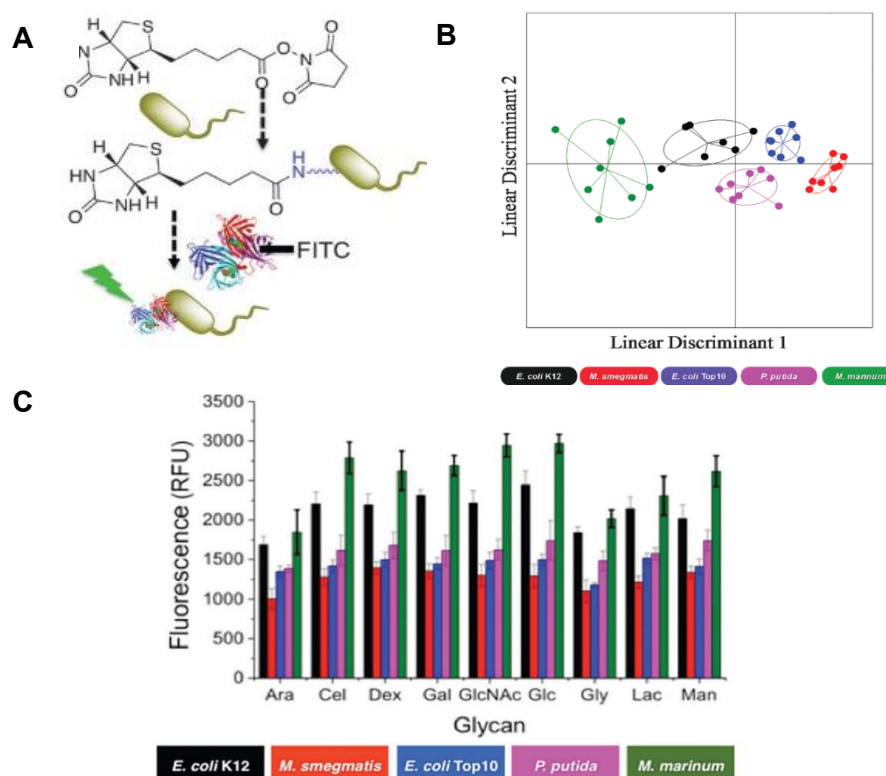
Figure 3.9: Summary of data of live bacterial binding to carbohydrate microplates by the Gibson group. (**A**) Schematic of binding and fluorescent labelling of bacteria. (**B**) Linear discriminant analysis model of bacterial binding. *E. coli* K12, *E. coli* Top10, *M. marinum*, *P. putida* and *M. smegmatis*. Each point represents a single binding profile to all the glycan surfaces, the centre point dictates the average value for all samples and the ellipses represent one standard deviation from the average. (**C**) Binding profiles of fluorescently labelled bacteria to glycan surfaces. Arabinose (Ara), cellobiose (Cel), dextran (Dex), galactose (Gal), glucose (Glc), N-acetyl-D-glucosamine (GlcNAc), glyceraldehyde (Gly), lactose (Lac) and mannose (Man). Reproduced with permission from the Royal Society of Chemistry [226].

3.3.2 Results and Discussion

3.3.2.1 Lectin Binding Assays

To generate a platform designed to determine the binding specificity of bacteria towards glycans requiring standard lab equipment, we used chemo-enzymatically synthesised bovine serum albumin (BSA) conjugated mannose (Figure 3.10) immobilised to polystyrene microtiter plates. We used the BSA linker as a functional group to immobilise the glycans by utilising surface amine-containing lysine residues of the BSA giving rise to BSA-mannose conjugates with differing numbers of mannose residues displayed on the surface (see Table 3.1).
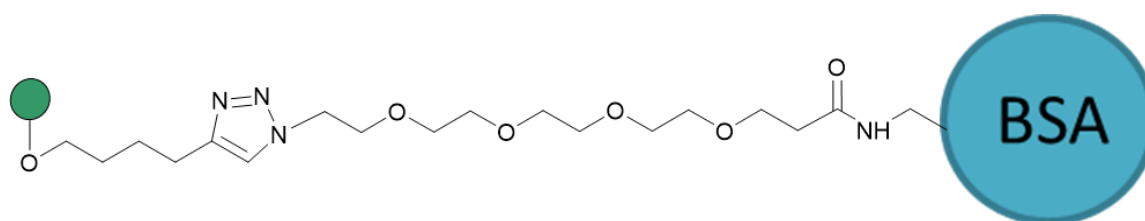
Figure 3.10: Composition of the mannosylated BSA-glycoconjugates used in this study.

Table 3.1: Table of mannose-containing compounds.

Compound No.	Compound Name
1	BSA-PEG$_4$-αMan (10 sugars)
2	BSA-PEG$_4$-αMan (15 sugars)
3	BSA-PEG$_4$-αMan (20 sugars)
4	BSA-PEG$_4$-αManα1,6Man (15 sugars)
5	BSA-PEG$_4$-αManα1,6Man (20 sugars)
6	BSA-PEG$_4$-αManα1,6Manα1,6Man (10 sugars)
7	BSA-PEG$_4$-αManα1,6Manα1,6Man (12 sugars)
8	BSA-PEG$_4$-αManα1,6Manα1,6Man (15 sugars)
9	Mannan, 1000 µg/mL
10	Mannan, 800 µg/mL
11	Mannan, 600 µg/mL
12	Mannan, 400 µg/mL

To examine the "printing" efficiency of the BSA-glycan conjugates, we first examined the fluorescent output with BSA-FITC conjugates (see Figure 3.11). This study revealed that non-covalent immobilisation of BSA-conjugates to polystyrene microtiter plates can be achieved under basic conditions. From 400 µg/mL to 50 µg/mL the fluorescent intensity remained around 120,000 FI due to surface saturation. At concentrations below 50 µg/mL the fluorescence intensity decreased as concentration decreased. Therefore, at BSA-FITC concentrations of 50 µg/mL or more, maximal binding was reached.
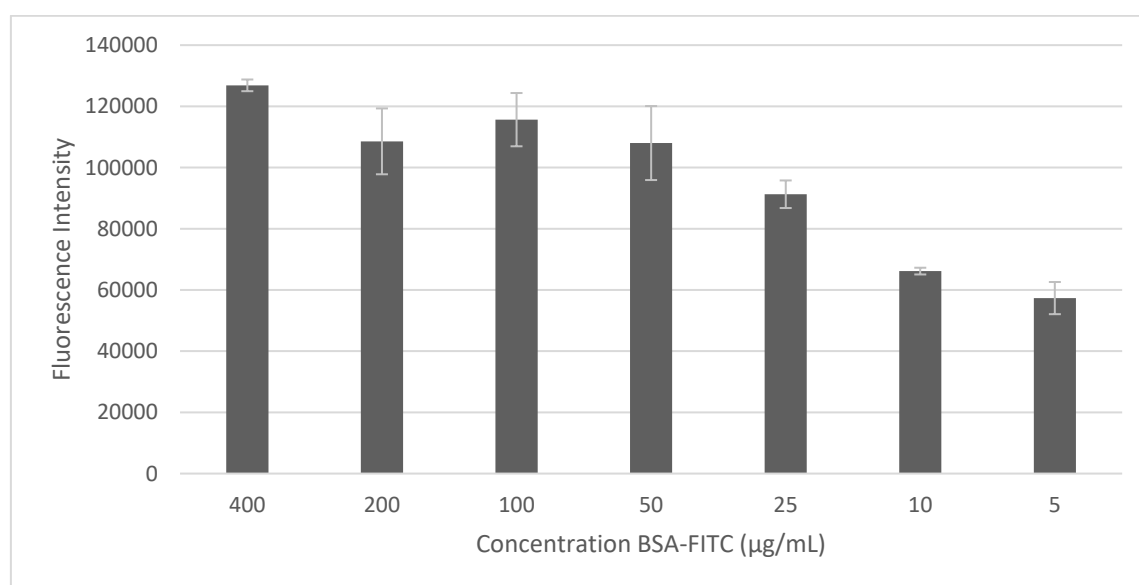
Figure 3.11: Fluorescence intensity of BSA-FITC directly bound to the polystyrene surface of the microplate. From 400 to 50 µg/mL BSA-FITC the surface is saturated and exhibits maximal fluorescence intensity. Below 50 µg/mL BSA-FITC, surface saturation decreases.

Plant lectins are essential tools in biological studies. The binding profiles of common plant lectins, which are naturally occurring glycan-binding proteins (GBPs), are very well studied and they are vital for identifying and characterizing specific glycan epitopes [227]. To validate the assay and determine the ability of the format to give specific binding properties of glycan-binding proteins, fluorescent plant lectins were used. Concanavalin A (ConA) from *Canavalia ensiformis* is a Man/Glc specific plant lectin commonly used lectin for verifying lectin-glycan interactions [228]. At pH 7, ConA is a tetramer with four active sites [229] and is stable upon surface immobilisation [230] and, therefore, can be functionalized on various surfaces with little effect from orientation as some active sites will be exposed to the surface. This makes ConA a good candidate for direct immobilisation onto a surface for use as a positive control, as it behaves as a model lectin.

First, the optimal concentration of fluorescein-ConA was evaluated for BSA-PEG$_4$-Man (10), i.e., ten single mannose residues conjugated to BSA via polyethylene glycol (PEG) - ylated linkers. Figure 3.12 shows that fluorescent intensity of ConA bound to BSA-PEG$_4$-Man was only half as strong as that of BSA-FITC but was of adequate intensity to identify binding over the negative control. Furthermore, 2 µg/mL of ConA was found to be the optimal concentration of fluorescent plant lectin, even at the lowest concentration of glycan, 50 µg/mL, with fluorescent intensity reducing dramatically as lectin concentration is reduced to 1 µg/mL or 0.5 µg/mL (see Figure 3.12).

When determining whether this format could be used to demonstrate binding specificity, BSA-2'FL was immobilised and probed with both a fluorescein-labelled fucose specific lectin, *Ulex europaeus* agglutinin I (UEA-I), and ConA. The lectins were tested with BSA-PEG$_4$-Man (15) to determine the ability of the assay to give specific binding interactions. Figure 3.13 shows that this assay format enables specific binding as the fluorescent output of BSA-2'FL with ConA is negligible when compared to its specific lectin, UEA-I. The same was seen when BSA-PEG$_4$-Man (15) with UEA-I - negligible binding was observed.
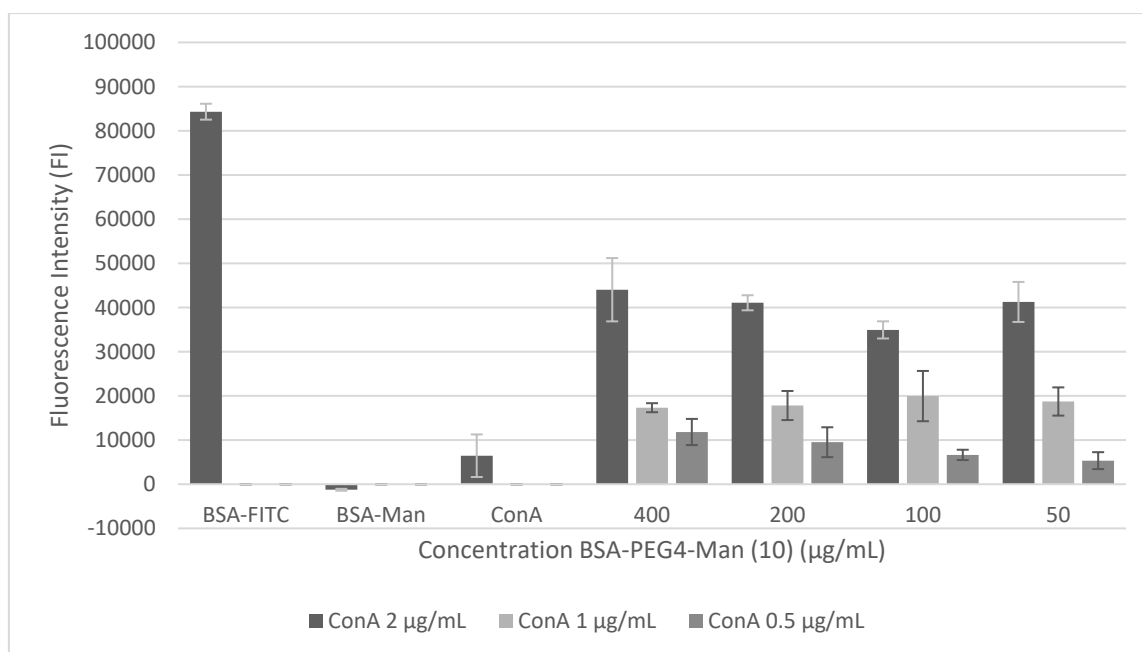
Figure 3.12: Fluorescence intensity of decreasing concentration of fluorescein-ConA bound to decreasing concentrations of BSA-PEG$_4$-Man (10) (compound 1). From 400 to 50 µg/mL BSA-PEG$_4$-Man (10) the surface is saturated. ConA (2 µg/mL) shows significantly higher fluorescence intensity than lower concentrations of ConA (1 µg/mL or 0.5 µg/ml).
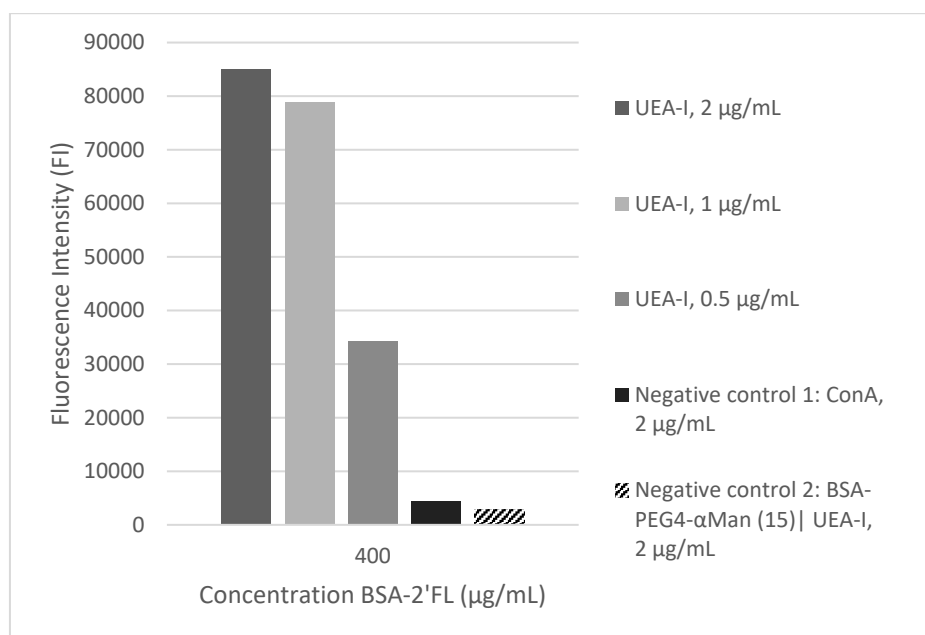
Figure 3.13: Fluorescence intensity of fluorescein-UEA-I or ConA bound to BSA-2'FL (400 µg/mL), including BSA-PEG$_4$-αMan (15) with UEA-I as an additional negative control. The assay offers binding specificity with plant lectins. UEA-I specifically bound to fucosylated glycans and not mannose. Equally, ConA did not bind to fucosylated glycans.

To investigate this further and examine the effect of glycan valency, glycan chain length and linker length on the efficiency of plant lectin binding, a series of mannose containing BSA-conjugates and mannan from *Saccharomyces cerevisiae* were immobilised and probed with fluorescein-ConA (see Table 3.1). Figure 3.14 shows that increasing glycan valency has a small effect on the binding ability of ConA. This may partly be explained by the presence of a monosaccharide and binding site and secondary extended binding site of ConA, where the lectin can also bind mannose trisaccharides [231].

Interestingly, mannan from *Saccharomyces cerevisiae* showed significantly greater ConA binding. Individual glycans display low affinity to their protein targets, but this is overcome in nature by multivalent display [232], thus plant lectins are usually dimers or tetramers to enable multivalent interactions [233]. The highly branched mannosylated structure of mannan may allow multivalent interactions resulting in greater binding capacity of the lectin to the mannan when compared to the BSA-conjugated mannose structures. The chemical nature of glycosidic linkages permits glycans a certain degree of flexibility, adopting a variety of three-dimensional shapes [234, 235]. Therefore, mannan is likely to be more inherently flexible enabling the correct conformational orientation for lectin binding.
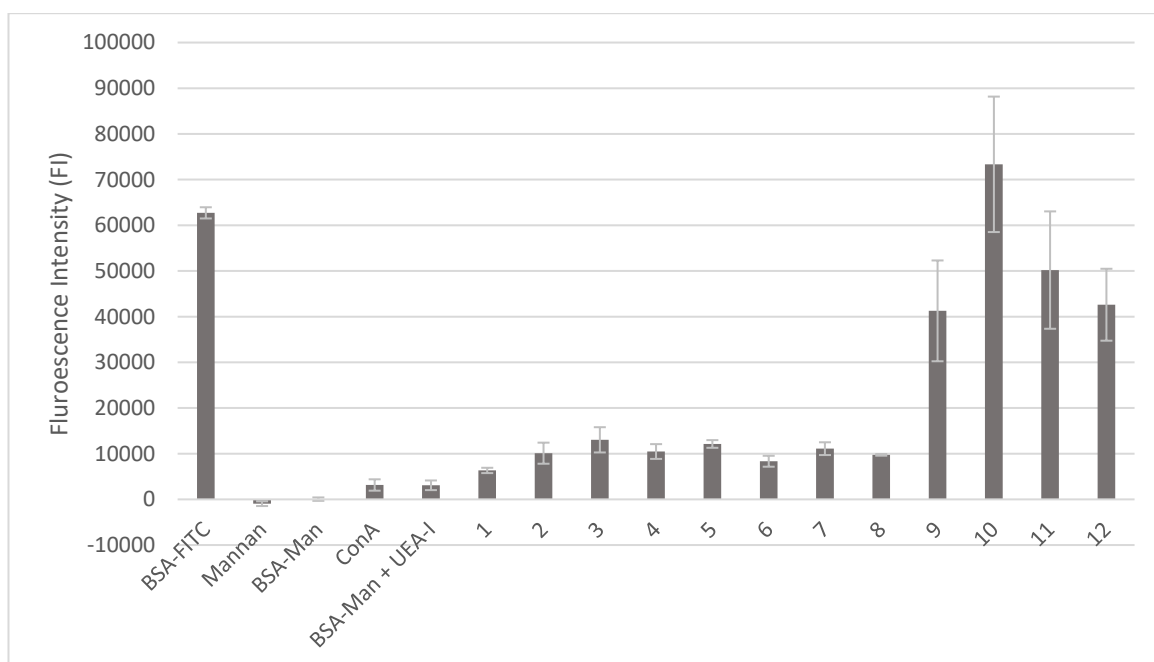
Figure 3.14: Fluorescence intensity of fluorescein-ConA bound to mannose-containing compounds 1 – 12 (see Table 3.1).

Furthermore, the use of BSA limits the number of mannose residues present in the correct orientation not only due to the size of BSA (BSA has a nominal size of 7.1 nm and a mass of 66.5kDa [236], but also due to its limited capacity to conjugate, and therefore, display sugar moieties. Bioconjugation of propargyl linkers on the surface of BSA to which sugar residues can be attached is restricted to the number of lysine residues present on the surface of the tertiary/quaternary BSA structure, of which BSA consists of 59 lysine residues, of which 30-35 have primary amines at the surface that can react [237]. Additionally, in this format, the orientation of the BSA-glycan to the surface is unknown, therefore not all the mannose residues present will be in the correct orientation for binding at any given time. It follows that adherence of mannan to the surface is much greater than mannosylated BSA conjugates due to the lack of protein component to facilitate covalent or non-covalent binding to the surface. In contrast to BSA-glycoconjugates, much higher concentrations of mannan can be bound to the surface with no surface saturation observed from 50 µg/mL to 400 µg/mL (see Figure 3.15), or concentrations up to 1 mg/mL (see compounds 9, 10, 11, 12 in Figure 3.14).
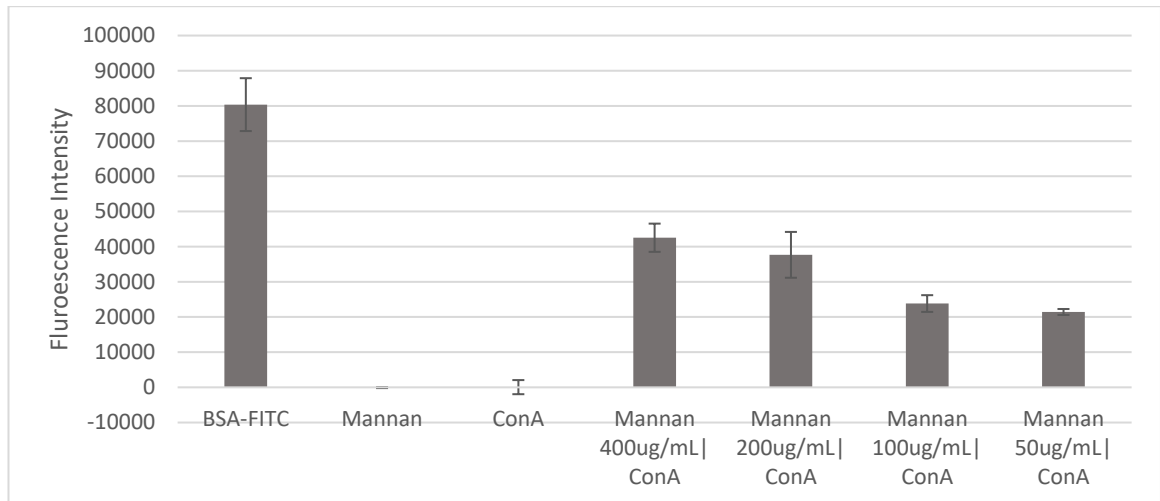
Figure 3.15: Fluorescence intensity of fluorescein-ConA bound to mannan from Saccharomyces cerevisiae.

3.3.2.2 Live Bacteria Binding Assays

With the validated assay in hand, we then moved to establish if the assay could show specific interactions of fluorescently labelled *E. coli* strains. The strains used were either WT (ORN178, FimH+) or mutants that do not express the mannose-binding FimH lectin (ORN208, FimH-). However, before the glycan-*E. coli* interactions could be probed, the labelling and dilution of *E. coli* cells were first optimised.

Initially, different methods of fluorescently labelled the cells was investigated: basic FITC labelling, CellTrace™ CFSE Cell Proliferation Kit (Thermo Fischer Scientific) and a two-step labelling procedure where bacteria were first biotinylated with biotin-NHS, followed by a reaction with FITC-labelled streptavidin [226]. Figure 3.16 summarises the respective fluorescent intensities of *E. coli* strains labelled using the techniques listed above. The CellTrace™ CFSE Cell Proliferation Kit showed the highest fluorescence intensity at 1 x 10^7 cells/mL (188444 ± 4654 and 129941 ± 1481 for ORN178 and ORN208, respectively). However, upon dilution to 1 x 10^6 cells/mL and beyond, fluorescent intensity decreased dramatically (24035 ± 326 and 17979 ± 112 for ORN178 and ORN208, respectively). The second most effective labelling method was the basic FITC procedure where 1 x 10^9 cells/mL showed a fluorescent intensity of 44840 and 62040 for ORN178 and ORN208, respectively. The two-step method using biotin-NHS and FITC-streptavidin appeared to produce the lowest fluorescent intensity (45822 ± 15530 and 34043 ± 956, respectively). Furthermore, the standard deviation of the data, represented by the error bars indicated an increased unreliability of this technique compared to the other labelling methods. Going

forward, the basic FITC protocol was employed due to its simplicity, cost effectiveness. Also, more cells/mL could be labelled at one time.

Fluorescent microscopy studies revealed that this technique did not appear to give to uniform labelling (see Figure 3.17A). Consequently, FITC labelling was optimised by repeating the labelling step one or two times to increase labelling efficiency, which resulted in a much higher fluorescent intensity of cells in solution. As FITC reacts with a primary amine on the bacterial surface to form a covalent amide bond [238-241], it is thought that free amines in solution were eliminated by first staining, which allowed for better cell staining in consequent labelling steps. For example, for 1×10^9 ORN178 cells/mL the fluorescence intensity increased from 13733 ± 260 after one labelling step, to 252428 ± 5325 following three repetitions (see Figure 3.17B). Therefore, the FITC labelling procedure with three repetitions was used for all future *E. coli* labelling.
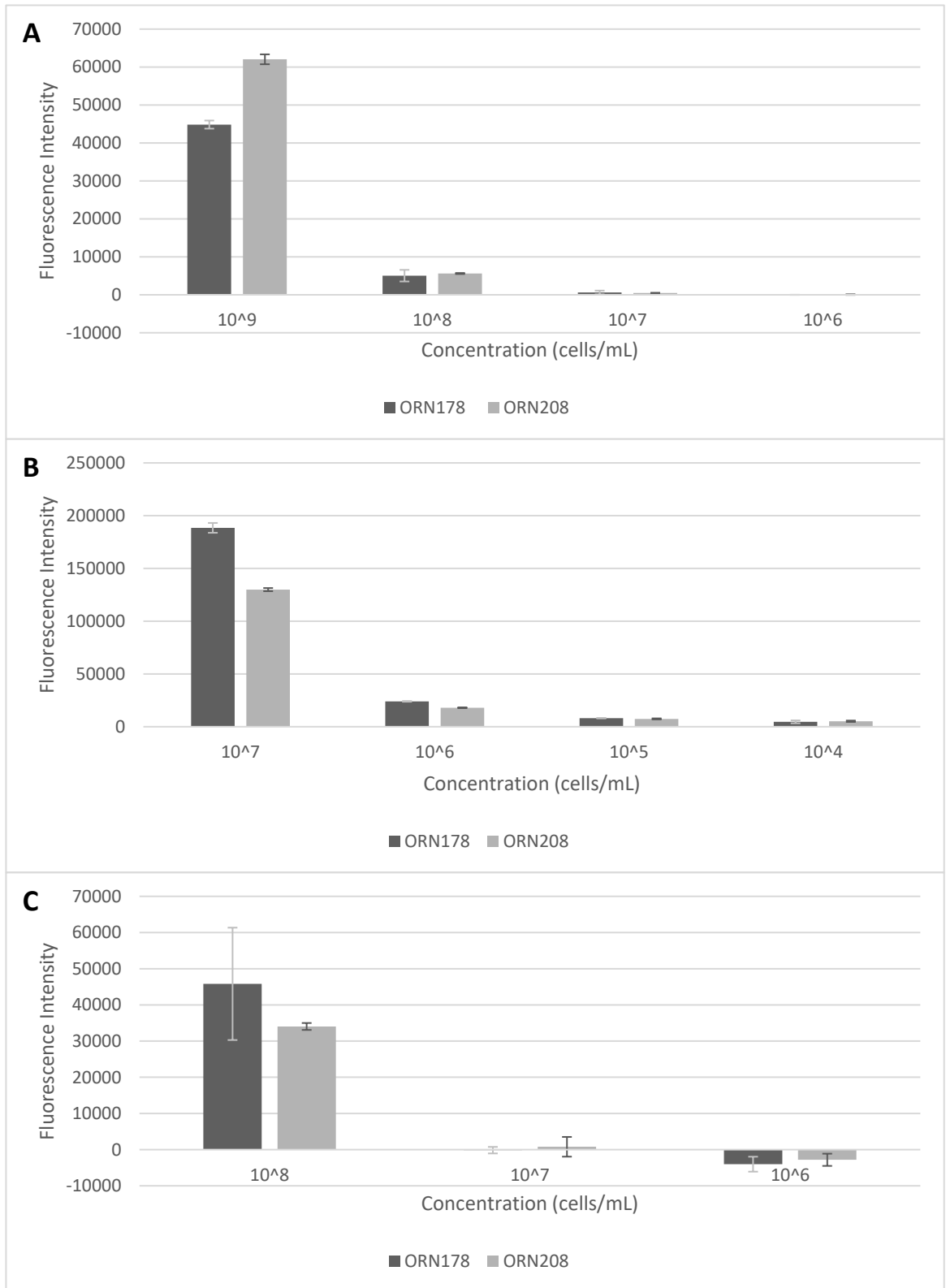
Figure 3.14: Summary of labelling procedures tested for fluorescent labelling of E. coli strains, ORN178 and ORN208. (**A**) basic FITC-labelling, (**B**) CellTrace™ CFSE Cell Proliferation Kit and (**C**) a two-step procedure using biotin-NHS followed by FITC-labelled streptavidin.
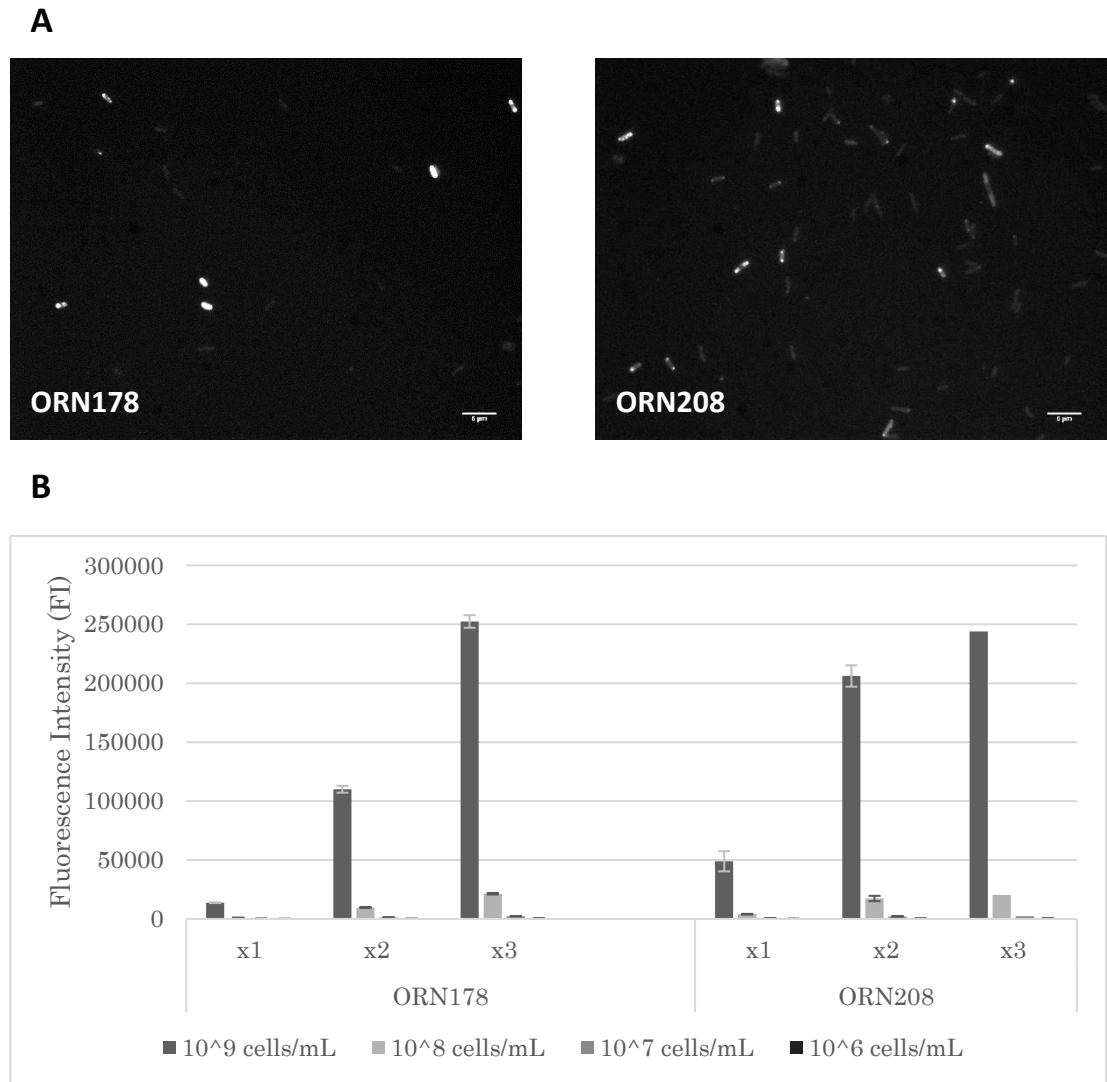
Figure 3.15: Fluorescent microscopy imaging of FITC-labelled *E. coli* strains ORN178 and ORN208 (A) and optimisation of FITC-labelling (B). (**A**) Un-optimised method for FITC labelling *E. coli* led to inconsistent labelling. (**B**) FITC labelling was optimised by carrying out the labelling step once (X1), twice (X2) or three times (X3) to improve the fluorescent intensity emitted by the bacterial cells.

With a functional glycan assay method and fluorescently labelled *E.* coli strains in hand, we probed the assay containing BSA-PEG$_4$-Man (10) (BSA-Man) and mannan from *Saccharomyces cerevisiae* (see Figure 3.18). When a glycan assay was either blocked or not blocked with 3% BSA, bound BSA-Man and mannan were treated with FimH+/- *E. coli* strains. We saw that blocking made no difference to the binding capabilities of the *E. coli* in the case of BSA-Man, however, completely diminished binding to mannan. Furthermore, in the case of mannan, *E. coli* FimH+ was found to bind specifically to mannan, with no significant binding seen for FimH- (106400 ± 10100 and 55600 ± 6000 for FimH+ and FimH-, respectively). However, residual binding of ORN208 was seen,

similarly to data from Disney et al., (see Figure 3.18A). In this case, it appears that there is some non-specific binding of *E. coli* to BSA or binding of other receptors to mannose as seen by high background from both the bacterial negative controls.
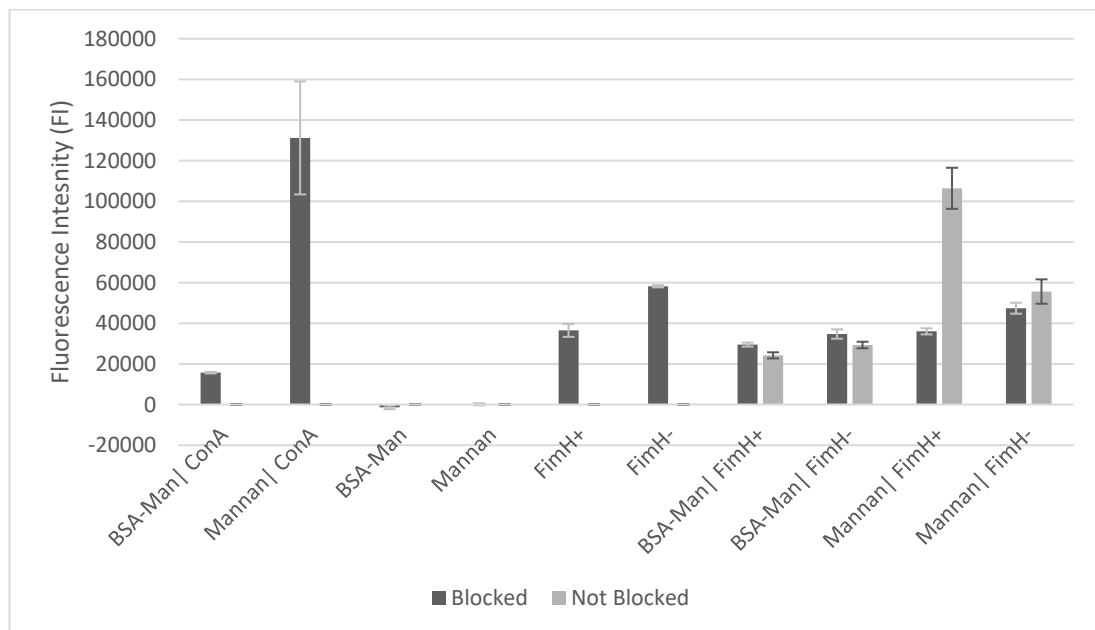
Figure 3.16: Fluorescence intensity of FITC-labelled *E. coli* (FimH+ or FimH-) bound to BSA-PEG4-Man (10) (BSA-Man) or mannan from *Saccharomyces cerevisiae*. Specific binding of ORN178 (FimH+) to mannan under non-blocking conditions was seen when compared to ORN208 (FimH-).

This data indicates that mannan could be used as a basis for the design of immunostimulatory compounds, inhibitory compounds or receptor decoys for mannose-binding pathogens allowing selective colonisation. Mannose-based treatments for uropathogenic *Escherichia coli*-induced urinary tract infections are already in development [242]. Moreover, mannan has previously been shown to enhance IL-12 production by increasing bacterial uptake and endosomal degradation in dendritic cells stimulated by Gram positive bacteria, *L. acidophilus* and *S. aureus* [243]. Furthermore, mannan from *Saccharomyces cerevisiae* can induce beta-defensin-1 expression via Dectin-2-Syk-p38 pathways in ovine ruminal epithelial cells [244] and mannan-oligosaccharide supplementation attenuates intestinal inflammation and barrier dysfunction induced by *Escherichia coli* in broilers [245]. There is also precedent for this in other species, where recent studies have shown that cell wall mannan from fungal species play a pivotal in host immune responses, allowing protection against fungal infection or facilitate immune invasion [246-248]. Mannans can induce anti-fungal protective immunity in *Candida* species [249, 250], for example, cell wall mannan of *Candida krusei* mediates dendritic cell apoptosis and orchestrates Th17 polarization [251]. In the literature, the inhibitory and

selective adhesion properties of mannosylated compounds have been examined in recent years. One example used hemagglutination assays to examine the inhibitory properties of mannose-tipped ligands and those containing a triazole ring have potential as a potent mannosyl ligands for targeting the FimH fimbrial lectin [252]. Alternatively, using saccharides to selectively colonise surfaces may be beneficial to many biotechnologies where promoting adhesion of benign microbes could prevent surface attachment of pathogens. This hypothesis was investigated by Young *et al.,* who presented that mannose-decorated glycopolymer surfaces could demonstrate the selective adhesion of *Shewanella oneidensis* [253].

When evaluating this system as a useful tool for glycan assays, there are some limitations to consider. Throughout the optimisation and validation process, the fluorescent intensity of bound entities was inconsistent. Possible reasons for this regarding the use of BSA as a bioconjugate were discussed in 3.3.2.1. To overcome these issues, future works could use multivalent dendrimers with long flexible linkers to increase incidence and ease of bacterial binding. Many studies have shown that multivalency improves the efficiency of antiadhesive compounds and an increase in flexibility of both protein and ligand can result in an increase in the stability of macromolecular complexes [254, 255]. Kim et al., showed that the binding affinities of toxins and pathogens for glycans are highly glycan density dependent. Specifically, toxins and pathogens bind to glycans more strongly as the valency of the glycans on the microarrays is increased from 1 to 4 [256]. Furthermore, dodecamannosylated fullerenes demonstrated improved adhesion for live uropathogenic *E. coli* [190]. Surface plasmon resonance measurements showed that glycodendrimers bearing aryl α-D-mannopyranoside residue clusters (tetramers) are the best ligands known to date toward *Escherichia coli* K12 FimH with subnanomolar affinities, where these clusters were 1000 times more potent than mannose for their capacity to inhibit the binding of E. coli to erythrocytes in vitro. [257].

A similar approach has been investigated with fucosylated compounds. When investigating the antibiofilm properties of fucose, a calix[4]arene scaffold in which multiple α-L-fucose units are linked to the platform via a flexible spacer of 6-aminocaproic acid, exhibited greater inhibition than its non-macrocyclic counterpart, which suggested a multivalency effect [258]. Similarly, Verkhnyatskaya *et al.,* found that cyclodextrins decorated with multiple fucose moieties showed greater antiadhesive properties than compounds mono-fucosylated compounds [259]. Alternatively, a different method of detection could be employed. Surface plasmon resonance (SPR) techniques have been known to give supramolecular recognition of *E. coli* by fluorescent oligo(phenyleneethynylene)s with mannopyranoside termini groups [215]. Furthermore,

this technique identified that two mannose groups are necessary to have a specific interaction with type 1 pilli.

To this end, in collaboration with the Blanford group (Manchester Institute of Biotechnology), similar lectin- or bacterial-binding interactions were investigated using an alternative glycan-based biosensor format. A mannose-functionalised surface of gold or graphene were used for probing the detection mechanism via quartz crystal microbalance (QCM) with dissipation and graphene field-effect transistor (GFET) to allow for both acoustic and electronic modes of sensing (see Figure 3.19). The QCM is a biosensor platform that incorporates a mechanical transducer, which operates on the principle of mass detection from changes to resonance frequency. A similar scheme is employed with GFET, but a graphene surface is used, and where modified conductivity of the surface is indicative of binding interactions. This format allowed selective and sensitive adhesion of *E. coli* to the mannose-functionalised biosensor; where ORN178 showed strong and irreversible binding to the mannose functionalised surface, whilst the ORN208 mutant (FimH-) did not bind (see Figure 3.20 and Figure 3.21), the methods of which can be seen in section 5.3.5 [260]. The improved selectivity and sensitivity of this biosensor may be due to the use of a small sugar probe, which circumvented the stability associated problems seen with large protein based bioprobes, such as the BSA glycoconjugates.
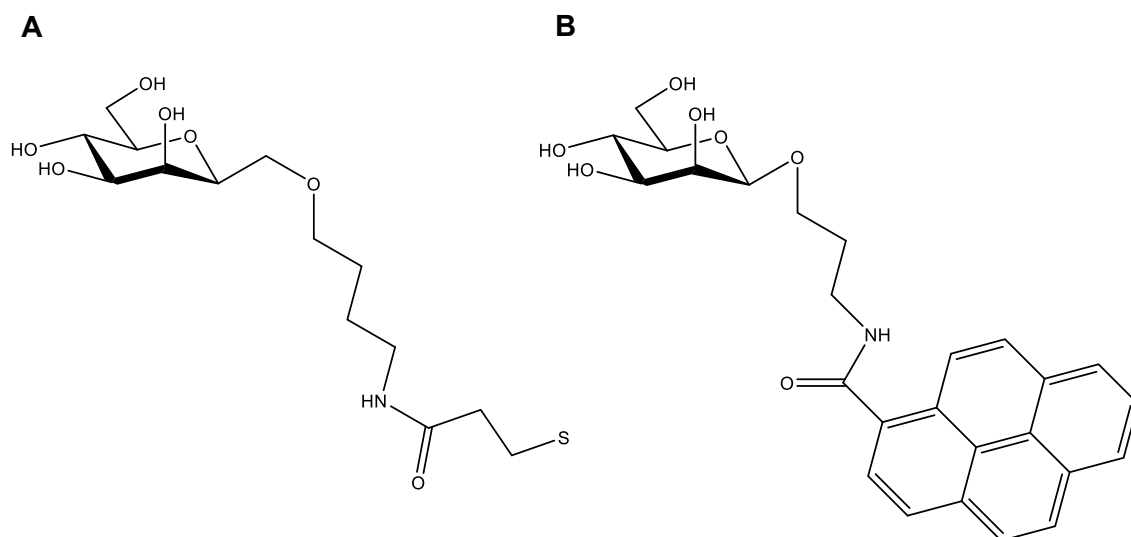
Figure 3.17: Mannose-functionalised gold surface for QCM (A) and graphene for GFET (B). (**A**) Drop casting of 3,3'-Dithiodipropionic acid di(N-hydroxysuccinimide ester) dissolved in DMF was incubated on the gold before rinsing with water and ethanol. This was then repeated with 3-azidopropyl 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranoside (3-APM) to obtain a mannose-functionalised gold surface. (**B**) 1-Pyrenebutanoic acid succinimidyl ester (in methanol) was drop casted onto the graphene surface and incubated followed by addition of 3-APM overnight to obtain functionalized surface.
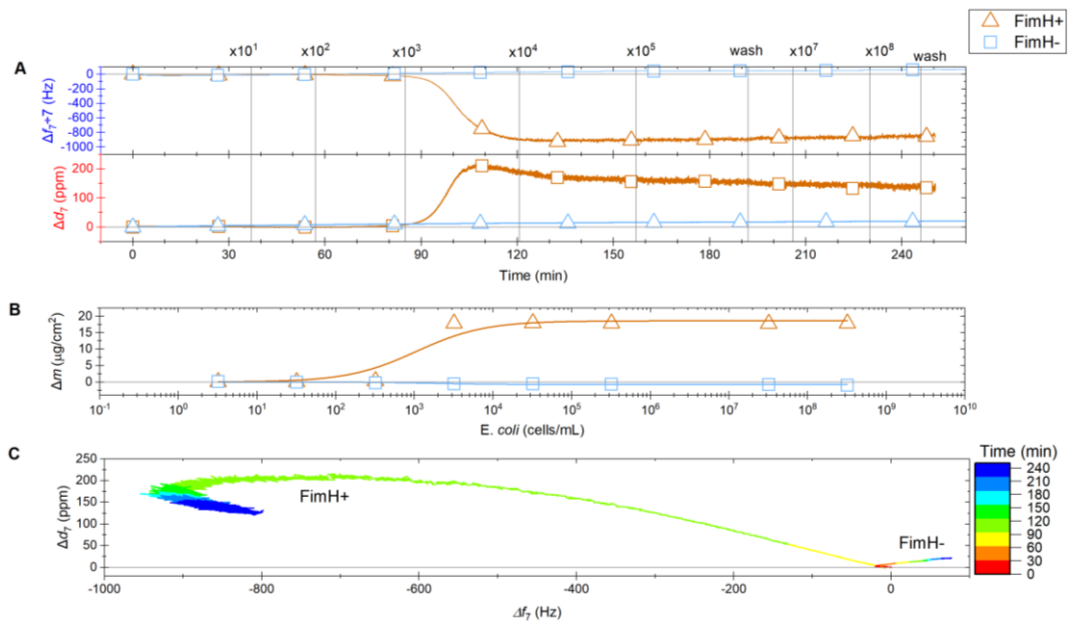
Figure 3.20: QCM of *E. coli* variants binding to mannose functionalized surface. (**A**) Dissipation and frequency traces are plotted over increasing *E. coli* concentration. (**B**) Calculated areal mass density using a viscoelastic model of the 3rd, 5th and 7th harmonics. (**C**) Dissipation plotted as a function of frequency to show conformational binding of *E. coli* on the functionalized surfaces [260].
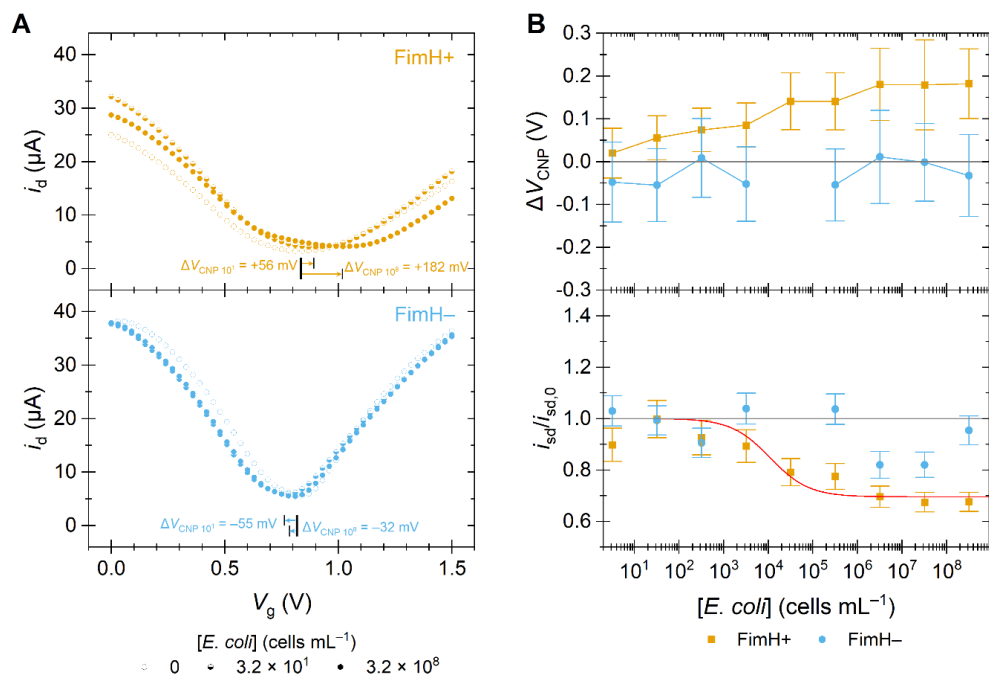
Figure 3.21: Binding of *E. coli* FimH variants on GFET surfaces functionalized with mannose. (**A**) Full transfer curves showing shift in voltage at the charge neutrality point (V_{CNP}) for FimH+ and FimH- variants at initial and final *E. coli* concentrations. (**B**) The shift V_{CNP} plotted for both variants and baselined with the V_{CNP} of the initial PBS solution and the normalised i_{sd} change at 1.2V Vg [260].

3.3.2.3 Flow Cytometry Aggregation Assay

From the assay data, we know that specific binding can be seen between mannan and FimH+ (ORN178) *E. coli* cells compared with the FimH- (ORN208) strain. To observe if mannan was capable of causing selective aggregation of *E. coli* strains, a dilution of mannan was incubated with *E. coli* cell lines (1 x 10⁴ cells/mL) and was analysed by cell cytometry to count cell aggregates. Data was plotted as an average across replicates. Initially, it was observed that only mannan at 100 μg/mL can facilitate aggregation of ORN178 cells, where 73.1% of cells form two or more aggregates compared to a mean of 26%, but not any other concentration of mannan or any ORN208 samples. However, upon repetition data suggests that there is no correlation between mannan concentration and aggregation in either strain (see Figure 3.22) but further optimisation and repeats are needed to draw firm conclusions. If successful, mannan from *Saccharomyces cerevisiae*, provides an easily accessible food supplement with potential inhibitory capabilities for uropathogenic *E. coli*.
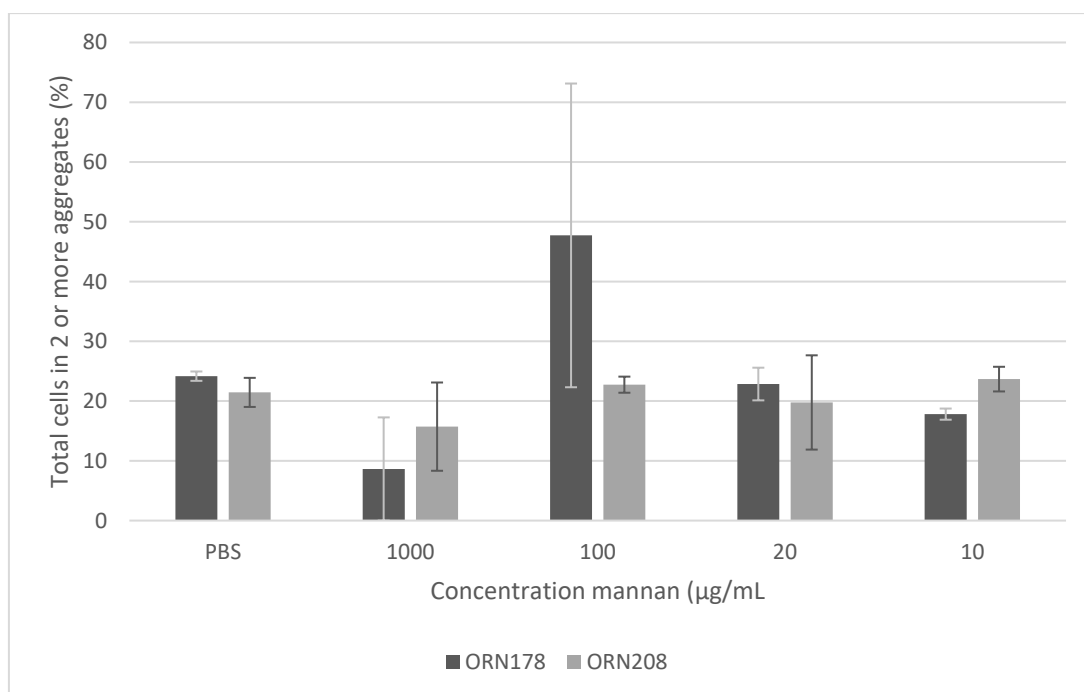
Figure 3.22: Flow cytometry analysis of *E. coli* cell lines (FimH+, ORN178 and FimH-, ORN208) in the presence of decreasing concentrations of mannan showing the total number of cells present in two or more aggregates. Initial data suggests that there is no aggregation of either strain in presence of mannan. Data is plotted as an average across experimental duplicates with the error bars representing standard deviation.

3.3.3 Conclusions

When evaluating simple microtiter plate glycan assays, where BSA-mannose conjugated compounds were non-covalently immobilised to a polystyrene microplate. To show that this binding principle was valid, a FITC conjugated BSA carrier protein was used. Using dilutions to determine minimum binding concentrations, maximal binding was achieved above concentrations of 50 µg/mL BSA-FITC.

The selectivity of the assay was then tested with fluorescent-plant lectins as a proof of concept. BSA-conjugated mannose was non-covalently bound to the surface of the microplate, displaying different numbers of glycans on the surface and different mannose chain lengths. Additionally, a highly valent, mannosylated polysaccharide, mannan from *Saccharomyces cerevisiae,* was bound to the surface. All compounds elicited selective binding with the fluorescent-plant lectin, ConA, and negligible binding to UEA-I. Moreover, when BSA-conjugated 2'FL was bound, the reverse was true, where selective binding of UEA-I was seen with negligible binding to ConA.

The assay was tested with live bacteria, *E coli* ORN178 (FimH+) and ORN208 (FimH-); strains with varied FimH mannose lectin-binding domains to test the selectivity of the test. The assay was able to identify selective binding interactions with live bacteria, where mannan showed the greatest binding affinity. However, the assay had limited sensitivity and appeared to be limited by the size of the carrier protein, and number and flexibility of the displayed glycans as mannan was able to elicit a much high fluorescent response, presumably due to its greater glycan valency. By modifying the method of immobilisation, i.e., eliminate the use of BSA and instead use a smaller functionalised probe with a long flexible linker for multivalent glycan display, the binding affinity of *E. coli* may be improved. The methodology also lends itself for adaption to other mediums, such as gold nanoparticles, quantum dots and gold/graphene chips for use as highly specific biosensors for detection of pathogens or identification of pathogen-specific glycan motifs for the development of targeted drug therapies. This was exemplified by the collaborative project with the Blanford group, where mannose-functionalised small probe biosensors showed highly selective and irreversible binding with *E. coli* strains ORN178 and ORN208. This is advantageous in biosensor design as it allows for robust pathogen detection.

Finally, flow cytometry methodology was investigated to determine whether mannan could facilitate selective aggregation of *E. coli* strains ORN178 and ORN208. Following brief optimisation, our studies did not show any indication of bacterial aggregation after pre-treatment with mannan. However, this was only a preliminary study and with more time, the correct conditions may provoke a different response.

3.4 References

10. Bode, L. (2006) Recent advances on structure, metabolism, and function of human milk oligosaccharides, *J Nutr.* **136**, 2127-2130.

40. Masi, A. C. & Stewart, C. J. (2022) Untangling human milk oligosaccharides and infant gut microbiome, *iScience.* **25**, 103542.

157. Pacheco, A. R., Curtis, M. M., Ritchie, J. M., Munera, D., Waldor, M. K., Moreira, C. G. & Sperandio, V. (2012) Fucose sensing regulates bacterial intestinal colonization, *Nature.* **492**, 113-117.

158. Garber, J. M., Hennet, T. & Szymanski, C. M. (2021) Significance of fucose in intestinal health and disease, *Mol Microbiol.* **115**, 1086-1093.

188. Disney, M. D. & Seeberger, P. H. (2004) The use of carbohydrate microarrays to study carbohydrate-cell interactions and to detect pathogens, *Chem Biol.* **11**, 1701-1707.

189. Mukhopadhyay, B., Martins, M. B., Karamanska, R., Russell, D. A. & Field, R. A. (2009) Bacterial detection using carbohydrate-functionalised CdS quantum dots: a model study exploiting *E. coli* recognition of mannosides, *Tetrahedron Lett.* **50**, 886-889.

190. Kalograiaki, I., Abellan-Flos, M., Fernandez, L. A., Menendez, M., Vincent, S. P. & Solis, D. (2018) Direct evaluation of live uropathogenic *Escherichia coli* adhesion and efficiency of antiadhesive compounds using a simple microarray approach, *Anal Chem.*

191. Sprenger, N., Tytgat, H. L. P., Binia, A., Austin, S. & Singhal, A. (2022) Biology of human milk oligosaccharides: From basic science to clinical evidence, *Journal of Human Nutrition and Dietetics.* **35**, 280-299.

192. Elison, E., Vigsnaes, L. K., Krogsgaard, L. R., Rasmussen, J., Sorensen, N., McConnell, B., Hennet, T., Sommer, M. O. A. & Bytzer, P. (2016) Oral supplementation of healthy adults with 2-O-fucosyllactose and lacto-N-neotetraose is well tolerated and shifts the intestinal microbiota, *British Journal of Nutrition.* **116**, 1356-1368.

193. Chen, Y. M., Zhu, Y. & Lin, E. C. (1987) The organization of the fuc regulon specifying L-fucose dissimilation in *Escherichia coli* K12 as determined by gene cloning, *Mol Gen Genet.* **210**, 331-7.

194. Chen, Y. M., Zhu, Y. & Lin, E. C. (1987) NAD-linked aldehyde dehydrogenase for aerobic utilization of L-fucose and L-rhamnose by *Escherichia coli*, *J Bacteriol.* **169**, 3289-94.

195. Baldomà, L. & Aguilar, J. (1988) Metabolism of L-fucose and L-rhamnose in *Escherichia coli*: aerobic-anaerobic regulation of L-lactaldehyde dissimilation, *J Bacteriol.* **170**, 416-21.

196. Bradley, S. A., Tinsley, C. R., Muiry, J. A. & Henderson, P. J. (1987) Proton-linked L-fucose transport in *Escherichia coli*, *Biochem J.* **248**, 495-500.

197. Zhu, Y. & Lin, E. C. (1988) A mutant crp allele that differentially activates the operons of the fuc regulon in Escherichia coli, *J Bacteriol.* **170**, 2352-8.

198. Tao, Y.-M., Bu, C.-Y., Zou, L.-H., Hu, Y.-L., Zheng, Z.-J. & Ouyang, J. (2021) A comprehensive review on microbial production of 1,2-propanediol: micro-organisms, metabolic pathways, and metabolic engineering in *Biotechnol Biofuels* pp. 216

199. Sridhara, S. & Wu, T. T. (1969) Purification and properties of lactaldehyde dehydrogenase from *Escherichia coli*, *J Biol Chem.* **244**, 5233-8.

200. Cocks, G. T., Aguilar, T. & Lin, E. C. (1974) Evolution of L-1, 2-propanediol catabolism in *Escherichia coli* by recruitment of enzymes for L-fucose and L-lactate metabolism, *J Bacteriol.* **118**, 83-8.

201. Boronat, A. & Aguilar, J. (1981) Metabolism of L-fucose and L-rhamnose in *Escherichia coli*: differences in induction of propanediol oxidoreductase, *J Bacteriol.* **147**, 181-5.

202. Kim, J., Cheong, Y. E., Jung, I. & Kim, K. H. (2019) Metabolomic and transcriptomic analyses of *Escherichia coli* for efficient fermentation of L-fucose in *Mar Drugs*

203. You, J., Lin, S. & Jiang, T. (2019) Origins and evolution of the α-L-fucosidases: from bacteria to metazoans, *Frontiers in Microbiology.* **10**.

204. Luijkx, Y. M. C. A., Bleumink, N. M. C., Jiang, J., Overkleeft, H. S., Wösten, M. M. S. M., Strijbis, K. & Wennekes, T. (2020) *Bacteroides fragilis* fucosidases facilitate growth and invasion of *Campylobacter jejuni* in the presence of mucins, *Cell Microbiol.* **n/a**, e13252.

205. Rodríguez-Díaz, J., Monedero, V. & Yebra, M. J. (2011) Utilization of natural fucosylated oligosaccharides by three novel alpha-L-fucosidases from a probiotic *Lactobacillus casei* strain, *Appl Environ Microbiol.* **77**, 703-705.

206. Seeberger, P. H. (2017) Glycan arrays and other tools produced by automated glycan assembly, *Perspectives in Science.* **11**, 11-17.

207. Liu, H.-Z., Liu, L., Hui, H. & Wang, Q. (2015) Structural characterization and antineoplastic activity of *Saccharomyces cerevisiae* mannoprotein, *Int J Food Prop.* **18**, 359-371.

208. Oba, S., Sunagawa, T., Tanihiro, R., Awashima, K., Sugiyama, H., Odani, T., Nakamura, Y., Kondo, A., Sasaki, D. & Sasaki, K. (2020) Prebiotic effects of yeast mannan, which selectively promotes *Bacteroides thetaiotaomicron* and *Bacteroides ovatus* in a human colonic microbiota model, *Scientific Reports.* **10**.

209. Vinogradov, E., Petersen, B. & Bock, K. (1998) Structural analysis of the intact polysaccharide mannan from *Saccharomyces cerevisiae* yeast using H-1 and C-13 NMR spectroscopy at 750 MHz, *Carbohydrate Research.* **307**, 177-183.

210. Harris, S. L., Spears, P. A., Havell, E. A., Hamrick, T. S., Horton, J. R. & Orndorff, P. E. (2001) Characterization of *Escherichia coli* type 1 pilus mutants with altered binding specificities, *J Bacteriol.* **183**, 4099-102.

211. Busch, A., Phan, G. & Waksman, G. (2015) Molecular mechanism of bacterial type 1 and P pili assembly, *Philos Trans A Math Phys Eng Sci.* **373**.

212. Oba, S., Sunagawa, T., Tanihiro, R., Awashima, K., Sugiyama, H., Odani, T., Nakamura, Y., Kondo, A., Sasaki, D. & Sasaki, K. (2021) Author Correction: Prebiotic effects of yeast mannan, which selectively promotes *Bacteroides thetaiotaomicron* and *Bacteroides ovatus* in a human colonic microbiota model, *Scientific Reports.* **11**, 3741.

213. Nimrichter, L., Gargir, A., Gortler, M., Altstock, R. T., Shtevi, A., Weisshaus, O., Fire, E., Dotan, N. & Schnaar, R. L. (2004) Intact cell adhesion to glycan microarrays, *Glycobiology.* **14**, 197-203.

214. Disney, M. D., Zheng, J., Swager, T. M. & Seeberger, P. H. (2004) Detection of bacteria with carbohydrate-functionalized fluorescent polymers, *J Am Chem Soc.* **126**, 13343-13346.

215. Arias, E., Méndez, M. T., Arias, E., Moggio, I., Ledezma, A., Romero, J., Margheri, G. & Giorgetti, E. (2017) Supramolecular recognition of *Escherichia coli* bacteria by fluorescent oligo(phenyleneethynylene)s with mannopyranoside termini groups, *Sensors (Basel, Switzerland).* **17**.

216. Mader, A., Gruber, K., Castelli, R., Hermann, B. A., Seeberger, P. H., Rädler, J. O. & Leisner, M. (2012) Discrimination of *Escherichia coli* strains using glycan cantilever array sensors, *Nano Lett.* **12**, 420-3.

217. Shin, I., Park, S. & Lee, M. R. (2005) Carbohydrate microarrays: an advanced technology for functional studies of glycans, *Chemistry (Weinheim an der Bergstrasse, Germany).* **11**, 2894-901.

218. Kim, Y., Hyun, J. Y. & Shin, I. (2022) Glycan microarrays from construction to applications, *Chemical Society Reviews.* **51**, 8276-8299.

219. Fukui, S., Feizi, T., Galustian, C., Lawson, A. M. & Chai, W. (2002) Oligosaccharide microarrays for high-throughput detection and specificity assignments of carbohydrate-protein interactions, *Nat Biotechnol.* **20**, 1011-7.

220. Wang, D., Liu, S., Trummer, B. J., Deng, C. & Wang, A. (2002) Carbohydrate microarrays for the recognition of cross-reactive molecular markers of microbes and host cells, *Nat Biotechnol.* **20**, 275-81.

221. Fazio, F., Bryan, M. C., Blixt, O., Paulson, J. C. & Wong, C.-H. (2002) Synthesis of sugar arrays in microtiter plate, *J Am Chem Soc.* **124**, 14397-14402.

222. Houseman, B. T. & Mrksich, M. Carbohydrate arrays for the evaluation of protein binding and enzymatic modification.

223. Willats, W. G., Rasmussen, S. E., Kristensen, T., Mikkelsen, J. D. & Knox, J. P. (2002) Sugar-coated microarrays: A novel slide surface for the high-throughput analysis of glycans, *Proteomics.* **2**.

224. Bryan, M. C., Plettenburg, O., Sears, P., Rabuka, D., Wacowich-Sgarbi, S. & Wong, C.-H. (2002) Saccharide display on microtiter plates, *Chem Biol.* **9**, 713-720.

225. Bryan, M. C., Fazio, F., Lee, H.-K., Huang, C.-Y., Chang, A., Best, M. D., Calarese, D. A., Blixt, O., Paulson, J. C., Burton, D., Wilson, I. A. & Wong, C.-H. (2004) Covalent display of oligosaccharide arrays in microtiter plates, *J Am Chem Soc.* **126**, 8640-8641.

226. Otten, L., Fullam, E. & Gibson, M. I. (2016) Discrimination between bacterial species by ratiometric analysis of their carbohydrate binding profile, *Molecular bioSystems.* **12**, 341-344.

227. Gao, C., Wei, M., McKitrick, T. R., McQuillan, A. M., Heimburg-Molinaro, J. & Cummings, R. D. (2019) Glycan microarrays as chemical tools for identifying glycan recognition by immune proteins, *Frontiers in Chemistry.* **7**.

228. Bhattacharyya, L. & Brewer, C. F. (1986) Precipitation of concanavalin A by a high mannose type glycopeptide, *Biochem Biophys Res Commun.* **137**, 670-4.

229. Cavada, B. S., Pinto-Junior, V. R., Osterne, V. J. S. & Nascimento, K. S. (2018) ConA-like lectins: high similarity proteins as models to study structure/biological activities relationships, *International journal of molecular sciences.* **20**.

230. Komolov, K. E., Senin, I. I., Philippov, P. P. & KW, K. (2002) Elucidating kinetic and thermodynamic constants for interaction of G protein subunits and receptors by surface plasmon resonance spectroscopy, *Methods Enzymol.* **344**, 15-42.

231. Mandal, D. K., Bhattacharyya, L., Koenig, S. H., Brown, R. D., Oscarson, S. & Brewer, C. F. (1994) Studies of the binding specificity of concanavalin A. Nature of the extended binding site for asparagine-linked carbohydrates, *Biochemistry.* **33**, 1157-1162.

232. Lundquist, J. J. & Toone, E. J. (2002) The cluster glycoside effect, *Chemical Reviews.* **102**, 555-578.

233. Reeke, G. N., Jr., Becker, J. W., Cunningham, B. A., Wang, J. L., Yahara, I. & Edelman, G. M. (1975) Structure and function of concanavalin A, *Adv Exp Med Biol.* **55**, 13-33.

234. Gimeno, A., Valverde, P., Ardá, A. & Jiménez-Barbero, J. (2020) Glycan structures and their interactions with proteins. A NMR view, *Curr Opin Struct Biol.* **62**, 22-30.

235. Woods, R. J. (2018) Predicting the structures of glycans, glycoproteins, and their complexes, *Chem Rev.* **118**, 8005-8024.

236. Carter, D. C. & Ho, J. X. (1994) Structure of serum albumin, *Advances in protein chemistry.* **45**, 153-203.

237. Dammer, U., Hegner, M., Anselmetti, D., Wagner, P., Dreier, M., Huber, W. & Güntherodt, H. J. (1996) Specific antigen/antibody interactions measured by force microscopy, *Biophys J.* **70**, 2437-41.

238. The, T. H. & Feltkamp, T. E. (1970) Conjugation of fluorescein isothiocyanate to antibodies. I. Experiments on the conditions of conjugation, *Immunology.* **18**, 865-73.

239. Twining, S. S. (1984) Fluorescein isothiocyanate-labeled casein assay for proteolytic enzymes, *Anal Biochem.* **143**, 30-4.

240. Barbero, N., Barolo, C. & Viscardi, G. (2016) Bovine serum albumin bioconjugation with FITC, *World Journal of Chemical Education.* **4**, 80-85.

241. Chaganti, L. K., Venkatakrishnan, N. & Bose, K. (2018) An efficient method for FITC labelling of proteins using tandem affinity purification, *Bioscience reports.* **38**.

242. Hatton, N. E., Baumann, C. G. & Fascione, M. A. (2021) Developments in mannose-based treatments for uropathogenic *Escherichia coli*-induced urinary tract infections, *Chembiochem : a European journal of chemical biology.* **22**, 613-629.

243. Mathiesen, R., Eld, H. M. S., Sørensen, J., Fuglsang, E., Lund, L. D., Taverniti, V. & Frøkiær, H. (2019) Mannan Enhances IL-12 Production by Increasing Bacterial Uptake

and Endosomal Degradation in L. acidophilus and S. aureus Stimulated Dendritic Cells, *Frontiers in Immunology.* **10**.

244. Jin, X., Zhang, M., Cao, G.-f. & Yang, Y.-f. (2019) *Saccharomyces cerevisiae* mannan induces sheep beta-defensin-1 expression via Dectin-2-Syk-p38 pathways in ovine ruminal epithelial cells, *Vet Res.* **50**, 8.

245. Wang, W., Li, Z., Han, Q., Guo, Y., Zhang, B. & D'Inca, R. (2016) Dietary live yeast and mannan-oligosaccharide supplementation attenuate intestinal inflammation and barrier dysfunction induced by *Escherichia coli* in broilers, *Br J Nutr.* **116**, 1878-1888.

246. Netea, M. G., Joosten, L. A. B., van der Meer, J. W. M., Kullberg, B.-J. & van de Veerdonk, F. L. (2015) Immune defence against *Candida* fungal infections, *Nature Reviews Immunology.* **15**, 630-642.

247. Erwig, L. P. & Gow, N. A. R. (2016) Interactions of fungal pathogens with phagocytes, *Nature Reviews Microbiology.* **14**, 163-176.

248. Brown, G. D. (2011) Innate antifungal immunity: the key role of phagocytes, *Annu Rev Immunol.* **29**, 1-21.

249. Netea, M. G., Gow, N. A., Munro, C. A., Bates, S., Collins, C., Ferwerda, G., Hobson, R. P., Bertram, G., Hughes, H. B., Jansen, T., Jacobs, L., Buurman, E. T., Gijzen, K., Williams, D. L., Torensma, R., McKinnon, A., MacCallum, D. M., Odds, F. C., Van der Meer, J. W., Brown, A. J. & Kullberg, B. J. (2006) Immune sensing of *Candida albicans* requires cooperative recognition of mannans and glucans by lectin and Toll-like receptors, *J Clin Invest.* **116**, 1642-50.

250. Saijo, S., Ikeda, S., Yamabe, K., Kakuta, S., Ishigame, H., Akitsu, A., Fujikado, N., Kusaka, T., Kubo, S., Chung, S. H., Komatsu, R., Miura, N., Adachi, Y., Ohno, N., Shibuya, K., Yamamoto, N., Kawakami, K., Yamasaki, S., Saito, T., Akira, S. & Iwakura, Y. (2010) Dectin-2 recognition of alpha-mannans and induction of Th17 cell differentiation is essential for host defense against *Candida albicans*, *Immunity.* **32**, 681-91.

251. Nguyen, T. N. Y., Padungros, P., Wongsrisupphakul, P., Sa-Ard-Iam, N., Mahanonda, R., Matangkasombut, O., Choo, M.-K. & Ritprajak, P. (2018) Cell wall mannan of *Candida krusei* mediates dendritic cell apoptosis and orchestrates Th17 polarization via TLR-2/MyD88-dependent pathway, *Scientific Reports.* **8**, 17123.

252. Al-Mughaid, H., Nawasreh, S., Naser, H., Jaradat, Y. & Al-Zoubi, R. M. (2022) Synthesis and hemagglutination inhibitory properties of mannose-tipped ligands: The effect of terminal phenyl groups and the linker between the mannose residue and the triazole moiety, *Carbohydrate Research.* **515**, 108559.

253. Young, T. D., Liau, W. T., Lee, C. K., Mellody, M., Wong, G. C. L., Kasko, A. M. & Weiss, P. S. (2020) Selective promotion of adhesion of *Shewanella oneidensis* on mannose-decorated glycopolymer surfaces, *ACS Applied Materials & Interfaces.* **12**, 35767-35781.

254. Gordo, S., Martos, V., Vilaseca, M., Menéndez, M., de Mendoza, J. & Giralt, E. (2011) On the role of flexibility in protein–ligand interactions: the example of p53 tetramerization domain, *Chemistry – An Asian Journal.* **6**, 1463-1469.

255. Tsouka, A., Hoetzel, K., Mende, M., Heidepriem, J., Paris, G., Eickelmann, S., Seeberger, P. H., Lepenies, B. & Loeffler, F. F. (2021) Probing multivalent carbohydrate-protein interactions with on-chip synthesized glycopeptides using different functionalized surfaces, *Frontiers in Chemistry.* **9**.

256. Kim, H., Hyun, J., Park, S.-H. & Shin, I. (2018) Analysis of binding properties of pathogens and toxins using multivalent glycan microarrays, *RSC Advances.* **8**, 14898-14905.

257. Touaibia, M., Wellens, A., Shiao, T. C., Wang, Q., Sirois, S., Bouckaert, J. & Roy, R. (2007) Mannosylated G(0) dendrimers with nanomolar affinities to *Escherichia coli* FimH, *ChemMedChem.* **2**, 1190-1201.

258. Granata, G., Stracquadanio, S., Consoli, G. M. L., Cafiso, V., Stefani, S. & Geraci, C. (2019) Synthesis of a calix[4]arene derivative exposing multiple units of fucose and preliminary investigation as a potential broad-spectrum antibiofilm agent, *Carbohydrate Research.* **476**, 60-64.

259. Verkhnyatskaya, S. A., Kong, C., Klostermann, C. E., Schols, H. A., de Vos, P. & Walvoort, M. T. C. (2021) Digestion, fermentation, and pathogen anti-adhesive properties of the hMO-mimic di-fucosyl-β-cyclodextrin, *Food & Function*.

260. Mohamed, A. A. A., S.; Hassan, M. H.; Hassan, O.; Lloyd, J. E.; Field, R. A.; Blanford, C. F. Fabrication of a lectin biosensor for the detection of pathogenic bacteria in

Chapter 4

Chapter 4 – Conclusions and Future Outlook

Human breast milk is widely regarded as the 'gold standard' diet for infants, providing complete postnatal nutrition during early neonate and further development, fined tuned over millions of years of evolution [2]. Contained within breast milk are a myriad of biologically active compounds that possess diverse roles aiding the development of the immune system and intestinal microbiota [4] and HMOs are one such component. They are non-digestible, complex carbohydrates [6] that exert prebiotic effects, as well as acting as anti-adhesion agents and immune system modulators [5]. However, in many cases the bioactive properties of complex, low abundance HMO structures are ill-defined. This is due to many factors, including the limited supply of human milk samples and tedious and labour-intensive purification procedures [89], extremely difficult chemical synthesis [99], and lack of a complete biosynthetic pathway to more complex structures. Nevertheless, they possess huge potential for the detection, prevention, and intervention of bacterial infection through diet or novel biosensors and therapeutics. In brief, this thesis set out to develop improved methodologies for the synthesis, characterisation, and purification of HMO analogues as well as to probe glycan-bacterial interactions via in vitro culture studies and the development of a glycan microplate-based binding assess capable of distinguishing specific interactions between plant lectins and live bacteria. This final chapter summarises the progress made during this thesis and the implications this research has for future work.

4.1 Analysis of Human Milk Oligosaccharides

There are many methods currently available for the characterisation of HMOs, principally LC methods coupled with MS for analysis of pooled HMO samples. However, in this study, LC methods were developed for characterisation of HMO mixtures where MS was used only to confirm structure identification. Mixtures of HMOs (M1 and M2) provided by DSM were used as a test case where the identity of contained compounds was known but not ratios. Following analysis by HPAEC-PAD and using commercial standards and co-injections, these mixtures were found to be incorrectly labelled. The method was able to resolve all peaks, apart from 3FL and LNFPIII. It also revealed the presence of additional compounds in M2 – Lac and 2'FL. To confirm peak identities and create a method towards the characterisation of non-derivatised HMOs and their semi-preparative purification, an improved HILIC HPLC coupled to CAD detector was developed. From the literature a

provisional linear gradient of 95% to 10% (v/v) acetonitrile over 12 min was used and developed to achieve baseline resolution of HMOs. The final gradient used was 95% to 60% (v/v) acetonitrile over 40 min. This method was able to resolve all compounds and confirmed the presence of both 3FL and LNFPIII in M1. All assignments were confirmed by MS. The misidentification of components is thought to be due to the method of synthesis used by DSM – enzymatic synthesis by α-L-fucosidases, which suffers from reversible equilibrium and relaxed regioselectivity.

The use of HILIC pre-packed cartridges was also explored for either the benchtop purification of HMOs or as a robust method of pre-purification to allow for improved subsequent semi-preparative HILIC HPLC purification. The 500 mg cartridges were able to resolve the components of both M1 and M2 at both 1 mg and 10 mg loading capacity using a linear gradient of acetonitrile to water (90% to 72% acetonitrile). Additionally, larger capacity cartridges are available, allowing for, potentially, higher resolution benchtop separations or higher glycan loading and thus larger-scale purification capacity.

The HPAEC-PAD methodology was further optimised so that it could be used for the monitoring and quantification of enzymatic fucosyltransferase reactions, where the starting material and product differ by only one fucose residue. Different CarboPac columns, gradients and eluents were tested to achieve improved resolution of substrate and product. An advantage of the method over TLC is that the conversion of the starting material to product can be quantified by integration of the resultant peaks. Additionally, the instrument requires only 5 μL of sample at micromolar range (~10 μM) for optimal resolution therefore, when sampling precious enzymatic reactions, very little is used. The instrument is non-destructive; however, it operates in high electrolyte conditions with minimal loading concentrations, therefore, recovery of samples or peaks of interest is difficult. Furthermore, this is particularly bothersome when further MS analysis are required on individual peaks. However, it should be noted that the instrument can be coupled with an in-line desalting unit and MS which would allow simultaneous reaction quantification and MS confirmation which would potentially negate the need for expensive or unobtainable commercial standards – a limitation of the technique. To give a more thorough examination of the improved methods mentioned above, they should also be tested with more complex HMO mixtures and different enzymatic reactions to ensure wider applications are also successful.

4.2 Synthesis of Novel HMO Analogues

Bacterial fucosyltransferases have been extensively used for the enzymatic synthesis of biorelevant and disease-associated oligosaccharides including fucosylated HMOs and Lewis antigens. However, much of the literature characterises the specific activity of α1,3/4-fucosyltransferases for the synthesis of Lewis blood group antibodies [115, 117-121] while the α1,2-fucosyltransferases have been consistently overlooked and understudied. In this research project, the substrate specificity of HmFucT, a previously unreported enzyme, was examined for the synthesis of complex HMO analogues by HPAEC-PAD. The enzyme is promiscuous and can accept and convert a variety of di-, tri-, tetra- and pent- saccharides, including core HMO structures and fucosylated glycans with greater than 60% efficiency in the majority of cases. The enzyme appears to have a slight preference towards type-1 glycans and cannot catalyse the fucosyltransfer to an internal Gal residue. Minimal turnover of fucosylated pentasaccharides was seen. Functionalised lactosides were synthesised to incorporate different functional groups at the reducing end terminus for probing not only the activity of the fucosyltransferase, but also to allow for further applications such as copper-catalysed click chemistry. An anomeric azide group, trifluoracetmidopropyl, aminopropyl and PEGylated azide group were used. The azide group was the most well tolerated (98%), with negligible turnover of the other lactosides (< 10%). The ability of HmFucT to catalyse the fucosylation of functionalised lactose derivatives could unlock the potential for using functionalised building blocks in a multi-enzyme cascade to chemoenzymatically synthesise novel HMO analogues with functionality for a multitude of uses, including glycan biosensors, arrays and selective inhibitors.

Due to the expense of the GDP-L-fucose sugar nucleotide donor that is required, the bifunctional enzyme FKP was examined for 1) its promiscuity and 2) as a part of a one pot reaction with the HmFucT. FKP was found to convert D-Ara and L-Gal into their respective GDP-sugar nucleotides, which enabled access to a range of novel non-natural HMO analogues with controllable chemistries at the C-5 position of the carbon ring. Furthermore, these donors can be processed by HmFucT giving access to novel non-natural HMO analogues.

In the future this work would benefit from optimising the one pot HmFucT/FKP reaction and upscaling it to produce useable quantities of functionalised HMOs and non-natural HMO analogues. These compounds could then be used for the synthesis of inhibitors, probes, biosensors and arrays to investigate the molecular nature of pathogen-glycan interactions. For example, *Campylobacter jejuni* is known to bind to α1,2-Fuc or H1

antigen [24, 58, 59], but the mechanism of this interaction is currently unknown. Site-directed mutagenesis could be employed to improve the activity of HmFucT with more complex acceptor substrates, however, the protein would first need to be crystallised and its structure active site elucidated, as it is currently unknown. Furthermore, FKP may benefit from active site mutation to improve its activity with non-natural or modified substrates.

4.3 Probing Bacterial-Glycan Interactions

The gut microbiota has been established as an important player influencing many aspects of human physiology. HMOs are known to act as prebiotics and most studies on HMO utilisation by microbes focus on the interaction of naturally occurring mixtures of HMOs with gut bacteria. As such, HMOs are of intense interest, not only to the infant formula industry, but also for use in adult health food products where they can be used to restore a beneficial microbiota in dysbiotic adults and provide health benefits. The first part of this section focussed on establishing the capabilities of wild type non-pathogenic *Escherichia coli* K12 MG1655 as a model organism to study HMO utilisation of fucosylated HMOs for pathogens of interest. *E. coli* was chosen as a model organism for this study due to its well-characterised catabolic pathways and it was found to utilise L-fucose as its sole energy source, however, above 5% (w/v) L-Fuc supplementation to the media, microbial growth was severely inhibited. This was thought to be due to the organism's inefficient carbon metabolism of L-Fuc which resulted in the accumulation of L-fuculose from ATP deficiency, a consequent of upregulation of gluconeogenesis and quorum sensing and downregulation of glycolysis. Overall, *E. coli* was not able to utilise 2'FL as a substitute to L-Fuc, nor do they effect growth when co-fed with L-Fuc. *E. coli* proved to be a reliable organism for modelling the metabolism of HMOs but only for microbes that also possess the *fuc* operon and lack fucosidase expression. Next steps would be to co-culture *E. coli* with microbes that have endogenous fucosidase expression allowing the scavenging of released L-Fuc to see if there is an effect when free fucose is released from HMOs. QCM could also be used to model bacterial biofilm formation or microbial culture dynamics in the presence or absence of HMOs, requiring only minimal quantities of HMOs.

Part II aimed to investigate the unique binding specificities of different bacterial strains. Glycan motifs have a strong and well-characterised interaction, a relatively high surface coverage and a known functioning role used by multiple different pathogenic species,

therefore, they serve as a promising target for detection of infection and targeted drug therapies. A microplate-based assay was designed, which used BSA as a protein carrier to immobilise glycans non-covalently to a standard polystyrene microplate, to probe the effect of glycan valency, linker flexibility and chain length on the selectivity of bacterial binding. To see if this method was viable, BSA-FITC was non-covalently immobilised to a microplate which resulted in maximal BSA binding at 50 µg/mL.

The assay was then validated with mannose conjugated BSA immobilised to the microplate with fluorescein-labelled plant lectins, ConA and UEA-I. The BSA glycans displayed different numbers of glycans on the surface and different mannose chain lengths. Highly valent, mannosylated polysaccharide, mannan from *Saccharomyces cerevisiae,* was also bound to the surface. All compounds elicited selective binding with the fluorescent-plant lectin, ConA, and negligible binding to UEA-I. Moreover, when BSA-conjugated 2'FL was bound, the reverse was true, where selective binding of UEA-I was seen with negligible binding to ConA. A much higher fluorescent intensity was observed for mannan versus the BSA-glycoconjugates as a result of highly branched mannosylated structure of mannan, which may allow multivalent interactions resulting in greater binding capacity of the lectin to the mannan. Furthermore, the use of BSA limits the number of mannose residues displayed in the correct orientation. Moreover, the size of the BSA limits the number of glycoconjugates that can bind to the surface.

The assay was tested with live bacteria, *E coli* ORN178 (FimH+) and ORN208 (FimH-); strains with varied FimH mannose lectin-binding domains to test the selectivity of the test. It was able to identify selective binding interactions with live bacteria, where mannan showed the greatest binding affinity. However, the assay had limited sensitivity and appeared to be limited by the size of the carrier protein, and number and flexibility of the displayed glycans, as mannan was able to elicit a much high fluorescent response, presumably due to its greater glycan valency. By modifying the method of immobilisation, i.e., eliminate the use of BSA and instead use a smaller functionalised probe with a long flexible linker for multivalent glycan display, the binding affinity of *E. coli* may be improved. Microplates with functionalised surfaces could also be used to facilitate covalent binding with functionalised glycans, e.g., an aminoreactive NHS-activated microplate surface with aminopropyl functionalised HMO analogues [261]. The methodology also lends itself for adaption to other mediums, such as gold nanoparticles, quantum dots and gold/graphene chips for use as highly specific biosensors for detection of pathogens or identification of pathogen-specific glycan motifs for the development of targeted drug therapies.

4.4 Summary

There is still a long way to go before the research conducted in this thesis can be developed into a viable targeted inhibitor, or glycan functionalised gold nanoparticles, quantum dots and gold/graphene chips for use as highly specific biosensors for detection of pathogens, and the identification of pathogen-specific glycan motifs for the development of targeted drug therapies or diet-mediated treatments. The results obtained underpin ongoing research into the development of chemoenzymatic approaches to the synthesis and analysis of biorelevant glycans and subsequent glycan innovations currently being conducted in the Field group and, in the wider sense, at the Manchester Institute of Biotechnology. Carbohydrates play an important role in disease development, prevention, diagnosis and care and further research is required to fully exploit their potential uses within therapeutics and diagnostics.

4.5 References

2. Blackburn, D. G. (1993) Lactation: historical patterns and potential for manipulation, *J Dairy Sci.* **76**, 3195-212.

4. Andreas, N. J., Kampmann, B. & Le-Doare, K. M. (2015) Human breast milk: A review on its composition and bioactivity, *Early Human Development.* **91**, 629-635.

5. Akkerman, R., Faas, M. M. & de Vos, P. (2018) Non-digestible carbohydrates in infant formula as substitution for human milk oligosaccharide functions: Effects on microbiota and gut maturation, *Crit Rev Food Sci Nutr*, 1-12.

6. Sprenger, N., Lee, L. Y., De Castro, C. A., Steenhout, P. & Thakkar, S. K. (2017) Longitudinal change of selected human milk oligosaccharides and association to infants' growth, an observatory, single center, longitudinal cohort study, *PloS one.* **12**, e0171814.

24. Newburg, D., Ruiz-Palacios, G. & Morrow, A. (2005) Human milk glycans protect infants against enteric pathogens, *Annu Rev Nutr.* **25**, 37-58.

58. Morrow, A. L., Ruiz-Palacios, G. M., Altaye, M., Jiang, X., Guerrero, M. L., Meinzen-Derr, J. K., Farkas, T., Chaturvedi, P., Pickering, L. K. & Newburg, D. S. (2004) Human milk oligosaccharides are associated with protection against diarrhea in breast-fed infants, *J Pediatr.* **145**, 297-303.

59. Morrow, A. L., Ruiz-Palacios, G. M., Altaye, M., Jiang, X., Guerrero, M. L., Meinzen-Derr, J. K., Farkas, T., Chaturvedi, P., Pickering, L. K. & Newburg, D. S. (2004) Human milk oligosaccharide blood group epitopes and innate immune protection against *Campylobacter* and calicivirus diarrhea in breastfed infants, *Protecting Infants through Human Milk.* **554**, 443-446.

89. Prudden, A. R., Barile, D., Boons, G.-J., Bode, L., Pohl, N., Contractor, N., Jennewein, S. & Jin, Y.-S. (2016) Overcoming the limited availability of human milk oligosaccharides: challenges and opportunities for research and application, *Nutr Rev.* **74**, 635-644.

99. Guo, J. & Ye, X. S. (2010) Protecting groups in carbohydrate chemistry: influence on stereoselectivity of glycosylations, *Molecules.* **15**, 7235-65.

115. Ye, J., Xia, H., Sun, N., Liu, C.-C., Sheng, A., Chi, L., Liu, X.-W., Gu, G., Wang, S.-Q., Zhao, J., Wang, P., Xiao, M., Wang, F. & Cao, H. (2019) Reprogramming the enzymatic assembly line for site-specific fucosylation, *Nature Catalysis*.

117. Dumon, C., Priem, B., Martin, S. L., Heyraud, A., Bosso, C. & Samain, E. (2001) *In vivo* fucosylation of lacto-N-neotetraose and lacto-N-neohexaose by heterologous expression of *Helicobacter pylori* alpha-1,3 fucosyltransferase in engineered *Escherichia coli*, *Glycoconjugate Journal.* **18**, 465-474.

118. Martin, S. L., Edbrooke, M. R., Hodgman, T. C., vandenEijnden, D. H. & Bird, M. I. (1997) Lewis X biosynthesis in *Helicobacter pylori* - Molecular cloning of an alpha(1,3)-fucosyltransferase gene, *J Biol Chem.* **272**, 21349-21356.

119. Tsai, T. W., Fang, J. L., Liang, C. Y., Wang, C. J., Huang, Y. T., Wang, Y. J., Li, J. Y. & Yu, C. C. (2019) Exploring the synthetic application of *Helicobacter pylori* alpha 1,3/4-fucosyltransferase FucTIII toward the syntheses of fucosylated human milk glycans and Lewis antigens, *Acs Catalysis.* **9**, 10712-10720.

120. Rasko, D. A., Wang, G., Monteiro, M. A., Palcic, M. M. & Taylor, D. E. (2000) Synthesis of mono- and di-fucosylated type I Lewis blood group antigen by *Helicobacter pylori*, *Eur J Biochem.* **267**, 6059-6066.

121. Bai, J., Wu, Z., Sugiarto, G., Gadi, M. R., Yu, H., Li, Y., Xiao, C., Ngo, A., Zhao, B., Chen, X. & Guan, W. (2019) Biochemical characterization of *Helicobacter pylori* alpha1-3-fucosyltransferase and its application in the synthesis of fucosylated human milk oligosaccharides, *Carbohydrate Research.* **480**, 1-6.

261. Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M. E., Alvarez, R., Bryan, M. C., Fazio, F., Calarese, D., Stevens, J., Razi, N., Stevens, D. J., Skehel, J. J., van Die, I., Burton, D. R., Wilson, I. A., Cummings, R., Bovin, N., Wong, C. H. & Paulson, J. C. (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins, *Proceedings of the National Academy of Sciences of the United States of America.* **101**, 17033-17038.

Chapter 5

Chapter 5 – Materials and Methods

5.1 General Methods

All chemicals were purchased from Sigma-Aldrich (UK) and used as received unless otherwise stated. Milli-Q H_2O (MQ) was used for the preparation of aqueous buffers. All reagents and solvents used for analytical applications were of analytical quality. All HMOs were provided by DSM (formerly Glycom) as part of a HMO donation program, including, for example, 2'-fucosyllactose (2'FL), 3-fucosyllactose (3FL), lacto-N-fucopentaose III (LNFPIII), lacto-*N*-neotetraose (LNnT), as well as the HMO mixtures.

5.1.1 Thin-Layer Chromatography

Analytical thin layer chromatography (TLC) was performed on Silica Gel 60 F_{254} aluminium backed sheets (Merck). Carbohydrates were detected by staining with sulphuric acid reagent (5% sulphuric acid in methanol) and developed by charring.

5.1.2 Mass Spectrometry

Low-resolution electrospray ionisation mass spectrometry (ESI MS) was performed using an Advion expression L Compact Mass Spectrometer. Prior to injection, samples were filtered through a 0.22 μm membrane disk filters. Data was recorded in both positive and negative ionisation modes and processed using Advion Mass Express software.

High resolution, accurate mass spectrometry (HR-MS) was performed using an Agilent 1290 infiniti series LC system, coupled to an Agilent 6560 QTOF mass spectrometer, ESI, QTOF Only, Positive (or Negative) mode. 5ul of sample was flow injected into 50% (v/v) acetonitrile with 0.1% (v/v) formic acid. The data was analysed using Agilent MassHunter software.

5.1.3 HPAEC-PAD Analysis

High-performance anion-exchange chromatography coupled with pulsed electrochemical detection (HPAEC-PAD) analysis was performed on a Dionex IC system equipped with a CarboPac anion-exchange column and guard cartridge, and an electrochemical detector

to register peaks and determine their retention times. Flow rates remained constant as 0.25 mL/min for all analysis. Detector settings were as follows: carbohydrates Gold quad waveform, PdH reference electrode, Int cell mode: on, ED_1 and ED_1_Total channels, a data collection rate of 20 Hz and pH 10-13. Data was analysed using Chromeleon™ (all Thermo Fisher Scientific).

5.1.4 Gel Permeation Chromatography

All gel permeation chromatography was carried out using a Gilson system, equipped with a Precision Instruments IOTA 2 refractive index detector and a fraction collector, using as stationary phase a Toyopearl TSK-HW40S column (1.6 cm x 90 cm) at room temperature (RT) in water with flow rate 0.5 mL/min.

5.1.5 Nuclear Magnetic Resonance Spectroscopy

NMR spectra were recorded on Bruker Avance III 400 MHz spectrometer equipped with a broadband BBFO probe or on Bruker Neo 600 MHz spectrometer equipped with a TCI cryoprobe. Chemical shifts (δ) are reported in parts per million (ppm) using the water peak at 4.79 ppm for referencing.

5.1.6 Colorimetric, Fluorescent and Cell Density Assays

A BMG Labtech FLUOStar Omega microplate reader equipped with suitable absorbance filters was used for all assays. All data is presented blank corrected, as an average across replicates. Colorimetric assays were performed in NUNC 96 plates, cell density assays were performed used sterile COSTAR 96 plates and fluorescent assays were performed using black, flat bottom, 384-well µCLEAR® microplates (Greiner Bio-One).

5.1.7 HPLC Analysis

All HPLC was performed using the UltiMate™ 3000 Binary semi-preparative system (Thermo Fisher Scientific) with an in-line Diode-Array Detection (DAD) and coupled to a charged aerosol detector (CAD) (Thermo Fisher Scientific) via an analytical adjustable flow splitter (Analytical Scientific Instruments US). Data was analysed using Chromeleon™ (Thermo Fisher Scientific).

5.1.8 SDS-PAGE Analysis

Proteins were analysed by SDS-PAGE using RunBlue precast 4-20% (v/v) gradient or 8% (w/v) polyacrylamide gels (Expedeon). Before analysis, protein samples (10 µL) were mixed with Laemmli buffer (10 µL, BioRad) and boiled at 95 °C for 10 mins. Samples were allowed to cool, centrifuged briefly and then loaded on to SDS-PAGE gels as 10 µL samples. All SDS-PAGE analysis was performed alongside a Precision Plus Protein™ Dual Color Ladder (5 µL, BioRad) and run at 200 V for 50 mins in 1x RunBlue running buffer (Expedeon). Gels were stained with Coomassie Brilliant Blue Protein Stain (Sigma) for a minimum of 2 h to visualise proteins. Gels were imaged using a G:Box F3 Gel Imaging System (Syngene).

5.2 Chapter 2 Experimental

5.2.1. Analysis of HMO Mixtures by LC-MS

LC-MS analysis of HMO mixtures was carried out using Thermo Q-Exactive Plus/Dionex Ultimate 3000 (Thermo Fisher Scientific) with a SeQuant ZIC-cHILIC, 3 µm, 100 Å (100 x 2.1 mm) (Merck) column. Samples of 150 µM were injected (5 µL) via autosampler. Buffer A = acetonitrile with 0.1% (v/v) formic acid, buffer B = MQ. Flow rate was 0.6 mL/min. T_0 = 99% B; T_1 = 0% B; $T_{0.6}$ = 99% B; $T_{1.6}$ = 99% B; T_6 = 1% B; $T_{7.5}$ = 1% B; $T_{7.51}$ = 1% B. Data was analysed using Xcalibur software (Thermo Fisher Scientific).

5.2.2 Analysis of HMO Mixtures by HPAEC-PAD

Samples (standards 100 µM, HMO1 200 µM and HMO2 400 µM in MQ) were injected (5 µL) via autosampler to a CarboPac™ PA100, 8.5 µm (2 x 250 mm) (Thermo Fisher Scientific) and eluted by a multi-step gradient at 0.2 mL/min: T_0 = 0% B; T_1 = 0% B; T_{20} = 50% B; T_{25} = 50% B; T_{30} = 0% B; T_{40} = 0% B. Buffer A = 150 mM NaOH ; buffer B = 150 mM NaOH with $_{600}$ mM NaOAc and buffer C = MQ. Samples for co-injection were prepared by spiking the HMO mixtures (200 or 400 µM) with a standard (100 µM) before injection. Data was analysed using Chromeleon™ Chromatography Data System (CDS) software (Thermo Fisher Scientific).

5.2.3 Analysis of HMO Mixtures by HILIC HPLC

Sample was injected (50 µL) manually at total carbohydrate concentration of 20 mg/mL to a Luna 5 µm HILIC 200 Å (250 x 10 mm) (Phenomenex) column. The flow rate was 5 mL/min using water as buffer A and acetonitrile as buffer B, both HPLC-grade. The method was optimised as follows:

1. T_0 = 95% B; $T_{0.6}$ = 95% B; $T1_{.6}$ = 95% B; T_7 = 10% B; $T_{7.05}$ = 95% B; T_{12} = 95%
2. T_0 = 95% B, $T_{1.5}$ = 95% B; T_{34} = 10% B; T_{35} = 95% B; T_{40} = 95% B
3. T_0 = 95% B; $T_{1.5}$ = 95% B; T_{34} = 50% B; $T_{34.5}$ = 95% B; T_{40} = 95% B
4. T_0 = 95% B; $T_{1.5}$ = 95% B; T_{34} = 60% B; $T_{34.5}$ = 95% B; T_{40} = 95% B

Following optimisation, all analyses were carried out using the multistep gradient (4) of: T_0 = 95% B; $T_{1.5}$ = 95% B; T_{34} = 60% B; $T_{34.5}$ = 95% B; T_{40} = 95% B. Fractions were collected manually, and eluent evaporated to dryness with centrifugal evaporation (Genevac EZ-2) for further analysis. Data was analysed using Chromeleon™ Chromatography Data System (CDS) software (Thermo Fisher Scientific).

5.2.4 Purification of HMO Mixtures with HILIC Cartridges

For benchtop separation of HMO mixtures, iSPE HILIC, 50 µm (6 mL, 500 mg) 50 µm particles column cartridges (HILICON) were used. Samples of 20 mg/mL or 200 mg/mL (50 µL) were applied to the column. Fractions were eluted with a gradient of 90% to 70% (v/v) acetonitrile by gravity and fractions (500 µL) were collected dropwise. Eluted fractions were analysed using TLC with butanol/acetic acid/water (2:1:1, v/v).

### 5.2.5 Synthesis of 4-*O*-(β-D-galactopyranosyl)-β-D-glucopyranosyl azide (12)
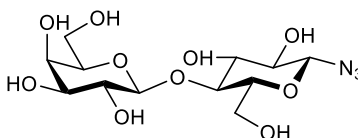
12

Anomeric azide was introduced following Shoda's methodology [174]. 2-Chloro-1,3-dimethylimidazolinium chloride [DMC (24.7 mg, 146 µmol, 5 eq.)] was added to a lactose (10 mg, 29.2 µmol), DIPEA (76.6 µL, 438 µmol, 15 eq.) and NaN_3 (19 mg, 292 µmol, 10

eq.) in MQ (0.5 mL), and the reaction mixture was stirred for 36 h at RT. After 36 h the reaction was analysed by TLC and 50% conversion of lactose to 1-azido lactose (**12**) was observed. More DMC (2 eq.), NaN_3 (5 eq.) and DIPEA (7.5 eq.) was added to the reaction mixture. The reaction mixture was left for a further 36 h, stirring at RT. The reaction mixture was purified by gel permeation chromatography. Fractions containing carbohydrates were identified by TLC and staining with orcinol and then pooled. Lyophilisation gave a crystalline solid of *1-azido lactose* (**12**) (10.1 mg, 27.5 µmol, 94%). ^{1}H NMR (500 MHz, D_2O) δ 4.76 (1H, d, H1, $J_{1,2}$ = 8.6 Hz), 4.45 (1H, d, H1', $J_{1',2'}$ = 8.1 Hz,), 3.98 (1H, dd, H6a, $J_{6a,6b}$ = 12.4 Hz, $J_{5,6a}$ = 1.3 Hz), 3.92 (1H, d, H4', $J_{3',4'}$ = 3.2 Hz,), 3.87 – 3.81 (1H, m, H6b), 3.81 – 3.74 (2H, m, H6a', H6b'), 3.75 – 3.64 (5H, m, H3, H4, H5, H3', H5'), 3.54 (1H, dd, H2', $J_{1',2'}$ = 9.0 Hz, $J_{2',3'}$ = 8.9 Hz), 3.31 (1H, t, H2, $J_{1,2}$ = J_{2-3} = 8.7 Hz). ^{13}C NMR (126 MHz, D_2O) δ 102.9 (C1), 90.0 (C1'), 77.7, 76.7, 75.4, 74.4 (C3, C4, C5, C5'), 72.6, 72.5 (C2, C3'), 71.0 (C2'), 68.6 (C4'), 61.1 (C6'), 59.9 (C6). ESI-MS; calculated $C_{12}H_{21}N_3O_{10}$ $[M+Na]^+$: 390.12; found: 390.10.

5.2.6 Synthesis of 3-trifluoracetamidopropyl-4-*O*-(β-D-galactopyranosyl)-β-D-glucopyranoside (14)


14
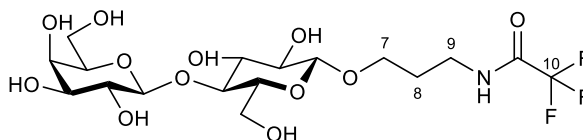
3-Trifluoracetamidopropyl lactoside (**14**) was chemically synthesised by acylation of 3-aminopropyl β-D-lactose with ethyl trifluoroacetate based on a protocol from Lau, *et al.,* 2010 [262]. 3-Aminopropyl lactose (**13**) (100 mg, 0.25 mmol, 1 eq.) and sodium carbonate (53 mg, 0.5 mmol, 2 eq.) were added to a round bottom flask under nitrogen. Dry methanol (1 mL) was added and left to equilibrate for several minutes before ethyl trifluoroacetate (71 mg, 1.194 g/mL at 25 °C, 0.5 mmol, 2 eq.) was added and the reaction was left to incubate overnight under nitrogen stirring at RT. The reaction was followed by TLC (DCM/MeOH, 1:1 v/v). The crude reaction was first purified by ENVI™-Carb 2g charcoal column (Sigma Aldrich) to remove excess ethyl trifluoracetate. The column was first conditioned with 1 CV of MQ followed by 1 CV of ACN and a further 1 CV of MQ. Following

sample loading fractions (5 mL) were then collected for washes consisting of 3 x 5 mL MQ and 3 x 5 mL ACN:MQ (50:50, v/v). Fraction containing the product were then combined and lyophilised, recovering 200 mg of crude compound.

Impure compound was further purified by Biotage SP4. It was dissolved in 2 mL MQ and 100 mg fine silica powder (Fluorochem, 35-70 μM) was added before lyophilisation. A further 1 mL MQ and 100 mg of silica was added and it was re-lyophilised. An empty 10g SNAP cartridge was filled with 35-70 μM silica gel and the sample was dry loaded onto the top of the column. The Biotage was primed with DCM and the sample was eluted across a gradient of 0 – 20% (v/v) MeOH for 5 CV, followed by 20% (v/v) MeOH for 20 CV at a flow rate of 15 mL min^{-1} with 1 CV waste. Fractions of 9 mL were collected and analysed by TLC (DCM/MeOH, 7:3, v/v and stained with orcinol). Fractions containing pure compound were lyophilised to give a white solid (25 mg) at a yield of 20% after purification.

^{1}H NMR (600 MHz, D$_2$O) δ 4.40 (1H, d, H1, $J_{1,2}$ = 8.0 Hz,), 4.37 (1H, d, H1', $J_{1',2'}$ = 7.9 Hz), 3.90 (2H, m, H6a, H7a), 3.85 (1H, dd, H4', $J_{3',4'}$ = 3.4, 1.0 Hz), 3.76 – 3.62 (5H, m, H5', H6b, H6a', H6b', H7b), 3.61 – 3.54 (3H, m, H3, H3', H4), 3.51 (1H, m, H5), 3.47 (1H, dd, H2', $J_{2',3'}$ = 10.0, $J_{1',2'}$ = 7.8 Hz,), 3.37 (2H, t, H9a, H9b, J = 6.9 Hzf), 3.24 (1H, dd, H2, $J_{2,3}$ = 9.0 Hz, J_{1-2} = 7.6 Hz), 1.84 (2H, q, H8a, H8b, J = 6.5 Hz). ^{13}C NMR (151 MHz, D$_2$O) δ 102.9 (C1'), 102.0 (C1), 78.4 (C4), 75.3 (C5'), 74.8 (C5), 74.4 (C3'), 72.8 (C2), 72.5 (C3), 70.9 (C2'), 68.5 (C4'), 67.6 (C7), 61.0 (C6'), 60.1 (C6), 36.8 (C9), 27.8 (C8). ESI-MS; calculated C17H28F3NO12 [M+Na]$^{+}$: 518.1564; found: 518.1466 (∆ = 1.9 ppm).

5.2.7 [3-amidopropyl-((2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethoxy)ethyl)]-4-*O*-(β-D-galactopyranosyl)-β-D-glucopyranoside (15)

15
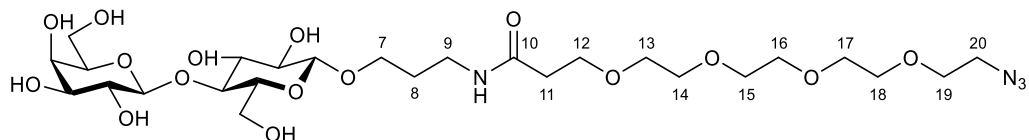
1-Azido-PEG$_4$-3-aminopropyl lactoside (**15**) was synthesised by coupling 3-aminopropyl lactose (**13**) with an azido-PEG$_4$-N-Hydroxysuccinimide (NHS) ester. The azido-PEG$_4$-NHS ester (Jena Bioscience GmbH) was first dissolved to 100 mg/mL in DMF. 3-Aminopropyl lactose (**13**) (20 mg, 50 μmol, 1 eq.) was dissolved in 200 μL MQ, azido-

PEG$_4$-NHS ester was added (25 mg, 64 µmol, 1.3 eq.) to give a total volume of 0.5 mL and the reaction was incubated, shaking, overnight at RT. The reaction was monitored by TLC and stained with orcinol and ninhydrin until it went to completion.

The crude reaction was lyophilised, syringe filtered (PTFE 0.45 µM, GE Healthcare Life Sciences) and purified by Toyopearl HW-40S gel permeation chromatography (GPC) (110 cm x 22 mm). Fractions were monitored by TLC and those containing the purified product were combined and lyophilised to yield 25.8 mg (77 %) of colourless solid. ^{1}H NMR (500 MHz, D$_2$O) δ 4.46 (1H, d, H1, $J_{1,2}$ = 8.0 Hz), 4.43 (1H, d, H1', $J_{1',2'}$ = 7.8 Hz), 3.99 – 3.93 (2H, m, H6a, H7a), 3.91 (1H, d, H4', $J_{3',4'}$ = 3.4 Hz), 3.82 – 3.74 (4H, m, H6a', H6b', H19a, H19b), 3.73 – 3.62 (18H, m, H5, H6b, H7b, H5', H13-H19), 3.65 – 3.62 (2H, m, H3, H3'), 3.53 (1H, dd, H2', $J_{1',2'}$ = 7.8 Hz, $J_{2',3'}$ = 10.0), 3.51 – 3.48 (2H, t, H12a, H12b, $J_{12,13}$ = 5.0 Hz), 3.30 (3H, m, H2, H9a, H9b), 2.51 (2H, t, H20a, H20b, $J_{19,20}$ = 6.1 Hz), 1.83 (2H, p, H8a, H8b, $J_{7,8}$ = $J_{8,9}$ = 6.6 Hz). ^{13}C NMR (126 MHz, D$_2$O) δ 174.0 (CO, -CONH-), 102.9 (C1'), 102.0 (C1), 78.4 (C4), 75.3 (C5'), 74.7 (C5), 74.3 (C3'), 72.8 (C2), 72.5 (C3), 70.9 (C2'), 69.6, 69.6, 69.5, 69.5, 69.5, 69.5, 69.2 (C13-C19), 68.5 (C4'), 67.7 (C6), 66.7 (C20), 61.0 (C6'), 60.0 (C7), 50.1 (C12), 36.2 (C9), 36.0 (C21), 28.4 (C8). ESI-MS; calculated C$_{26}$H$_{48}$N$_4$O$_{16}$ [M+H]$^+$ 673.2065, [M+Na]$^+$ 695.3065; found 673.3138 (Δ = 0.7 ppm) and 695.2958 (Δ = 0.1 ppm), respectively.

5.2.8 Acceptor Substrate Specificity of *H. mustelae* α1,2-FucT

All acceptors were first purified by HILIC HPLC to remove impurities. Reaction mixture contained 50 mM HEPES (pH 7.0), 5 mM acceptor, 5 mM GDP-L-fucose disodium salt (Carbosynth), 10 mM MgCl$_2$, alkaline phosphatase (1U) and α1,2-fucosyltransferase (HmFucT) (0.01U, Chemily Glycoscience) to give a total volume of 15 µL. The reaction mixture was incubated at 37 °C overnight with shaking (225 rpm) and monitored by TLC analysis. TLC solvent system was isopropan-2-ol/ammonium hydroxide/water (6:3:1, v/v) and stained with 5% (v/v) sulphuric acid in ethanol before charring.

Typical enzymatic synthesis work-up involved the addition of methanol (1:1) to quench the reaction. The insoluble material was removed by centrifugation (14000 rpm for 2 mins) and the supernatant was removed and concentrated by evaporation (Genevac EZ-2) to dryness. The crude product was re-suspended in MQ for HPAEC-PAD analysis. Samples (50 µM) were injected (5 µL) via autosampler. To improve resolution of reaction components, the method was developed as follows:

1. CarboPac^{TM} PA100, 8.5 μm (2 x 250 mm) (Thermo Fisher Scientific) with a multistep gradient of buffer A (150 mM NaOH) and B (150 mM NaOH with 600 mM NaOAc) ; T_0 = 0% B; T_1 = 0% B; T_{40} = 100% B; T_{45} = 100% B; T_{50} = 0% B; T_{60} = 0% B.
2. CarboPac^{TM} PA100, 8.5 μm (2 x 250 mm) (Thermo Fisher Scientific) with an extended multistep gradient of buffer A (150 mM NaOH) and B (150 mM NaOH with $_{600}$ mM NaOAc) ; T_0 = 0% B; T_1 = 0% B; T_{40} = 30% B; $T_{40.5}$ = 100% B; T_{45} = 100% B; $T_{45.5}$ = 0% B; T_{55} = 0% B.
3. CarboPac^{TM} PA100, 8.5 μm (2 x 250 mm) (Thermo Fisher Scientific) with isocratic 50 mM NaOH.
4. CarboPac^{TM} PA1, 10 μm (2 x 250 mm) (Thermo Fisher Scientific) with isocratic 50 mM NaOH.
5. CarboPac^{TM} PA20, 6.5 μm (3 x 150 mm) (Thermo Fisher Scientific) with isocratic 50 mM NaOH.

Following optimisation, method 5 was used for all HPAEC-PAD analyses of reaction products. Co-injections were carried out to identify the starting material and product peaks. All reactions were spiked with a relevant acceptor (50 μM).

Functionalised lactosides were analysed by LC-MS. All analysis was conducted on a QExactive Plus equipped with an Ultimate 3000 UHPLC (Thermo, UK). The UHPLC was equipped with a ZIC®-cHILIC column (C18 -2.1 mm x 100 mm; 3.0 μm particle size). The solvents employed were (A) water + 0.1% (v/v) formic acid and (B) acetonitrile + 0.1% (v/v) formic acid. The flow was 300 μL/min and the gradient was programmed to equilibrate at 99% B for 2 min followed by a linear gradient to 60% B over 8 min and a further linear gradient to 5% B over 5 minutes before returning to 99% B for 2 min. The column was maintained at 40 °C and the samples chilled in the autosampler at 10 °C. A sample volume of 5 μL was injected onto the column. Blank injections were analysed at the start and end of the analytical batch to assess the background and carryover. Data acquisition was conducted in full MS mode in the scan range of 90-1350 m/z with a resolution of 70,000, an AGC target of 3e^6 and a maximum integration time of 200 ms. The samples were analysed in positive and negative mode in separate acquisitions. Data was analysed by identifying the retention from extract total ion chromatograms using Xcalibur (Thermo Fisher Scientific). The raw data was then process using TidyMass [175], R script that extracts m/z values, retention times and relative areas from which starting material and product masses could be quantified.

5.2.9 Isomer Identification by β1,3/4-Galactosidase Assay

First the reactions from 5.2.8 were purified by HILIC HPLC (as stated in 5.2.3) to recover pure fucosylated product. Sample was injected (50 µL) manually. The flow rate was 5 mL/min using water as buffer A and acetonitrile as buffer B, both HPLC-grade and detector settings used were recommended values. The method consisted of a multistep gradient outlined as follows: T_0 = 95% B; $T_{1.5}$ = 95% B; T_{34} = 60% B; $T_{34.5}$ = 95% B; T_{40} = 95% B. Products were eluted and collected manually, and data was analysed using Chromeleon™ (Thermo Fisher Scientific). Following lyophilisation, fractions were analysed by LR-MS to identify fucosylated reaction products. Samples were resuspended in MQ (1 mg/mL) and the β1,3/4-galactosidase assay was carried out as follows using the β1,3/4-galactosidase kit (New England Biolabs). The reaction contained acceptor (5 µg), X1 Glycobuffer 4 and β1,3/4-galactosidase (8U) to give a total volume of 16 µL. Lactose was used a positive control and 2'FL as a negative control. Following ON incubation at 37°C, methanol (1:1) was used to quench the reaction and insoluble material was removed by centrifugation (14000 rpm for 2 mins). The supernatant was removed and concentrated by evaporation (Genevac EZ-2) to dryness. The crude reaction was re-suspended in MQ for HPAEC-PAD analysis. Samples (50 µM) were injected (5 µL) via autosampler. Reactions were analysed by HPAEC-PAD following resuspension in MQ (100 µM) using CarboPac™ PA20, 6.5 µm (3 x 150 mm) (Thermo Fisher Scientific) with isocratic elution in 50 mM NaOH at 0.2 mL/min.

Samples were also analysed by LC-MS, using Thermo Q-Exactive Plus/Dionex Ultimate 3000 (Thermo Fisher Scientific) with a SeQuant ZIC-cHILIC, (C18 -2.1 mm x 100 mm; 3.0 µm particle size) column as above (see section 5.2.8). Data was analysed using Xcalibur software (Thermo Fisher Scientific).

5.2.10 FKP Expression and Purification

FKP was expressed and purified following the protocol outlined in Wang *et al.,* 2009 [146]. Clonal populations of BL21(DE3) cells harbouring the FKP-encoding plasmid were incubated in 1 L culture of 2XYT broth with carbenicillin (100 µg/mL) with shaking at 37 °C until OD_{600} = ≤1, at which time FKP expression was induced with 500 µM isopropyl β-D-1-thiogalactopyranoside (IPTG). The temperature was lowered to 25 °C. After 12–16 h, cells were harvested and resuspended in 40 mL of lysis buffer (20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 10 mM imidazole, 2 units/mL DNase, 1 Complete EDTA-free protease inhibitor mixture tablet) per litre of culture and lysed by sonication (Vibra-Cell™ Ultrasonic liquid

processor). Cell debris was removed by centrifugation (15,000 X *g*, 4 °C, 15 min) and the supernatant filtered (Corning, 0.2 µM syringe filter). Proteins in the supernatant were purified at 4 °C using an ÄKTA pure FPLC system (GE Healthcare). The supernatant was passed through a HisTrap™ HP column (5 ml, GE healthcare), pre-equilibrated with buffer A (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 20 mM imidazole). Unbound proteins were washed with five column volumes (CV) of buffer A, followed by elution with buffer B (50 mM Tris-HCl, pH 8, 500 mM NaCl, 500 mM imidazole). Protein fractions were collected, combined and dialysed overnight in Q_A buffer (20 mM Tris-HCl (pH 7.5), 20 mM NaCl). FPK was confirmed and assessed for purity by SDS-PAGE gel (Criterion, Bio-Rad) and the concentration of the enzyme was determined by Bradford assay (~ 12 mg protein). The purified enzyme was stored in Q_A buffer at -80 °C.
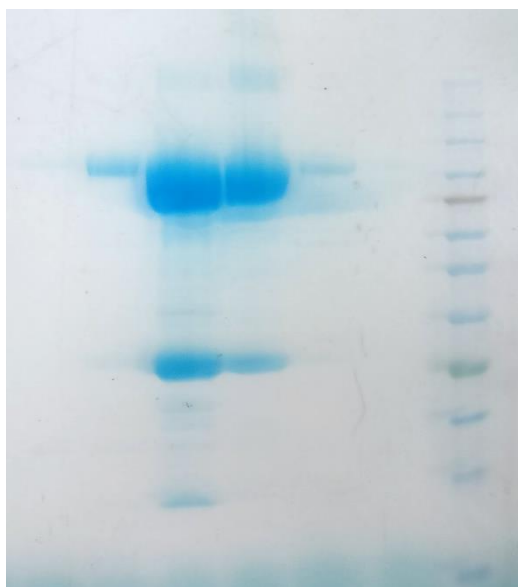
Figure 5.1: SDS-PAGE analysis of FKP following purification. Expected MW = 97 kDa. Bands at ~25 kDa are likely monomers of the FKP tetramer.

5.2.11 Synthesis of GDP-sugar nucleotides

GDP-sugar nucleotides (L-Fuc, D-Ara and L-Gal) were synthesised using the FKP enzyme. Reaction mixture contained 50 mM Tris-HCl (pH 7.5), 5 mM acceptor, 5 mM ATP, 5 mM GTP, 5 mM $MnCl_2$, inorganic pyrophosphatase (0.25 units, lyophilised form containing $MgCl_2$) and 200 µL FKP (1 mg/ml) to give a total volume of 1 mL (10 x 100 µL). The reaction mixture was incubated at 37 °C overnight to 48 h with shaking (225 rpm) and monitored by TLC analysis. TLC solvent system was isopropan-2-ol/ammonium hydroxide/water (6:3:1, v/v) and stained with 5% (v/v) sulphuric acid in ethanol before charring. Methanol (1:1) was used to quench the reaction. The insoluble material was

removed by centrifugation (14000 rpm for 2 min) and the supernatant was removed and concentrated by Genevac to dryness. The crude product was re-suspended in MQ (500 μL).

Further analysis was carried out by strong anion-exchange chromatography (SAX) performed with a Dionex Ulitmate 3000 HLPC (Thermo Fisher Scientific), equipped with an Ultimate 3000 variable wavelength detector, using a PerSeptive Biosystems Poros HQ 50 strong anion-exchange column (10 x 50 mm) (Thermo Fisher Scientific). Buffer A = 5 mM ammonium carbonate and buffer B = 250 mM ammonium carbonate. The column was first equilibrated with buffer A for 3 min, followed by linear gradient of ammonium bicarbonate at a flow rate of 8 ml/min and detection with an on-line detector to monitor A_{265nm}. The method was as follows: $T_0 = 0\%$ B; $T_3 = 0\%$ B; $T_{11} = 100\%$ B; $T_{14} = 100\%$ B; $T_{16} = 0\%$ B; $T_{18} = 0\%$ B and the sample was injected (50 μL) manually. Data was analysed using Chromeleon™ (Thermo Fisher Scientific). GDP-sugar nucleotides were confirmed by HR-MS of reaction mixtures. GDP-L-Fucose: ESI-MS; calculated $C_{17}H_{26}N_4O_{15}P_2$ [M-H]⁻: 588.0870; found: 588.0792. GDP-D-Arabinose: ESI-MS; calculated $C_{15}H_{23}N_5O_{15}P_2$ [M-H]⁻: 574.0667; found: 574.0637. GDP-L-Galactose: ESI-MS; calculated $C_{17}H_{25}N_5O_{16}P_2$ [M-H]⁻: 604.0772; found: 604.0790.

5.2.12 Synthesis of HMO Analogues by One Pot Enzymatic Reaction with FKP and α1,2-FucT from *H. mustelae*

HMO analogues were synthesised from the respective sugar nucleotide donors as a product of the FKP reaction to which α1,2-fucosyltransferase (HmFucT) enzyme, *E. coli* recombinant from *Helicobacter mustelae* (Chemily Glycoscience), with lactose as an acceptor. To the 20 μL of the reaction mixture detailed in 5.2.11; 5 mM acceptor, alkaline phosphatase (1U) and α1,2-fucosyltransferase (0.02U, Chemily Glycoscience) to give a total volume of 25 μL. The reaction mixture was incubated at 37 °C overnight with shaking (225 rpm) and monitored by TLC analysis. TLC solvent system was isopropan-2-ol/ammonium hydroxide/water (6:3:1, v/v) and stained with 5% (v/v) sulphuric acid in ethanol before charring. Methanol (1:1, v/v) was used to quench the reaction. The insoluble material was removed by centrifugation (14000 rpm for 2 mins) and the supernatant was removed and concentrated by Genevac to dryness. The crude product was re-suspended in MQ (100 μM) and analysed by HPAEC-PAD, as above, with CarboPac™ PA20, 6.5 μm (3 x 150 mm) (Thermo Fisher Scientific) with isocratic 50 mM NaOH. Product masses were confirmed by LC-MS.

5.3 Chapter 3 Experimental

5.3.1 HMO Growth Assays

E. coli K12 MG1655 were kindly provided by Dr Elizabeth Watton (Manchester Institute of Biotechnology, Manchester, UK). First, starter cultures (10 mL) of *E. coli* K12 MG1655 in LB were incubated overnight with shaking at 37 °C. The cultures centrifuged at 4,000 rpm for 10 min, the supernatant was discarded and the cell pellet was washed twice with PBS before re-suspending (10 mL) in M9 minimal media (93.0 mM sodium, 22.1 mM potassium, 18.7 mM ammonium, 1.0 mM calcium, 0.1 mM magnesium, 29.2 mM chloride, 0.1 mM sulphate, 42.2 mM phosphate) [263] supplemented with either 2% (w/v) Glc or 2% (w/v) Fuc and incubated overnight with shaking at 37 °C. Samples were repeated in quadruple. In a 96-well microplate, wells containing 250 uL M9 minimal media supplemented with either 2% (w/v) Glc or Fuc were inoculated with 1 mL starter culture and covered with an adhesive film (Microseal 'B' PCR Plate Sealing Film, Bio-Rad). Samples were incubated at 37 °C and optical density (OD_{600}) readings were recorded every 30 min for 16 or 36 h. Before each cycle, the plate underwent shaking at 100 rpm for 30 sec and the readings were taken measuring a spiral average (2 mm) with path length correction (250 µL). All data is presented blank corrected, as an average across replicates.

For co-feeding experiments, *E. coli* K12 MG1655 cultures were started as above. In a 96-well microplate, wells containing 250 uL M9 minimal media supplemented with L-Fuc (0%, 1.25% and 5%, w/v) and/or 2'FL (0%, 0.25%, 0.5%, 1% and 2%, w/v) were inoculated with 1 mL starter culture and covered with an adhesive film (Microseal 'B' PCR Plate Sealing Film, Bio-Rad). All samples were repeated in duplicate. Samples were incubated at 37 °C and optical density (OD_{600}) readings were recorded every 30 min for 36 h. Before each cycle, the plate underwent shaking at 100 rpm for 30 s and the readings were taken measuring a spiral average (2 mm) with path length correction (250 µL). All data is presented blank corrected, as an average across replicates.

5.3.3 Lectin Binding Assays

5.3.3.1 BSA-FITC Binding

BSA-FITC was kindly provided by Dr Irina Ivanova (John Innes Centre, Norwich, UK). BSA-FITC (1 mg/mL in PBS) was serially diluted (1:2) in bicarbonate buffer (50 mM, pH 10) from 400 µg/mL to 5 µg/mL. BSA-FITC sample (20 µL) was added in triplicate and incubated for 1 h, shaking (225 rpm) at 37 °C. Wells were washed three times with MQ

and dried with compressed air. Emitted fluorescence intensity was measured with a 485/520 nm filter, gain = 2000. All data is presented blank corrected, as an average across replicates.

5.3.3.2 Fluorescent Plant Lectins Binding Assay

BSA-mannose and BSA-2'FL conjugates were kindly provided by Dr Irina Ivanova and Dr Rositsa Karamanska (John Innes Centre, Norwich, UK). First, the optimal concentration of fluorescent plant lectin was established. BSA-PEG$_4$-Man (10) BSA-Man/mannan + lectin (Table 3.1, Compound 10) was diluted in bicarbonate buffer (50 mM, pH 10) from 400 µg/mL to 50 µg/mL. In a 384-well µCLEAR® microplate (black, flat bottom, Greiner Bio-One) 20 µL of each BSA-FITC sample was added in triplicate and incubated for 1 h shaking (225 rpm) at 37 °C. Wells were washed three times with MQ and dried with compressed air. Wells were blocked with 3% BSA overnight, shaking (225 rpm) at RT before being washed three times with MQ and dried with compressed air. For each concentration of BSA-PEG$_4$-Man (10), 2 µg/mL, 1 µg/mL or 0.5 µg/mL fluorescein-ConA (20 µL, Vector Labs), or 2 µg/mL fluorescein-UEA-I (20µL, Vector Labs) as a control, were incubated in binding buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 4 mM Ca^{2+} (4 mM Mn^{2+}, ConA only)) for 1 h, shaking (225 rpm) at 37 °C. They were then washed three times with MQ and dried with compressed air. Emitted fluorescence intensity was measured with a 485/520 nm filter, gain = 2000. All data is presented blank corrected, as an average across replicates.

With the method optimised, the same procedure was carried out on compounds 1-12 (Table 3.1). Glycans (BSA-mannose conjugates = 50 µg/mL, 20 µL and mannan *Saccharomyces cerevisiae* (Sigma Aldrich) = 1000 – 400 µg/mL, 20 µL), including BSA-FITC as a positive control, were physiosorbed in triplicate in wells of a 384-well microplate and incubated for 1 h, shaking (225 rpm) at 37 °C. Wells were washed three times with MQ and dried with compressed air. Wells were blocked with 3% BSA overnight, shaking (225 rpm) at RT before being washed three times with MQ and dried with compressed air. 2 µg/mL fluorescein-ConA (20 µL) in binding buffer was then added to wells and incubated for 1 h, shaking (225 rpm) at 37 °C. Emitted fluorescence intensity was measured with a 485/520 nm filter, gain = 2000. All data is presented blank corrected, as an average across replicates. Negative controls were mannan, BSA-Man, ConA and BSA-PEG$_4$-Man (10) with UEA-I.

5.3.4 Live Bacteria Glycan Arrays

E. coli strains ORN178 and ORN208 were provided by Professor P. E. Orndorff (North Carolina State University, USA). First, fluorescent labelling of *E. coli* strains ORN178 (WT, FimH+) and ORN208 (mutant, FimH-) was optimised and three different staining methods were reviewed. In all cases, ORN178 and ORN208 were first cultured in LB (10 mL) overnight at 37 °C, shaking (180 rpm). The cultures centrifuged at 4,000 rpm for 10 min, the supernatant was discarded, and the cell pellet was washed twice with PBS. Below details the methods used for each of the labelling methods:

1. Basic Fluorescein Isothiocyanate (FITC) Labelling

The cell pellet was re-suspended in PBS to OD_{600} = 1. To diluted cells (1 mL), 25 µL FITC (4 mg/mL in ethanol) was added, covered and incubated for 30 min, shaking (400 rpm) at 37 °C, protected from light. Excess FITC was then removed by centrifugation (5 mins at 10,000 rcf); the supernatant was discarded, and the cell pellet washed three times with 1 mL PBS. When this method was further optimised, the FITC staining step was repeated a further one or two times. Finally, following the centrifugation of the last labelling step, cells were re-suspended in *E. coli* binding buffer (PBS containing 1 mM $CaCl_2$ and 1 mM $MnCl_2$).

Fluorescence Microscopy:

Fluorescence microscopy was carried out using an LED Snapshot Widefield microscope (Zeiss Axioplan 2). FITC-labelled ORN178 and ORN208 cells were mounted using agarose pads. Agarose pads were made using 1% (w/v) low-melt agarose solution in water to which 2 µL cell culture was added (OD_{600} = 1.5) before the coverslip applied. Images were captured using a X100 oil-immersion lens and the FITC channel at 50% power and 500 ms exposure. Images were processed using ImageJ.

2. CellTrace ™ CFSE Cell Proliferation Kit (Thermo Fischer Scientific)

Cell labelling was achieved by following the manufacturers protocol. The cell pellet was re-suspended in PBS and diluted to contain 1×10^6 cells/mL. To 1 mL of cells, 1 µL of CellTrace ™ (stock = 5 mM in DMSO) was added and incubated for 20 mins at 37 °C, protected from light. To remove any free dye, 5 mL PBS with 3% BSA was added. Cells were pelleted by centrifugation (5 mins at 10,000 rcf), supernatant discarded and re-suspended in *E. coli* binding buffer (PBS containing 1 mM $CaCl_2$ and 1 mM $MnCl_2$).

3. Two-step labelling procedure: Biotin-Streptavidin

Protocol was based on methods of Otten *et al.,* 2016 [226]. The cell pellet was re-suspended in PBS to OD_{600} = 1. To diluted cells (1 mL), 2.5 µL (+)-Biotin N-

hydroxysuccinimide (NHS) ester (10 mg/mL in DMSO, Sigma Aldrich) was added and incubated at RT for 2.5 h. To remove unbound Biotin-NHS ester, cells were centrifuged (5 mins at 8000 rpm) and re-suspended in PBS (repeated three times). To 50 μL of biotinylated cells, 100 μL streptavidin-FITC (1:40, Thermo) was added and incubated for 1 h, shaking (400 rpm) at 37 °C. To remove unbound streptavidin-FITC, cells were centrifuged at (5 mins at 8000 rpm) and re-suspended in PBS (repeated three times). Following final centrifugation, cells were re-suspended in *E. coli* binding buffer (PBS containing 1 mM $CaCl_2$ and 1 mM $MnCl_2$).

To test the fluorescent intensity of each staining method across a range of *E. coli* dilutions, cells were diluted in *E. coli* binding buffer and added to wells of a 96-well plate in triplicate. Emitted fluorescence intensity was measured with a 485/520 nm filter, gain = 2000. All data is presented blank corrected, as an average across replicates.

For the live bacteria array, glycans (BSA-PEG$_4$-Man (10) = 50 μg/mL, 20 μL and mannan = 1000 μg/mL, 20 μL), were physiosorbed in triplicate in wells of a 384-well microplate and incubated for 1 h, shaking (225 rpm) at 37 °C. Wells were washed three times with MQ and dried with compressed air. For samples that were blocked, 3% BSA was added to well and incubated overnight, shaking (225 rpm) at RT before being washed three times with MQ and dried with compressed air. FITC-labelled *E. coli* were diluted in binding buffer to 1×10^8 cells/mL and added to wells (20 μL), before incubating for 1 h, shaking at 37 °C. Wells were washed three times with PBS and dried with compressed air. A BMG Labtech FLUOStar Omega microplate reader was used to measure emitted fluorescence intensity with a 485/520 nm filter, gain = 2000. All data is presented blank corrected, as an average across replicates. Positive controls were BSA-PEG$_4$-Man (10) with ConA and mannan with ConA. Negative controls were BSA-PEG$_4$-Man (10), mannan, ORN178 and ORN208.

5.3.5 Quartz Crystal Microbalance and Graphene Field Effect Transistors as Biosenors

All methods in this section pertain to the work done in collaboration with Mohamed, A *et al.*, [260].

5.3.5.1 General

Au-coated AT-cut α-quartz resonators with a fundamental frequency of (4.95 ± 0.05) MHz (QSX 301) were purchased from Biolin Scientific. All water was deionized before use to 18.2 MΩ cm at 25 °C.

Measurements using a quartz crystal microbalance with dissipation monitoring (QCM-D) were carried out on QSense E1 (Biolin Scientific) with a QFM 401 flow module. Deprotected sugars were lyophilized using a Büchi Lyovapor L-200 freeze dryer. Purification of the protected sugars via flash column chromatography was conducted using Biotage SP4 flash chromatography system with normal phase silica (pre-packed SNAP Ultra cartridges). ^{1}H, COSY, ^{13}C, DEPT-135 and HSQC NMR spectra were recorded on a Bruker Avance III 400 MHz at 298 K. Chemical shifts (δ) recorded in $CDCl_3$ and D_2O are reported with respect to the solvent residual peak at 7.26 and 4.79 ppm in ^{1}H NMR, respectively.

5.3.5.2 Cell Culture

E. coli ORN178 and ORN208 were grown overnight at 37 °C in LB media (10 mL) to attain an optical density measured at 600 nm (OD_{600}) of approximately 1.0 (~10^8 cells/mL). The cultures were centrifuged at 4,000 rpm for 10 min, the supernatant was discarded and the cell pellet was washed twice with PBS before re-suspending in fresh PBS buffer.

5.3.5.3 Quartz Crystal Microbalance Analysis

QCM-D resonators were washed with acid piranha etch 1:3 H_2SO_4:H_2O_2 (30%) and exposed to UV-ozone for 30 min. The resonator was then rinsed with ethanol then water and finally dried in a stream of N_2 gas. The resonator was functionalized with mannose by first drop casting 150µL of 27 mM 3,3'-Dithiodipropionic acid di(N-hydroxysuccinimide ester) (DTSP) dissolved in DMF and left overnight at room temperature in N_2 glovebox. The resonator was then rinsed with water and ethanol. The same procedure was then repeated with 150µL 27 mM 3-azidopropyl 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranoside (3-APM). The surface was then washed with water, ethanol, and acetone sequentially. All overnight incubation steps were carried out in N_2 glovebox to avoid thiol oxidation.

The sensor was loaded into a flow module (QFM 401) through which buffer was drawn through at 1 mL min^{-1} by peristaltic pump. The cell was maintained within 2 mK of 25 °C with the incorporated Peltier cooler. QSoft (v2.5.21) was used to acquire the frequency and dissipation changes for odd harmonics up to 13. The sensor was considered stabilized when the drift was <5kHz over 10 min. ConA adsorption were carried out in pH 7.4 HEPES buffer with added cations (10 mM HEPES, 150 mM NaCl, 4 mM $CaCl_2$, 4 mM $MnCl_2$) to facilitate ConA binding. *E. coli* adsorption was carried out in phosphate buffer pH 7.4.

The mass of dissolved adlayer was estimated using the Sauerbrey equation (eq. 1), which assumes a rigid adlayer, and by fitting to a Voigt viscoelastic model.

$$\Delta m = -C \frac{\Delta f_N}{N} \quad (1)$$

where m represents the adlayer mass density (ng/cm^{-2}), C is a proportionality constant that depends only on the intrinsic properties of the sensor (-17.9 ng Hz^{-1} cm^{-2} for 5 MHz crystals), N is the harmonic number (an odd positive number), and Δf_N is the frequency shift for that harmonic.

Sauerbrey analysis used the seventh harmonic because of its moderate penetration depth and low sensitivity to mounting effects. Viscoelastic fits used the frequency and dissipation responses from harmonics 3, 5, and 7, employing the chi squared minimization algorithm in QTools (v3.1.24.301), constraining the shear modulus to between 10^4 and 10^8 Pa and viscosity to between 10^{-4} and 10^{-1} Pa.

5.3.5.4 Graphene Field-Effect Transistors

10 mM 1-Pyrenebutanoic acid succinimidyl ester (PBASE) (in methanol) was drop casted onto the graphene surface and left overnight in N_2 glovebox followed by addition of 10 mM 3-APM overnight in N_2 glovebox to obtain functionalized surface. The surface was then rinsed with methanol and water. Phosphate-buffered saline (PBS) pH 7.4 from pre-formulated tablets was diluted 100-fold to 0.1 mM phosphate, 27 µM KCl, 1.37 mM NaCl. This solution was used in *E. coli* analysis. ConA analyses also included 40 µM $CaCl_2$ and 40 µM $MnCl_2$. Transfer characteristics were obtained by applying a constant 0.1 V drain-source voltage (V_{ds}) and sweeping the gate voltage (V_g) from 0–1.5 V with 0.03 V steps using an Agilent E5270B SMU. Raman spectra were measured with an InVia (Renishaw, UK) with a 532 nm laser and 2600 gr/mm. Raman maps were acquired from 1100 to 3200 cm^{-1} with 10 µm between points over a 50 µm^2 surface.

5.3.5.5 Data analysis

The currents from the GFET transfer curves acquired at gate voltages over a 0.24 V window centred on the minimum current were fit to a quadratic equation, and the value of V_{CNP} was estimated from that fit. The response (ΔV_{CNP}) was taken as the difference between V_{CNP} in the presence of ConA or *E. coli* and V_{CNP} in buffer only.

The QCM frequency response and shift in the GFET charge-neutral point were fit to Eq. 2 using OriginPro 2020b to obtain values for the equilibrium dissociation constant (K_d).

$$R([x]) = R_{max} \frac{[x]}{K_d + [x]} + C \qquad (2)$$

where R is the measurement response (Δf for QCM, ΔV_{CNP} for GFET), R_{max} is the limiting response at infinite concentration, [x] is the concentration of the species being titrated, and C is a constant offset. The offset was fixed to 0 Hz when fitting QCM data and allowed to vary for GFET data.

5.3.6 Flow Cytometry Aggregation Assay

To prepare samples, ORN178 and ORN208 were first cultured in LB (10 mL) overnight at 37°C, shaking (180 rpm). The cultures centrifuged at 4,000 rpm for 10 min, the supernatant was discarded, and the cell pellet was washed twice with PBS. The cell pellet was re-suspended in *E. coli* binding buffer (PBS containing 1 mM $CaCl_2$ and 1 mM $MnCl_2$) and diluted to 1 x 10^4 cells/mL. Diluted cells were then incubated, shaking (225 rpm) at 37 °C for 30 min with decreasing concentrations (1000, 100, 50, 20, 10, 0 μg/mL) of mannan from *Saccharomyces cerevisiae* (Sigma Aldrich) in a total volume of 500 μL.

Imaging cytometry was performed on an ImageStream X (Amnis). Images of 20,000 cells were captured with a 60x/NA 0.9 objective (pixel size 0.33um) and excited with 736nm laser at 3.75mW to generate back scatter. LED illumination was used to generate bright field images. The diameter of the core was maintained at 5um, with a bead percentage of 25%. Data was collected with 'remove beads' selected to prevent onboard focus beads being acquired in the data files. Analysis was performed on IDEAS v6.1.187 (Amnis). An analysis mask was generated to count the number of cells per image based on the back scatter signal. Utilizing the 'spot count' wizard within the software, cells per event, or image, were enumerated and exported as a text file. Data was plotted in Windows Excel.

5.4 References

146. Wang, W., Hu, T. S., Frantom, P. A., Zheng, T. Q., Gerwe, B., del Amo, D. S., Garret, S., Seidel, R. D. & Wu, P. (2009) Chemoenzymatic synthesis of GDP-L-fucose and the Lewis X glycan derivatives, *Proceedings of the National Academy of Sciences of the United States of America.* **106**, 16096-16101.

174. Noguchi, M., Tanaka, T., Gyakushi, H., Kobayashi, A. & Shoda, S. (2009) Efficient synthesis of sugar oxazolines from unprotected N-acetyl-2-amino sugars by using chloroformamidinium reagent in water, *J Org Chem.* **74**, 2210-2.

175. Shen, X., Yan, H., Wang, C., Gao, P., Johnson, C. H. & Snyder, M. P. (2022) TidyMass an object-oriented reproducible analysis framework for LC–MS data, *Nature communications.* **13**, 4365.

226. Otten, L., Fullam, E. & Gibson, M. I. (2016) Discrimination between bacterial species by ratiometric analysis of their carbohydrate binding profile, *Molecular bioSystems.* **12**, 341-344.

260. Mohamed, A. A. A., S.; Hassan, M. H.; Hassan, O.; Lloyd, J. E.; Field, R. A.; Blanford, C. F. Fabrication of a lectin biosensor for the detection of pathogenic bacteria in

262. Lau, K., Thon, V., Yu, H., Ding, L., Chen, Y., Muthana, M. M., Wong, D., Huang, R. & Chen, X. (2010) Highly efficient chemoenzymatic synthesis of beta1-4-linked galactosides with promiscuous bacterial beta1-4-galactosyltransferases, *Chemical communications (Cambridge, England).* **46**, 6066-8.

263. Barrick, J., Barnhart, C., Dasgupta, A., Renda, B., Suarez, G., Leon, D., Monk, J., Leonard, S., Perreau, J. & Deatherage , D. (2022) Media recipes: M9 minimal media in

Chapter 6

Chapter 6 – Appendix

6.1 Appendices from Chapter 1

Appendices 1.1: Characterised *Bifidobacterium* strains and genes, which utilise HMOs [40].

	Strain	Protein/Enzyme	Abbreviation/ Gene Locus	Enzymatic Activity	HMO Substrates	Ref
1	B. bifidum JCM1254	1,2-α-L-fucosidase	AfcA	Extracellular	2'FL, LNFPI, limited activity on 3FL and LNFPV	[264]
2	B. bifidum JCM1254	1,3/4-α-L-fucosidase	AfcB	Extracellular	3FL, LNFPII, LNFPIII	[265]
3	B. bifidum JCM1254	Exo-α-sialidase	SiaBb2	Extracellular	3'SL, DSLNT, 6'SL	[266]
4	B. bifidum JCM1254	β-galactosidase	BbgIII	Extracellular	LacNAc, LNnT, LNH, Lac	[42]
5	B. bifidum JCM1254	Lacto-N-biosidase	LNbB	Extracellular	LNT, LNH	[267]
6	B. bifidum JCM1254	β-N-Acetylglucosaminidase	BbhI	Extracellular	LNT1	[42]
7	B. bifidum JCM1254	GNB/LNB phosphorylase	LnpA1	Intracellular	LNB/GNG	[268]
8	B. bifidum JCM1254	GNB/LNB phosphorylase	LnpA2	Intracellular	LNB/GNB	[268]
9	B. longum sp. infantis ATCC 15697	Transporter SBP	FLP1-BP	-	2'FL	[269]
10	B. longum sp. infantis ATCC 15697	Transporter SBP	FLP2-BP	-	2'FL, 3FL, LDFT, LNFPI	[269]
11	B. longum sp. infantis ATCC 15697	α-L-fucosidase	AfcA	Intracellular	LNFPI, 2'FL, 3FL	[270]
12	B. longum sp. infantis ATCC 15697	1,3/4-α-L-fucosidase	AfcB	Intracellular	LNFPIII, 3FL	[270]
13	B. longum sp. infantis ATCC 15697	1,3/4-α-L-fucosidase	Blon_0248	Intracellular	LNFPIII	[270]
14	B. longum sp. infantis ATCC 15697	α-L-fucosidase	Blon_0426	Intracellular	LNFPIII	[270]

15	*B. longum sp. infantis* ATCC 15697	β-galactosidase	Bga2A	Intracellular	Lac, LacNAc, LNnT	[43]
16	*B. longum sp. infantis* ATCC 15697	LNT β-galactosidase	Bga42A	Intracellular	LNT, LNB	[43]
17	*B. longum sp. infantis* ATCC 15697	β-N-Acetylglucosaminidase	Blon_0459	Intracellular	LNT, LNH, LNT1	[271]
18	*B. longum sp. infantis* ATCC 15697	β-N-Acetylglucosaminidase	Blon_0732	Intracellular	LNT, LNH, LNT1	[271]
19	*B. longum sp. infantis* ATCC 15697	β-N-Acetylglucosaminidase	Blon_2355	Intracellular	LNT, LNH, LNT1	[271]
20	*B. breve* UCC2003	Transporter SBP	NahS		LNnT	[41]
21	*B. breve* UCC2003	β-galactosidase	LntA	Intracellular	LNT, LNnT, Lac	[41]
22	*B. breve* UCC2003	β-galactosidase	LazZ2	Intracellular	LNnT, Lac	[41]
23	*B. breve* UCC2003	β-galactosidase	LacZ6	Intracellular	LNnT, Lac	[41]
24	*B. breve* UCC2003	β-N-Acetylglucosaminidase	NahA	Intracellular	LNT2	[41]
25	*B. breve* UCC2003	GNB/LNB phosphorylase	LnbP	Intracellular	LNB	[41]
26	*B. longum sp. longum* JCM1217	Lacto-N-biosidase	LnbX	Extracellular	LNT, LNH, LNFPI, LST a	[272]
27	*B. longum sp. longum* JCM1217	Chaperone for LnbX	LnbY	Extracellular	-	[272]
28	*B. longum sp. longum* JCM1217	Transporter SBP	GL-BP	-	LNB, GNB	[273]
29	*B. longum sp. longum* JCM1217	β-N-Acetylglucosaminidase	BLLJ_1391	Intracellular	LNT1	[274]
30	*B. longum sp. longum* JCM1217	N-acetylhexosamine 1-kinase	NahK	Intracellular	GlcNAc/GalNAc	[275]
31	*B. longum sp. longum* JCM1217	GNB/LNB phosphorylase	LnpA	Intracellular	LNB/GNB	[275]

6.2 Appendices from Chapter 2

A

B
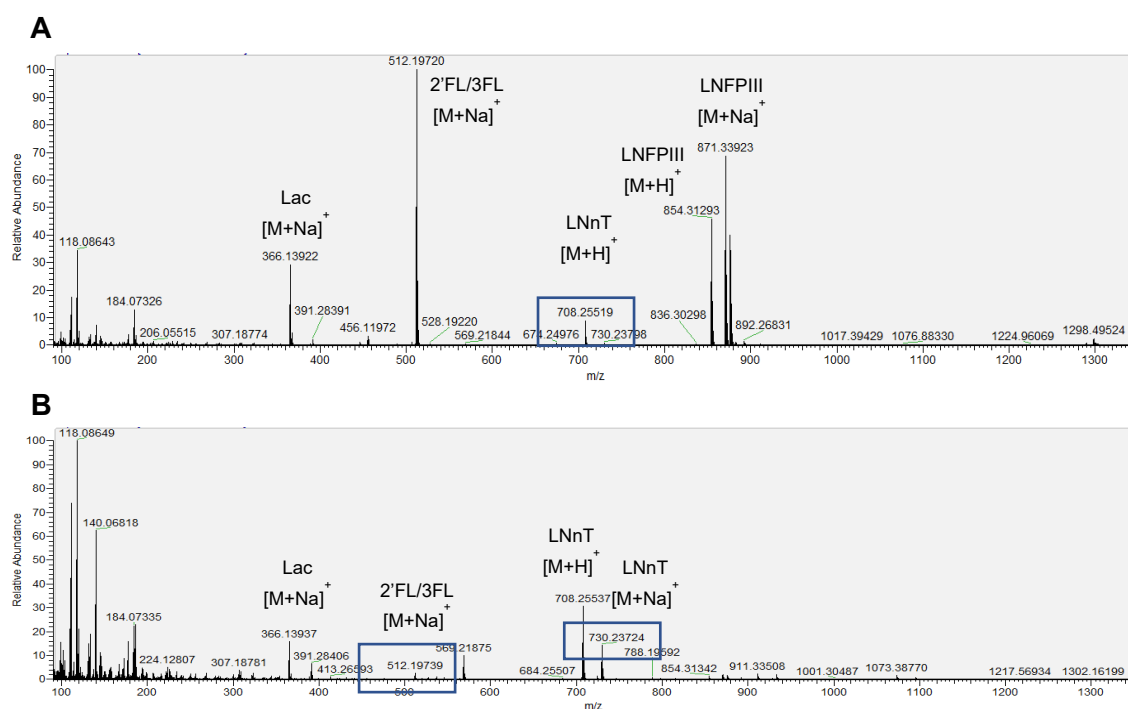
Appendices 2.1: HILIC LC-MS analysis of HMO mixtures M1 (A) and M2 (B).

Appendices 2.2: HPAEC-PAD retention times of commercial HMO standards.

Commercial Standard	Retention Time (min)
Lac	6.065
2'FL	6.533
3FL	4.500
LNFPIII	4.505
LNnT	9.462

A

B

C

D
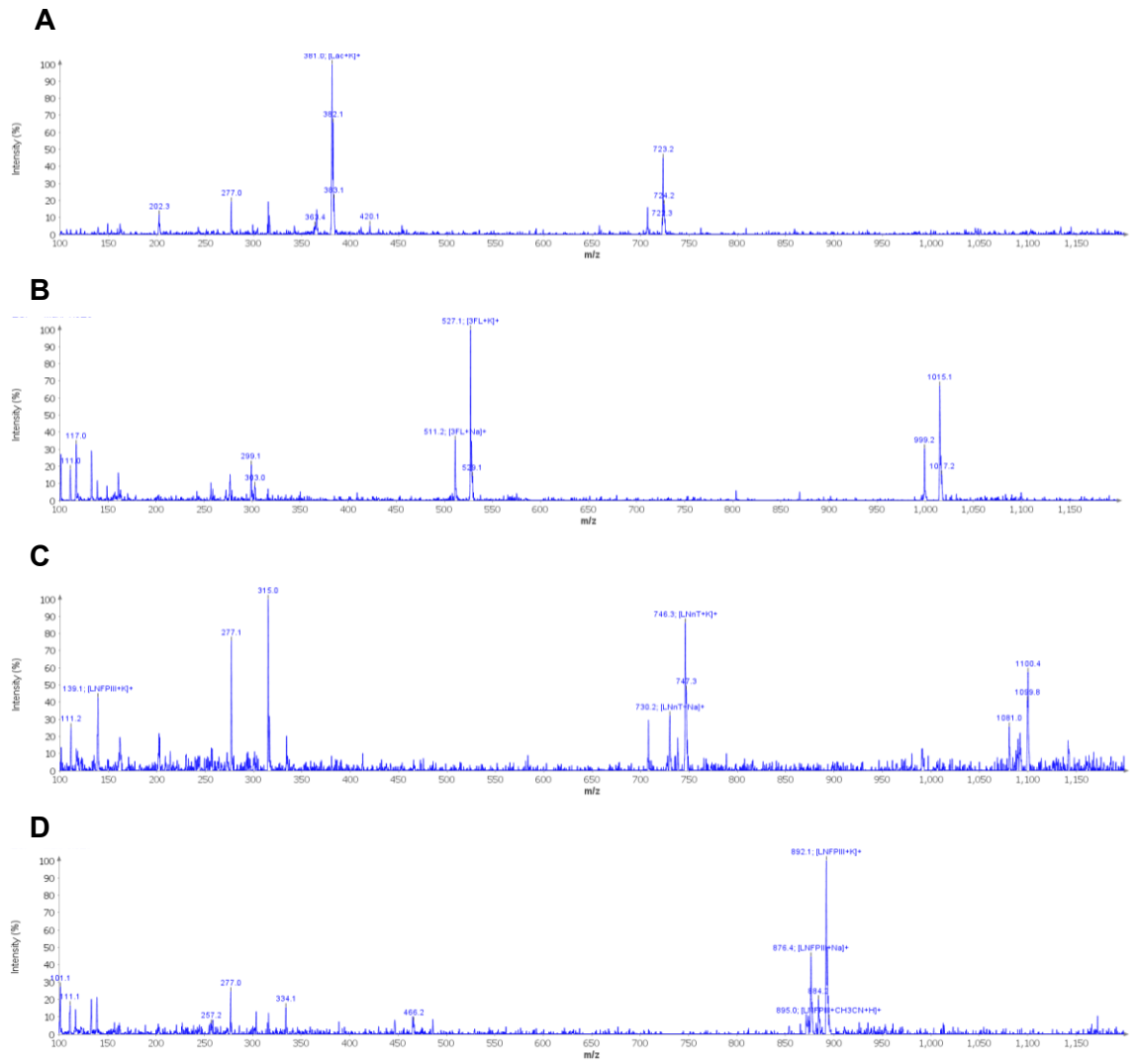
Appendices 2.3: Low-resolution mass spectra of M1 HILIC HPLC peaks 1 – 4. A) Lac; B) 3FL; C) LNnT; D) LNFPIII.
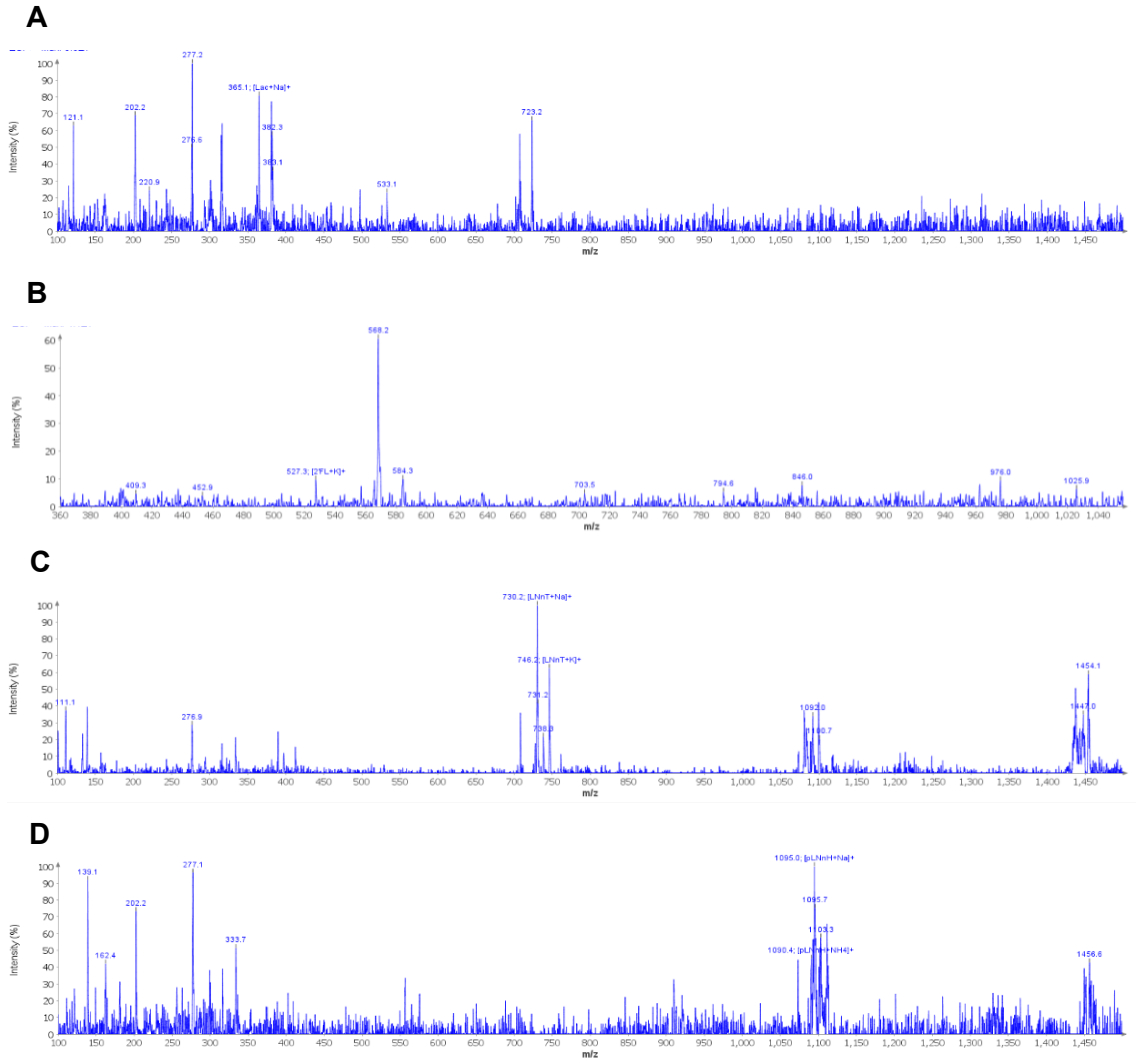
Appendices 2.4: Low-resolution mass spectra of M2 HILIC HPLC peaks 1 – 4. A) Lac; B) 3FL; C) LNnT; D) LNFPIII.

159
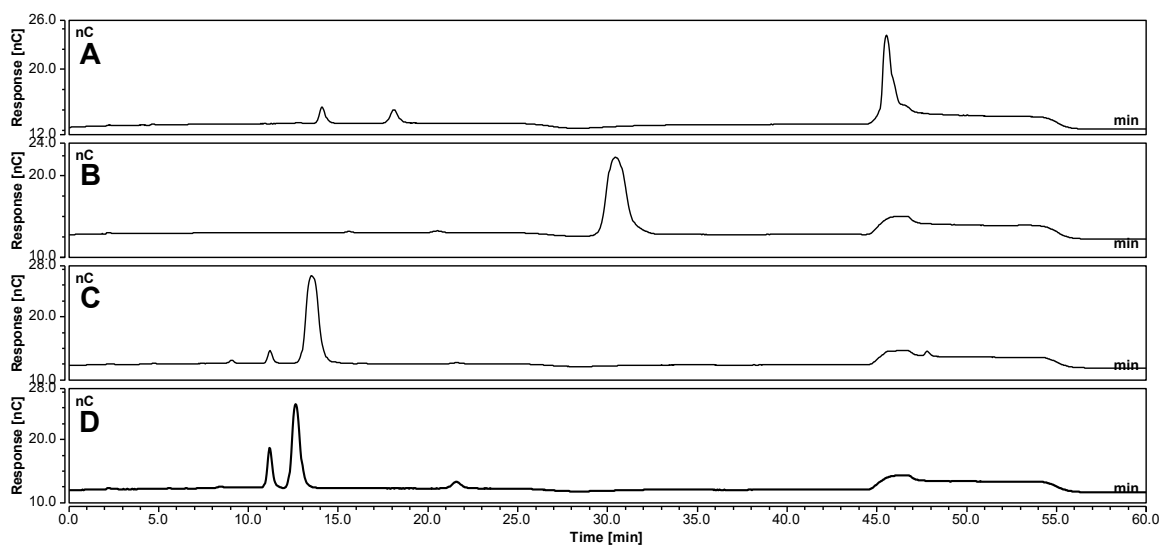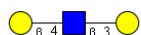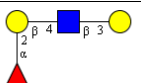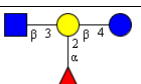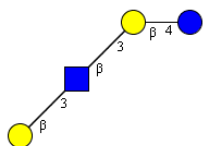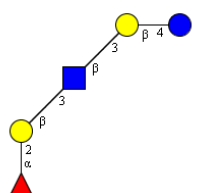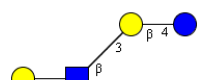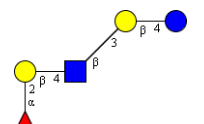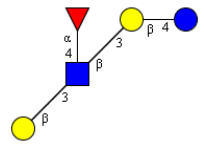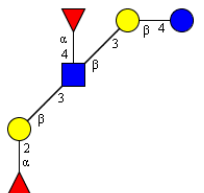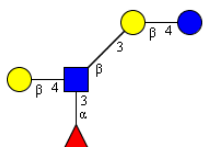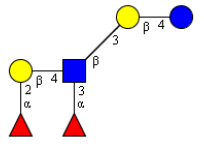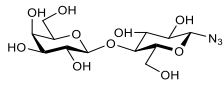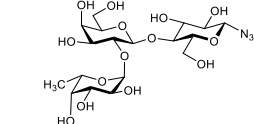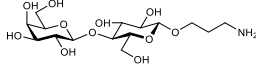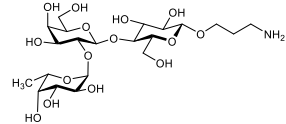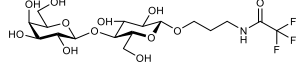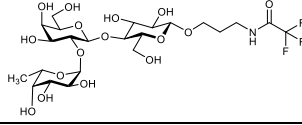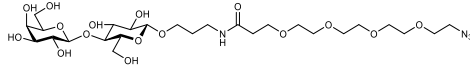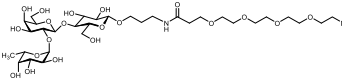
Appendices 2.5: HPAEC-PAD analysis of crude commercial compounds 8 – 11 (A to D).

Appendices 2 6: HPAEC-PAD retention times of commercial standards (compounds 1 to 11).

Compound	Name Abbreviation	Retention Time (min)
1	Gal	8.950
2	Lac	14.675
3	LacNB	18.417
4	LacNAc	14.817
5	3FL	12.225
6	LNT1	23.359
7	LNT2	20.550
8	LNT	47.209
9	LNnT	28.742
10	LNFPII	12.634
11	LNFPIII	11.792

Appendices 2.7: Summary of α-1,2-fucosyltransferase activity from *H. mustelae* for compounds 1 to 15.

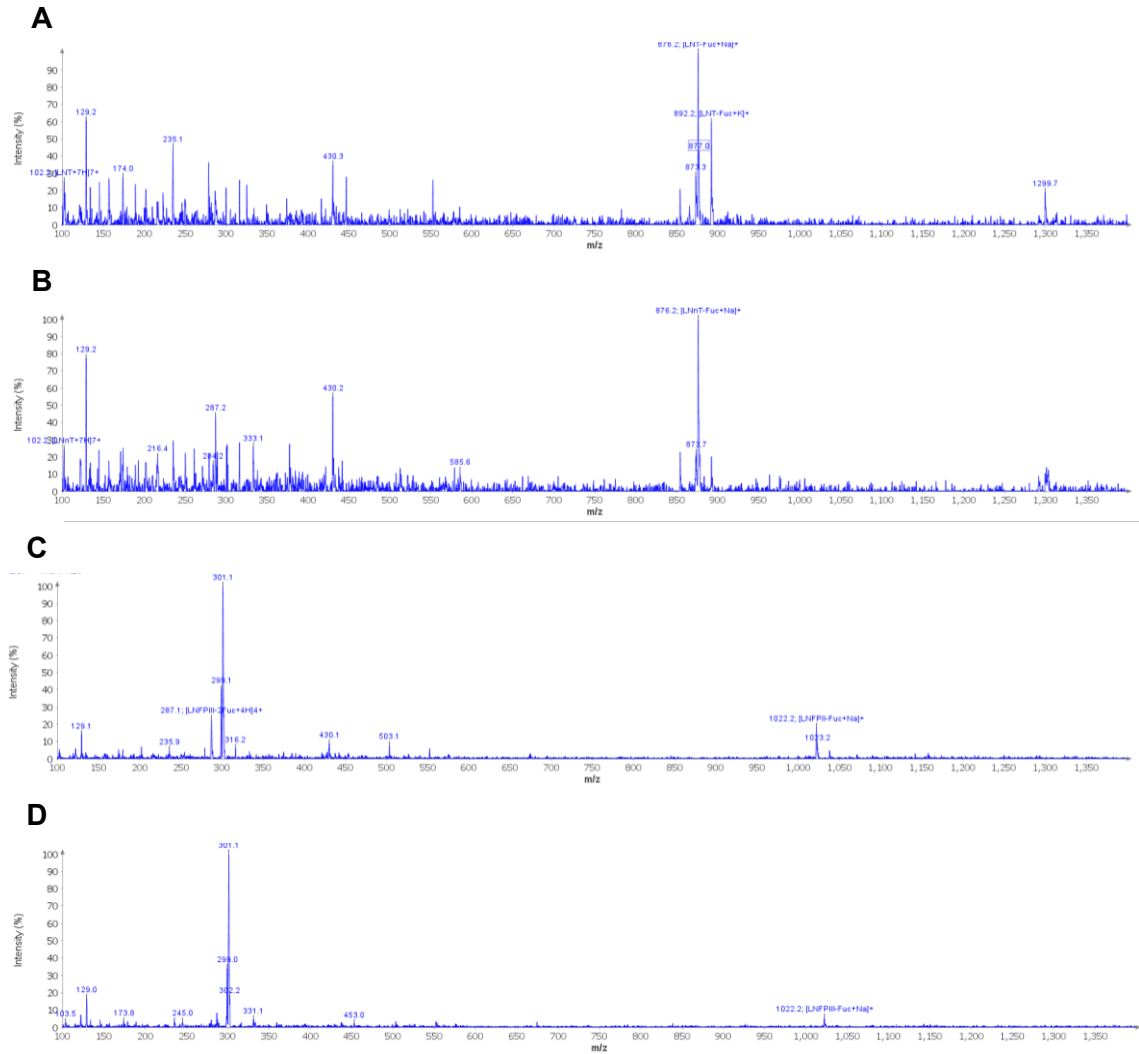| Entry	Substrate	Conversion (%)	Product
1		100	
2		76	
3		95	
4		71	
5		60	
6		61	
7		0	
8		100	
9		100	
10		21	
11		5	

12		98	
13		3	
14		7	
15		9	

Appendices 2.8: Low resolution mass spectra of fucosylated HMO analogue reaction products from α1,2-FucT following HILIC HPLC purification. (A) LNT-Fuc, (B) LNnT-Fuc, (C) LNFPII-Fuc and (D) LNFPIII-Fuc.

Appendices 2.9: SAX chromatography retention times of commercial nucleotide standards.

Commercial Standard	Retention Time (min)
AMP	5.540
ADP	7.367
ATP	9.230
GMP	6.230
GDP	8.244
GTP	10.234

6.2 References

40. Masi, A. C. & Stewart, C. J. (2022) Untangling human milk oligosaccharides and infant gut microbiome, *iScience.* **25**, 103542.

41. James, K., Motherway, M. O., Bottacini, F. & van Sinderen, D. (2016) *Bifidobacterium breve* UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways, *Sci Rep.* **6**, 38560.

42. Miwa, M., Horimoto, T., Kiyohara, M., Katayama, T., Kitaoka, M., Ashida, H. & Yamamoto, K. (2010) Cooperation of β-galactosidase and β-N-acetylhexosaminidase from *bifidobacteria* in assimilation of human milk oligosaccharides with type 2 structure, *Glycobiology.* **20**, 1402-9.

43. Yoshida, E., Sakurama, H., Kiyohara, M., Nakajima, M., Kitaoka, M., Ashida, H., Hirose, J., Katayama, T., Yamamoto, K. & Kumagai, H. (2012) *Bifidobacterium longum* subsp. *infantis* uses two different β-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides, *Glycobiology.* **22**, 361-8.

264. Katayama, T., Sakuma, A., Kimura, T., Makimura, Y., Hiratake, J., Sakata, K., Yamanoi, T., Kumagai, H. & Yamamoto, K. (2004) Molecular cloning and characterization of *Bifidobacterium bifidum* 1,2-alpha-L-fucosidase (AfcA), a novel inverting glycosidase (glycoside hydrolase family 95), *J Bacteriol.* **186**, 4885-93.

265. Ashida, H., Miyake, A., Kiyohara, M., Wada, J., Yoshida, E., Kumagai, H., Katayama, T. & Yamamoto, K. (2009) Two distinct alpha-L-fucosidases from *Bifidobacterium bifidum* are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates, *Glycobiology.* **19**, 1010-7.

266. Kiyohara, M., Tanigawa, K., Chaiwangsri, T., Katayama, T., Ashida, H. & Yamamoto, K. (2010) An exo-α-sialidase from *bifidobacteria* involved in the degradation of sialyloligosaccharides in human milk and intestinal glycoconjugates, *Glycobiology.* **21**, 437-447.

267. Wada, J., Ando, T., Kiyohara, M., Ashida, H., Kitaoka, M., Yamaguchi, M., Kumagai, H., Katayama, T. & Yamamoto, K. (2008) *Bifidobacterium bifidum* lacto-N-biosidase, a critical enzyme for the degradation of human milk oligosaccharides with a type 1 structure, *Appl Environ Microbiol.* **74**, 3996-4004.

268. Nishimoto, M. & Kitaoka, M. (2007) Identification of the putative proton donor residue of lacto-N-biose phosphorylase (EC 2.4.1.211), *Biosci Biotechnol Biochem.* **71**, 1587-91.

269. Sakanaka, M., Hansen, M. E., Gotoh, A., Katoh, T., Yoshida, K., Odamaki, T., Yachi, H., Sugiyama, Y., Kurihara, S., Hirose, J., Urashima, T., Xiao, J.-z., Kitaoka, M., Fukiya, S., Yokota, A., Lo Leggio, L., Abou Hachem, M. & Katayama, T. (2019) Evolutionary adaptation in fucosyllactose uptake systems supports bifidobacteria-infant symbiosis, *Science Advances.* **5**, eaaw7696.

270. Sela, D. A., Garrido, D., Lerno, L., Wu, S., Tan, K., Eom, H. J., Joachimiak, A., Lebrilla, C. B. & Mills, D. A. (2012) *Bifidobacterium longum* subsp. *infantis* ATCC 15697 α-fucosidases are active on fucosylated human milk oligosaccharides, *Appl Environ Microbiol.* **78**, 795-803.

271. Garrido, D., Ruiz-Moyano, S. & Mills, D. A. (2012) Release and utilization of N-acetyl-D-glucosamine from human milk oligosaccharides by *Bifidobacterium longum* subsp. *infantis*, *Anaerobe.* **18**, 430-5.

272. Sakurama, H., Kiyohara, M., Wada, J., Honda, Y., Yamaguchi, M., Fukiya, S., Yokota, A., Ashida, H., Kumagai, H., Kitaoka, M., Yamamoto, K. & Katayama, T. (2013) Lacto-N-biosidase encoded by a novel gene of *Bifidobacterium longum* subspecies *longum* shows unique substrate specificity and requires a designated chaperone for its active expression, *J Biol Chem.* **288**, 25194-25206.

273. Suzuki, R., Wada, J., Katayama, T., Fushinobu, S., Wakagi, T., Shoun, H., Sugimoto, H., Tanaka, A., Kumagai, H., Ashida, H., Kitaoka, M. & Yamamoto, K. (2008) Structural and thermodynamic analyses of solute-binding protein from *Bifidobacterium longum* specific for core 1 disaccharide and lacto-*N*-biose *J Biol Chem.* **283**, 13165-13173.

274. Honda, Y., Nishimoto, M., Katayama, T. & Kitaoka, M. (2013) Characterization of the Cytosolic β-Acetylglucosaminidase from *Bifidobacterium longum* subsp. *longum*, *Journal of Applied Glycoscience.* **60**, 141-146.

275. Kitaoka, M., Tian, J. & Nishimoto, M. (2005) Novel putative galactose operon involving lacto-*N*-biose phosphorylase in *Bifidobacterium longum*, *Appl Environ Microbiol.* **71**, 3158-3162.