<table>
<thead>
<tr>
<th>Title</th>
<th>Targeting the glycome of the milk fat globule membrane for anti-infective properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Ross, Sarah</td>
</tr>
<tr>
<td>Publication Date</td>
<td>2016-10-21</td>
</tr>
<tr>
<td>Item record</td>
<td><a href="http://hdl.handle.net/10379/6087">http://hdl.handle.net/10379/6087</a></td>
</tr>
</tbody>
</table>
Targeting the glycome of the milk fat globule membrane for anti-infective properties

A thesis presented to National University of Ireland, Galway for the degree of Doctor of Philosophy (Ph.D.) in the National Centre for Biomedical Engineering Science (NCBES)

October 2016

By

Sarah A. Ross B.Sc.

1 Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork
2 National University of Ireland Galway, Galway

Head of NCBES: Prof. Rhodri Ceredig

Research supervisors: Dr. Rita M. Hickey\textsuperscript{1} and Prof. Lokesh Joshi\textsuperscript{2}
## Contents

Declaration .................................................................................................................................................. i
Thesis abstract ........................................................................................................................................ ii
Publications ............................................................................................................................................... v
Abstracts ................................................................................................................................................ vi
Oral communications ........................................................................................................................ vii
Abbreviations .......................................................................................................................................... viii
List of tables ........................................................................................................................................... xi
List of figures ........................................................................................................................................... xiii
Acknowledgements ............................................................................................................................ xvi

**Thesis Introduction** ................................................................................................................................ xviii

**Chapter I. The milk fat globule membrane: A potential soure of health-promoting glycans** ................................................................. 1

Abstract .................................................................................................................................................. 2
Introduction ........................................................................................................................................... 3
MFGM composition ............................................................................................................................... 7
Glycoconjugates of the MFGM .......................................................................................................... 11
  Glycolipids ........................................................................................................................................ 11
  Glycoproteins ................................................................................................................................... 14
    Mucins ........................................................................................................................................... 14
    Butyrophilin .................................................................................................................................. 17
    Lactadherin ................................................................................................................................... 20
    CD59 ........................................................................................................................................... 22
    CD36 ........................................................................................................................................... 23
    Proteose peptone component 3 ................................................................................................. 24
    Xanthine oxidoreductase ........................................................................................................... 26
    Carbonic anhydrase .................................................................................................................... 27
    Clusterin ...................................................................................................................................... 29
Commercial potential .......................................................................................................................... 31
Chapter II. Defatted bovine milk fat globule membrane inhibits the association of enterohaemorrhagic Escherichia coli with human HT-29 cells .................................................................................. 67

Abstract ............................................................................................................. 68

Introduction .......................................................................................................... 69

Materials and methods ....................................................................................... 72

Preparation of a defatted MFGM-enriched fraction (dMFGM) from bovine milk .................................................................................................................. 72

Sodium dodecyl sulfate polyacrylamide gel electrophoresis analysis ........... 73

Bacterial strains and culture conditions ............................................................. 73

Mammalian cell culture ....................................................................................... 74

In vitro total association assays .......................................................................... 74

Standard competition assay ............................................................................. 75

Concentration dependency assay .................................................................... 75

Bacterial interaction assay ............................................................................... 76

Host cell interaction assay .............................................................................. 76

Bacterial growth assays .................................................................................... 77

Statistical analysis ............................................................................................ 77

Results and discussion ...................................................................................... 78

Isolation and characterisation of the dMFGM fraction ..................................... 78

dMFGM reduced the association of E. coli O157:H7 with HT-29 cells............. 79

Conclusions ......................................................................................................... 85

References ........................................................................................................... 86

Chapter III. Temporal alterations in the bovine buttermilk glycome from parturition to milk maturation ................................................................. 98

Abstract ............................................................................................................. 99

Introduction ......................................................................................................... 100
Materials and Methods ................................................................. 103
Materials ....................................................................................... 103
Sample collection ........................................................................... 103
Buttermilk generation ................................................................. 104
Characterisation of buttermilk samples ........................................ 104
Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis ................................................................. 104
Fluorescent labeling of MFGM and glycoproteins ............................ 105
Construction of lectin microarrays and MFGM profiling .................. 105
Data extraction, hierarchical clustering and statistical analysis ......... 106
Lectin blotting analysis ................................................................... 107
Results and discussion .................................................................. 109
Characterisation of bovine buttermilk over 3 month lactation ......... 109
MFGM glycoprofiling on lectin microarray ...................................... 111
Lectin blotting analysis of bovine buttermilk samples .................... 117
Conclusions ................................................................................... 119
References ..................................................................................... 120
Supplementary figures and tables ................................................... 133

Chapter IV. Lectin and neoglycoconjugate microarray profiling of enterohaemorrhagic and enteropathogenic Escherichia coli strains indicate serotype dependent glycan presentation and glycan binding affinity .... 148
Abstract ....................................................................................... 149
Introduction ................................................................................... 150
Material and methods .................................................................. 155
Materials ....................................................................................... 155
Bacterial strains and culture conditions ......................................... 155
Fluorescent labeling of bacterial strains ........................................ 155
Construction of lectin and NGC microarrays ................................... 156
Bacterial profiling with glycomic microarrays ................................. 157
Data extraction and processing ..................................................... 158
Statistical analysis ........................................................................ 158
Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and has not been submitted for any degree or examination at the National University of Ireland, Galway or elsewhere, and that all the sources I have used or quoted have been indicated and acknowledged by means of completed references.

Signed:

Sarah Ross

Date:
**Thesis abstract**

This thesis sets out to explore the potential bioactivities of bovine milk fat globule membrane (MFGM) to enhance our understanding of the roles bovine milk glycoconjugates may play in humans and to contribute towards further discovery and development of functional ingredients which may promote human health. Glycosylated MFGM proteins and lipids in particular, have been accredited with a number of bioactivities, but the full extent of their potential health promoting properties have yet to be realised.

Given the reported anti-infective properties associated with bovine MFGM glycoproteins, a defatted bovine MFGM glycoprotein fraction, rich in proteins and glycoproteins (dMFGM), was generated. The ability of this fraction to inhibit *Escherichia coli* O157:H7 association with human colonic adenocarcinoma, HT-29 cells was subsequently investigated. The dMFGM fraction was observed to reduce the association of several *E. coli* O157:H7 strains with the HT-29 cells. This activity was strain-specific and concentration dependent. The parameters of the experimental assays were varied to identify the potential mechanisms by which the dMFGM fraction exerted its activity. The anti-infective activity was shown to occur as a result of the dMFGM interacting with the *E. coli* rather than the HT-29 cells. This study identified the potential use of defatted bovine MFGM to prevent the onset of *E. coli* infections in humans.

Since the glycan portion of milk glycoconjugates may be an important factor in their health-promoting activities, the changes occurring in buttermilk, a source of MFGM, were profiled throughout the course of lactation. Buttermilk samples were generated at 13 time points during lactation for three multiparous animals and their glycosylation patterns were profiled through the use of lectin microarrays and lectin blotting. These techniques provided a platform for high-throughput analysis of the buttermilk sample glycosylation. The data suggested differences in glycosylation, including N-glycosylation, sialylation and fucosylation, between early and late time points. In addition, differences were evident between individual animals at various time points.
throughout the lactation cycle. The findings of this study may be invaluable in order to target the isolation of glycosylated ingredients for functional foods from particular lactation time points to maximise the abundance of particular glycan structures.

In relation to further investigation of the anti-infective activities described above, defining the glycosylation patterns of MFGM and/or buttermilk is of importance however defining the surface glycosylation of pathogenic bacteria is also of great importance. Similarly, the host’s glycan presentation is important for pathogen binding and thus identification of the glycan binding preferences of pathogenic species is necessary to better understand how to reduce the threat of infection. The cell surface glycosylation and carbohydrate binding profiles of 5 enterohaemorrhagic *E. coli* (EHEC) O157:H7 strains and 2 enteropathogenic *E. coli* (EPEC) strains were characterized and compared using lectin and neoglycoconjugate microarrays in order to investigate if their surface glycosylation patterns and binding preferences were markedly different. The lectin microarray profiles suggested the presence of a variety of cell surface glycan structures with different apparent abundance or accessibility for each strain screened. An apparent difference between the surface glycosylation of the EHEC and EPEC strains was evident. Furthermore, the neoglycoconjugate microarray profiling identified the bacterial strains could bind to a relatively high number of oligosaccharide structures and interestingly, the binding preferences of EHEC and EPEC were markedly different. These results demonstrate that EHEC and EPEC surface glycosylation profiles and binding preferences differ. This may partially explain the different colonisation properties observed in chapter 2. Furthermore, knowledge of these glycan profiles may be invaluable in the design of anti-infectives targeted at specific bacterial strains.

In addition to its anti-infective activity, bovine buttermilk and its components have been reported to possess immune-modulating properties. The transcriptional responses of colonic epithelial cells to buttermilk generated from colostrum and from mature milk were compared since the glycosylation status
is known to change through the milk maturation process. Genomic microarrays were employed to identify any processes influenced by exposure to buttermilk. A number of processes were shown to be affected by both treatments including apoptosis, reproduction, immune system processes, barrier function, metabolism and stimulus response. Subsequently, immune system processes were chosen for further analysis as these were the most significantly modulated by the colostrum and mature buttermilk treatments. Polymerase chain reaction (PCR) validation indicated that the buttermilk treatments influenced the expression of a number of immune-associated genes including chemokine, cytokine and anti-viral genes. Interestingly, an apparent anti-inflammatory effect was observed indicating the potential of bovine buttermilk in preventing inflammatory diseases.

Overall these studies indicate the potential of bovine MFGM fractions to promote human health by reducing the threat of infection and modulating immune response and inflammation. Moreover, the selection of bovine MFGM from certain lactation stages may be advantageous in order to increase the level of bioactivity by using the most appropriate glycosylation profile in the dairy fraction. Future research, including in vivo trials, are necessary to help identify the impact such fractions may have on the maintenance and preservation of human health.
Publications


Abstracts


Oral communications


Abbreviations

ADPH – Adipophilin
A/E – Attaching and effacing
AF – Alexafluor
AIDA – Adhesion involved in diffuse adherence
ASF - Asialofetuin
ATCC – American Type Culture Collection
BCA – Bicinchoninic acid
BHI – Brain heart infusion
BMO – Bovine milk oligosaccharides
BSA – Bovine serum albumin
BTN – Butyrophilin
CA – Carbonic anhydrase
cBM – Colostrum buttermilk
dMFGM – Defatted milk fat globule membrane
DPC – Dairy Products Research Centre
E. coli – Escherichia coli
EAE - Experimental autoimmune encephalomyelitis
EHEC – Enterohaemorrhagic E. coli
EPEC – Enteropathogenic E. coli
ER - Endoplasmic reticulum
Fuc – Fucose
Gal – Galactose
GalNAc – N-acetylgalactosamine
GD – Disialoganglioside
GIT – Gastrointestinal tract
Glc - Glucose
GlcNAc – N-acetylglucosamine
GM – Monosialoganglioside
Gp330 – Glycoprotein 330
GT – Trisialoganglioside
HIV – Human immunodeficiency virus
HMO – Human milk oligosaccharides
HRP – Horseradish peroxidase
IFN - Interferon
IgG – Immunoglobulin G
JAK – Janus kinase
LacdiNAc - N, N′-diacetyllactosamine
LacNAc - N-acetyllactosamine
LacCer – Lactosylceramide
LDH – Lactadherin
LDS – Lithium dodecyl sulfate
LP – Lactophorin
LPS - Lipopolysaccharide
Man – Mannose
mBM – Mature buttermilk
MFG – Milk fat globule
MFGM – Milk fat globule membrane
Muc – Mucin
NCTC – National Collection of Type Cultures
Neu5Ac – N-acetyl neuraminic acid
Neu5Gc – N-glycolyl neuraminic acid
NGC - Neoglycoconjugate
PAS - Periodic acid-schiff
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PIA</td>
<td>Polysaccharide intercellular adhesin</td>
</tr>
<tr>
<td>PP3</td>
<td>Proteose peptone 3</td>
</tr>
<tr>
<td>PrCA</td>
<td>Principal component analysis</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride</td>
</tr>
<tr>
<td>RFU</td>
<td>Relative fluorescence units</td>
</tr>
<tr>
<td>RGD</td>
<td>Arg-Gly-Asp</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulphate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SM</td>
<td>Sphingomyelin</td>
</tr>
<tr>
<td>STAT</td>
<td>Signal transducer and activator of transcription</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris buffered saline</td>
</tr>
<tr>
<td>TBS-T</td>
<td>Tris buffered saline with 0.05% Tween 20</td>
</tr>
<tr>
<td>VNTR</td>
<td>Variable number of tandem repeat</td>
</tr>
<tr>
<td>WGA</td>
<td>Wheat germ agglutinin</td>
</tr>
<tr>
<td>XDH</td>
<td>Xanthine dehydrogenase</td>
</tr>
<tr>
<td>XO</td>
<td>Xanthine oxidase</td>
</tr>
<tr>
<td>XOR</td>
<td>Xanthine oxidoreductase</td>
</tr>
</tbody>
</table>
List of tables

Chapter I

Table 1. Overview of the mammalian MFGM glycolipids and their properties

Table 2. Overview of the mammalian MFGM glycoproteins and their properties

Chapter III

Table S1. Lectins printed in order, their binding specificities, their simple print sugars (1 mM) and the supplying company

Chapter IV

Table S1. Lectins printed in order, their binding specificities, their simple print sugars (1 mM) and the supplying company

Table S2. Neoglycoconjugates printed in order and their substitution ratio

Chapter V

Table 1. Gene ontology groups affected by buttermilk samples

Table 2. Immune related gene ontology categories affected by buttermilk samples

Table 3. Differential expression of genes potentially involved in maintenance of epithelial barrier

Table 4. Differential expression of immune associated genes identified by RT-PCR
General summary and discussion

Table 1. Cases of *Escherichia coli* O157:H7 infection in the USA
List of figures

Chapter I

Figure 1. Structure of the MFGM

Figure 2. Outline of MFGM processing

Chapter II

Figure 1. Electrophoresis profile of bovine milk fractions

Figure 2. The effect of the dMFGM fraction on association of *E. coli* strains with human HT-29 cells

Figure 3. The effect of the dMFGM fraction on association of *E. coli* O157:H7 with HT-29 cells

Chapter III

Figure 1. Protein characterisation of buttermilk samples over the course of lactation

Figure 2. Acidic and neutral sugar characterisation of buttermilk samples over the course of lactation

Figure 3. Heatmap and hierarchical clustering results for lectin microarray profiles generated for labelled buttermilk samples

Figure 4. Lectin microarray profiles of bovine buttermilk generated for three animals over the course of lactation

Figure 5. Lectin blot profiles of bovine buttermilk generated for animal 1 over the course of lactation

Figure S1. Representative image of experimental subarrays
Figure S2. NuPage 4-12% SDS PAGE gels for buttermilk samples for all animals for days 1-90 post-partum

Figure S3. Differential gel band intensities for the bands at 45/50 kDa of the coomassie-stained SDS PAGE gel

Figure S4. Lectin microarray profiles of bovine buttermilk generated for three animals demonstrating all time points over the course of lactation

Figure S5. Line graphs representing lectin microarray profiles of bovine buttermilk generated for all animals demonstrating all time points over the course of lactation

Figure S6. (A) SBA lectin microarray profile of bovine buttermilk generated for all animals over the course of lactation (Days 1-10 and days 30, 70 and 90) in comparison to (B) densitometry of the SBA binding to bovine milk fat globule membrane glycoproteins (Days 1, 2, 4 and 6) as published by Ujita et al., (1993)

Figure S7. Lectin blot profiles of bovine buttermilk generated for animal 1 over the course of lactation

Figure S8. Differential gel band intensities were quantified using ImageJ software for LEL (A) and AIA (B) lectin blots. Arbitrary units were used

Chapter IV

Figure 1. Optimization of E. coli NCTC 12900-, ATCC 43888-, DPC 6055-, DPC 6054-, DAF454-, NCTC 8007- and NCTC 8623-staining with Syto-82 at different concentrations of fluorescent dye

Figure 2. Lectin microarray profile of E. coli strains NCTC 12900, ATCC 43888, DPC 6055 and DPC 6054 (A), NCTC 8007 and NCTC 8623 (B), and E. coli DAF454 (C)
Figure 3. Unsupervised hierarchical clustering and principal component analysis results for lectin microarray profiles generated for labelled *E. coli* strains

Figure 4. Neoglycoconjugate microarray profiles of EHEC and EPEC strains

Figure 5. Unsupervised hierarchical clustering and principal component analysis results for neoglycoconjugate microarray profiles generated for labelled *E. coli* strains

**Chapter V**

Figure 1. Protein characterisation of buttermilk samples over the course of lactation

Figure 2. The number of DET affected by buttermilk samples

Figure 3. Quantitative reverse transcription-polymerase chain reaction (qRT-PCR) analysis of immune-related genes in HT-29 cells treated with bovine colostrum buttermilk (cBM) or bovine mature buttermilk (mBM)
Acknowledgements

First and foremost, I am very grateful to my research supervisors, Dr. Rita Hickey and Prof. Lokesh Joshi, for their invaluable knowledge, expert opinions and guidance over the past few years. I would also like give a special mention to Jared in NUIG, for the extremely helpful advice and support throughout my PhD and to Jonny who was a wealth of knowledge – I was very lucky to have you as a post-doc in my first year.

My PhD was funded by the Food Institutional Research Measure and I would like to thank them for providing me with the financial support necessary to undertake my PhD project. I would also like to thank Teagasc Moorepark and National University of Ireland, Galway for the use of their facilities and equipment. In particular, I would like to thank the staff and students in Teagasc Moorepark for the friendly and encouraging atmosphere - it was a pleasure to be able to undertake my project here and I will always have great memories from my time in the Food Research Centre.

Of course I have to thank the Glyco girls – Sinead, Erinn, Aoife, Mariarosaria and Helen- and the boys, especially to JT for the delicious cakes and to Jonny for the endless advice and help down through the years. I also can’t forget Sinead Mackey – the office was never the same after you left. It has been a pleasure to work with such a great group of people.

Ciara, Ann, Kate and Bernie –even though you never set foot in Moorepark, I think you got to know more about my PhD woes than you had ever bargained for! I lived for our spa days! Thanks to you all for the friendship down through the years.

I definitely could not have gotten through this PhD without the amazing friends I made in Moorepark. A massive thank you has to go to Noelle, AnneMarie, Aoife, Susan, Sarah H, Sarah C, Owen, Cian, Colette, Lorna and Emma. I can honestly say that my time in Moorepark would not have been the same without you all. You made my four years here so worth it. I couldn’t have asked for a
funnier or friendlier bunch. Thanks for the hilarious nights out, dinners, cinema trips, get togethers and the unforgettable trip to Amsterdam! The copious amounts of cake, non-stop laughter and friendship I received from you all helped me get through those dark days! I don’t think I’ll ever be able to look at chocolate cheesecake or brownies again without thinking of our tea breaks.

A special thank you has to go to Andy, for keeping me sane through the tough times and for the weekend breaks! Thanks for keeping me grounded, motivated and for the love and encouragement. I’d have been lost without you.

Lastly, I have to give a huge thank you to my family. Thanks to my parents, Ann and George, without whom my academic achievements would not have been possible. Thank you to my sister Emma, for your help and support and thanks to my wonderful Gran, whose unwaivering encouragement was a great source of comfort. I dedicate this thesis to all of you.

True friendship multiplies the good in life and divides its evils. Strive to have friends, for life without friends is like life on a desert island... to find one real friend in a lifetime is good fortune; to keep him is a blessing.

- Baltasar Gracian
Thesis Introduction
**Mammalian milk glycans**

Human milk is known to contain oligosaccharides and glycoconjugate structures which play an important role in protecting the neonatal gut from infection (Kunz and Rudloff 2008), establishing commensal bacteria in the gut (Kavanaugh, O’Callaghan et al. 2013) and influencing immune function (Lane, O’Callaghan et al. 2013). The presence of oligosaccharides in human milk is associated with protection of breast fed infants from infection and diarrhoea (Morrow, Ruiz-Palacios et al. 2004). However, the use of human milk as a source of these oligosaccharides is not feasible given the limited quantities available for research and commercial purposes. As a result, researchers have begun to focus on the bioactivities associated with oligosaccharides and glycoconjugates isolated from domestic animal milks, including bovine, goat, camel and ovine (Urashima, Taufik et al. 2013). Domestic animal milks are more widely available and are already acceptable for use as a source of ingredients in the public eye. Indeed, bovine milk has been shown to contain some oligosaccharide and glycoconjugate structures similar to those found in human milk (Barile, Tao et al. 2009, Albrecht, Lane et al. 2014). In particular, a component of bovine milk termed milk fat globule membrane (MFGM) is an excellent source of highly glycosylated compounds (Singh 2006, Ross, Lane et al. 2015).

The MFGM is a trilayer membrane structure which surrounds and stabilizes milk fat droplets (Singh 2006). MFGM can be sourced from dairy fractions including butter serum and buttermilk (Dewettinck, Rombaut et al. 2008). Buttermilk is produced as a by-product of the butter-manufacture process. During butter-making, cream is churned and MFGM is disrupted, breaking milk fat globules and separating the MFGM from the lipid core. The MFGM is subsequently released into the aqueous phase along with most of the water soluble milk components, all of which are recovered in the buttermilk (Dewettinck, Rombaut et al. 2008). Glycoproteins such as mucins, butyrophilin and lactadherin (Pas6/7) and neutral and acidic glycolipids (gangliosides) have
been identified as components of MFGM (Singh 2006, Dewettinck, Rombaut et al. 2008). Recently, bovine MFGM gained interest due to the bioactivities associated with its glycosylated components. Such properties include anti-infective (Parker, Sando et al. 2010), anti-bacterial (Martin, Hancock et al. 2004) and immunomodulating (Zanabria, Tellez et al. 2014) activities.

Health-promoting properties of bovine milk glycoconjugates

Previous studies have indicated that purified bovine MFGM glycoproteins may aid in prevention of pathogenic infection. For example, bovine MFGM mucin demonstrated a concentration-dependent ability to reduce the binding of enteric pathogens to Caco-2 cells in vitro (Parker, Sando et al. 2010). Furthermore, purified bovine MFGM glycoproteins were shown to reduce infection of a gastric epithelial, NCI-N87, cell line with Helicobacter pylori in vitro (Horemans, Kerstens et al. 2012). Additionally, whole bovine MFGM displayed anti-infective activity against Listeria monocytogenes in a rat model (Sprong, Hulstein et al. 2012) and reduced gastric colonization by Helicobacter pylori in a mouse model (Wang, Hirmo et al. 2001). It is likely that the glycan structures act as decoy receptors, whereby they mimic host cell receptors which bacteria bind to instead of the host cell, thereby preventing bacterial adhesion and colonization (Sharon and Ofek 2000). Decoy receptor therapy provides an alternative method to prevent bacterial infection in susceptible individuals rather than treating infections with antibiotics which can lead to development of antibiotic resistance in pathogenic bacteria.

Strains of Escherichia coli are known to cause disease globally. In particular, E. coli serotype O157:H7 can result in serious diseases including haemolytic uremic syndrome (HUS), which can lead to kidney failure, and in serious cases can even lead to death. The devastating effects of E. coli O157:H7 infections have recently been highlighted in the mass outbreaks which occurred in the USA as a result of contaminated Romaine lettuce. This outbreak led to 58
people infected, 33 hospitalisations and 3 cases of HUS development (Control and Prevention 2012). It is possible that the development of a functional glycosylated ingredient could aid in prevention of \textit{E. coli} O157:H7 infections when consumed on a regular basis. Bovine MFGM glycoproteins could potentially interact with host and bacterial cell surfaces to prevent invading pathogens from establishing themselves in the gastrointestinal tract. However, obtaining highly pure fractions of these glycoproteins for human consumption is time consuming, technically challenging and costly. Therefore, whole bovine MFGM would not only be more commercially viable when compared to purifying individual MFGM components, but also can be sourced from buttermilk. Buttermilk is produced in large quantities as a by-product of butter- manufacture and therefore targeting this dairy stream for novel activities could increase the value of milk. Considering this rationale, the ability of a bovine MFGM fraction, enriched in glycoproteins and proteins, was screened for the potential to reduce the association of \textit{E. coli} O157:H7 with human colonic adenocarcinoma cells.

**Glycosylation of MFGM through milk maturation**

The glycosylation of bovine MFGM components has been shown to be an important factor in relation to their potential health-promoting properties. For instance, sialidase treatment of bovine MFGM mucin 1 reduced the ability of the mucin to prevent \textit{E. coli} and \textit{Salmonella enterica} serovar Typhimurium binding to Caco-2 cells \textit{in vitro} (Parker, Sando et al. 2010). In addition, deglycosylation of a mucin complex isolated from human milk reduced the anti-rotaviral properties of the fraction (Yolken, Peterson et al. 1992). Previous studies have demonstrated that bovine milk glycosylation differs in abundance and variety as lactation progresses from colostrum to mature milk (Ujita, Furukawa et al. 1993, Takimori, Shimaoka et al. 2011). Therefore, further knowledge on how lactation affects the glycosylation of bovine buttermilk, a source of MFGM, may be critical in order to identify particular lactation time
points with a high abundance of relevant glycosylation. By targeting certain time points in lactation from which MFGM is sourced, it may be possible to maximise the glycosylation of its components, thereby increasing the abundance and variety of glycosylation of the resulting functional ingredients. The use of such a strategy may increase the bioactivity of the fraction. Such a fraction could be used to supplement infant formula or baby food in order to increase the health benefits of formula in an effort to further emulate the gold standard that is breast milk. Previous studies have identified changes in bovine milk glycosylation as lactation progresses (Ujita et al., 1993, Wilson, Robinson et al. 2008) however, to the best of our knowledge, no study to date has investigated the changes in glycosylation over the first 10 days of milk maturation, nor compared these to multiple samples of mature milk taken at various time points. Additionally, such studies have not been conducted using multiple animals which may be of interest as biological variation may occur in the type and abundance of glycosylation of MFGM components between individual animals of the same species. Lectin arrays provide a platform from which to screen a large number of samples with a large number of lectins to facilitate the assessment of a wide variety of glycosylation profiles. Such methods have previously been used to observe trends in glycosylation of individual bovine milk glycoconjugates over the course of milk maturation (O'Riordan, Gerlach et al. 2014). In this thesis, lectin microarrays were employed to identify potential changes in glycosylation of buttermilk components which may occur over the course of bovine milk maturation.

**Glycosylation of bacterial cell surfaces**

The aim of decoy receptor therapy is to use decoy receptors such as glycans to adhere to pathogens and prevent pathogenic binding to host cell receptors (Zopf and Roth 1996). Therefore, knowledge of *E. coli* surface glycosylation and preferential host receptors would be of great benefit when identifying compounds to use as anti-infective agents. Additionally, understanding the
relationship between the pathogen’s surface receptors and the host could aid in engineering anti-infective compounds to target these pathogens. Bacteria possess a variety of oligosaccharide moieties on their surface which can be found as components of structures including lipopolysaccharide (Caroff and Karibian 2003), extracellular capsule (Whitfield and Roberts 1999), peptidoglycan layer (Adibekian, Stallforth et al. 2011) flagella and adhesins (Reid, Fulton et al. 2010, Iwashkiw, Vozza et al. 2013). These function in pathogenic processes including adhesion and colonization, biofilm formation and immune evasion (Schmidt, Riley et al. 2003, Itoh, Wang et al. 2005, Severi, Hood et al. 2007).

Strains of enterohaemorrhagic *E. coli* (EHEC) and enteropathogenic *E. coli* (EPEC) differ in their tissue tropisms, as EHEC colonizes and infects the colon while EPEC preferentially colonizes the small intestine (Mellies, Barron et al. 2007). This could be as a result of the different glycosylation patterns on their surfaces (Georgia-Korea, Ghigo et al. 2011), as well as varying preferences for host intestinal cell receptors. Methods previously used for identification of bacterial surface glycosylation have included nuclear magnetic resonance, electrophoresis, mass spectrometry and high performance liquid chromatography (Reid, Fulton et al. 2010). Lectin microarray approaches have also been employed. For instance Hsu et al. (2006) identified the surface glycosylation of three *E. coli* strains, however this study used a limited number of lectins (21) to screen the bacterial surface glycosylation. Here, *E. coli* O157:H7 strains (EHEC) and EPEC strains were screened in order to identify if their different localisation preferences were due to differing surface glycosylation or receptor preferences. Knowledge of *E. coli* surface glycosylation and glycan receptor binding preferences could aid in the discovery of anti-infective agents aimed at protecting consumers from these highly infective strains.
Thesis Introduction

Transcriptomic modulation by MFGM

Many bioactivities have been associated with bovine MFGM including anti-infective (Liu, Yu et al. 2012, Ross, Lane et al. 2016), anti-carcinogenic (Zanabria, Tellez et al. 2013), anti-rotaviral (Kvistgaard, Pallesen et al. 2004), and immunomodulatory (Zanabria, Tellez et al. 2014) properties. The transcriptional response of human cells to free milk oligosaccharides has been investigated previously, and studies have even compared the response to both human and bovine milk oligosaccharides (Lane, O'Callaghan et al. 2013). Since lactational stage may influence the abundance and type of glycosylation of bovine buttermilk, the associated bioactivities of bovine buttermilk may also be affected. Considering this, the aim here was to investigate if buttermilk generated from colostrum and mature milk could have different transcriptomic effects on human intestinal epithelial cells in vitro. Such a study should shed further light on the potential effects the two buttermilk fractions have on a variety of biological processes. By studying the differential regulation of genes, an insight may be gained on the possible mechanisms of actions by which buttermilk components exert their activity at the transcriptional level.

The work described in this thesis has focused on the bioactivities of bovine MFGM components focusing on the anti-infective capabilities of the fraction against E. coli strains and its potential immune-modulating properties on human intestinal epithelial cells in vitro. Additionally, the research examined how lactation stage may affect the glycosylation of bovine buttermilk, a source of MFGM, and also compared the surface glycosylation and receptor preferences of EPEC and EHEC strains to aid in the generation of decoy receptors to target specific pathogenic strains. This work highlights the bioactivities and commercial importance of bovine MFGM and its potential as a functional ingredient.
References


longum subsp. infantis to milk oligosaccharides increases adhesion to epithelial cells and induces a substantial transcriptional response." PloS one 8(6): e67224.


Chapter I
Literature review

The milk fat globule membrane: A potential source of health-promoting glycans

Sarah A. Ross, Jonathan A. Lane, Michelle Kilcoyne, Lokesh Joshi and Rita M. Hickey

Published in ‘Biotechnology of Bioactive Compounds: Sources and applications’. Chapter 26
Chapter I

Abstract

The milk fat globule membrane (MFGM) which surrounds the lipid globules in milk is assembled and secreted by the epithelial cells of the mammary gland and consists of a complex mixture of proteins, enzymes, glycoproteins, neutral lipids, polar lipids and glycolipids. Many nutritional and health promoting properties have been linked with these MFGM components including anti-carcinogenic and anti-cholesterol activity as well as improving long-term memory. Glycosylated MFGM proteins and lipids in particular, have been accredited with a number of health promoting properties. For instance, these glycoconjugates are thought to act as specific bacterial and viral ligands which, when present in the gastrointestinal tract of young mammals, contribute to the prevention of infection. MFGM glycoproteins have also been linked with the prevention of inflammatory diseases such as multiple sclerosis. In this chapter, we will provide an overview of the composition, structure and biological activities of glycosylated MFGM compounds sourced from various mammalian milks, using human MFGM for comparative purposes. We will also discuss the potential use of MFGM as a source of health-promoting glycans which could be incorporated into functional foods.
Chapter I

Introduction

The mammalian milk glycome has recently gained increased interest due to health promoting properties which have been associated with both free and conjugated glycans. Research in this area has mainly been focused on the structural and functional characterisation of human milk oligosaccharides (HMO) and glycoproteins. These molecules are highly abundant in human milk and therefore can be obtained in significant quantities for both structural and biological characterisation. For example, mature human milk contains 5-15 g/L of free oligosaccharides while human colostrum may contain up to 50 g/L (Kunz, Rudloff et al. 1999, Bode 2006, Kunz and Rudloff 2008). Initially, the majority of researchers focused on characterising the prebiotic activities of HMO (György, Norris et al. 1954, Ward, Niñonuevo et al. 2006, LoCascio, Ninonuevo et al. 2007). However, the fact that 90% of all HMO were found intact in infant faeces suggested that these compounds had additional effects once consumed (Bode 2009). Consequently, researchers discovered that HMO had other health-promoting properties such as the ability to promote development of the immature immune system, regulate immune responses, contribute towards brain development and protect against infection (Wang, McVeagh et al. 2003, Newburg, Ruiz-Palacios et al. 2005, Bode 2006, Kunz and Rudloff 2006, Bode 2012, Kavanaugh, O’Callaghan et al. 2013, Lane, O’Callaghan et al. 2013). Indeed, HMO have protective effects against both viral and bacterial infection as they resemble the complementary receptors present on human cells and can act as decoy receptors, thus preventing adhesion and subsequent colonisation and infection (Sharon and Ofek 2000). For example, Coppa et al. (2006) demonstrated that HMO, including 6′-sialyllactose and 3-fucosyllactose, can inhibit the adhesion of Salmonella fyris, Vibrio cholerae and enteropathogenic Escherichia coli serotype O119 to the human colonic adenocarcinoma intestinal epithelial cell line, Caco-2. In addition, the authors reported inhibition of E. coli and V. cholerae adhesion to Caco-2 cells by the neutral high molecular weight glycoconjugate fraction. The ability of HMO to reduce bacterial colonisation has also been demonstrated in in vivo
trials. Mysore et al. (1999) demonstrated that Helicobacter pylori colonisation in rhesus monkeys treated with the milk oligosaccharide 3′-sialyllactose was dramatically reduced. Furthermore, there is clear evidence that breast-fed infants acquire fewer diseases or less severe infections, such as diarrhoeal diseases, when compared to formula-fed infants (Duffy, Byers et al. 1986, Newburg, Peterson et al. 1998, Morrow, Ruiz-Palacios et al. 2004), a property which has been associated with the protective nature of HMO. The immune-regulatory activities of HMO have been widely reported. For example, roles for HMO in lymphocyte maturation and cytokine production (Velupillai and Harn 1994, Eiwegger, Stahl et al. 2004) have been demonstrated. Interestingly, the levels of sialic acid, a negatively charged sugar which caps the terminus of many HMO structures, on brain-related glycoproteins and gangliosides have been linked to improved cognitive function and learning ability. The beneficial role of HMO in brain development has been demonstrated by Wang et al. (2003) who reported increased concentrations of sialylated gangliosides in the brain of breast-fed infants compared to formula-fed infants (Wang, McVeagh et al. 2003).

Human milk glycoconjugates have also demonstrated health-promoting properties similar to those associated with free milk oligosaccharides. For example, human milk gangliosides have been shown to bind and neutralize bacterial toxins including H. pylori vacuolating toxin (Wada, Hasegawa et al. 2010) and cholera toxin (Iwamori, Takamizawa et al. 2008). Additionally, the human milk gangliosides GM1 and GM3 have demonstrated an ability to prevent binding of enterotoxigenic E. coli to Caco-2 cells in vitro (Idota and Kawakami 1995) and human milk fat globule membrane (MFGM) mucin has prevented rotavirus-associated gastroenteritis in a mouse model (Yolken, Peterson et al. 1992). Schroten et al. (1992) also reported that S-fimbriated E. coli adherence to buccal epithelial cells can be inhibited by human MFGM. This activity was attributed to the presence of heavily glycosylated mucins. Furthermore, human milk mucins have been shown to inhibit Salmonella enterica serovar Typhimurium invasion of human Caco-2 and FHs 74 Int
(human small intestine) cells (Liu, Yu et al. 2012) while Newburg et al. (1998) demonstrated the human milk glycoprotein, lactadherin (LDH), provides protection to infants from rotavirus infection. These studies highlight the important role HMO and glycoconjugates play in the early stages of life.

Although human milk is clearly a source of biologically active oligosaccharides and glycoconjugates, the difficulty is that it is not readily available in large enough quantities to be a commercially viable source. Consequently, research has begun to focus on domestic animal milks as an alternative source of functional oligosaccharides and glycoconjugates (O’Riordan, Kane et al. 2014). The use of mature domestic animal milks as a source of free milk oligosaccharides has proven challenging as only trace quantities of oligosaccharides are present in these milks. For example, the oligosaccharide content of bovine, ovine and caprine milk is 0.03–0.06 g/L, 0.02–0.04 g/L and 0.25–0.30 g/L, respectively (Urashima, Saito et al. 2001, Lane, Mehra et al. 2010). This has lead researchers to focus on fractions from dairy sources, which are rich in glycoconjugates, as an alternative to free oligosaccharides as they may demonstrate similar biological activity. For example, buttermilk is a sustainable source of MFGM-associated glycosylated ingredients. As a by-product of butter production, buttermilk is low in cost and is available in large quantities. In the United States, 610 million kg of butter is produced annually resulting in the production of 35.4 million kg of buttermilk which is condensed or evaporated (Jiménez-Flores and Brisson 2008). These figures highlight the potential of buttermilk, and in particular buttermilk derived from bovine milk, as a sustainable source of functional food ingredients. Furthermore, bovine MFGM-associated ingredients have demonstrated health-promoting activities similar to those of HMO, e.g., prevention of pathogen colonisation in *in vitro* and *in vivo* (Sanchez-Juanes, Alonso et al. 2009, Sprong, Hulstein et al. 2012). Considering this, our review will focus on the composition and functional properties of mammalian MFGM, with particular attention paid to human and bovine MFGM, and the important roles glycosylated MFGM components could play as nutraceuticals such as providing
Chapter I

protection against infection by enteric pathogens for immune-compromised individuals including newborns and the elderly.
Chapter I

MFGM composition

The MFGM is composed of lipids and proteins which surround and stabilise milk fat droplets (Cebo, Caillat et al. 2010). These fat droplets form within the endoplasmic reticulum (ER) of the mammary gland epithelial cells (Zaczek and Keenan 1990) and move towards the apical membrane. The droplets then bud into the lumen of the alveolar epithelial cells and as they do so, become surrounded by the apical membrane which forms the main part of the MFGM (Mather and Keenan 1998). The average size of the milk fat globule (MFG) differs between species. For example, bovine MFGs have been found to be slightly larger than those of goat milk (4.2 and 3.7 µm respectively) (Guri, Griffiths et al. 2012). In the absence of MFGM, the lipid droplets in milk would exist as aggregates rather than in a dispersed form (Spitsberg 2005). For instance, if the MFGM is disrupted by physical agitation such as churning, butter is formed due to lipid aggregation while the aqueous phase of the MFGM forms the buttermilk as a by-product (Spitsberg 2005, Harrison 2006).

MFGM consists of three parts; the first is an inner monolayer derived from the ER which surrounds the lipid droplet. This monolayer is surrounded by a proteinaceous coat which is surrounded by an outer bilayer (Keenan and Mather 2002, Evers, Haverkamp et al. 2008). Thus, the MFGM is composed of a trilayer structure (Abrahamse, Minekus et al. 2012), with the outer bilayer originating from the apical membrane (Mather and Keenan 1998, Reinhardt and Lippolis 2006). When the fat globule is extruded into the alveolar epithelial cell lumen, a 10-20 nm space exists between the inner monolayer and the nascent outer bilayer (Wooding 1971, Heid and Keenan 2005). It is within this space that the proteinaceous coat is formed and that the cytoplasmic organelles can be found (Wooding 1971, Heid and Keenan 2005). Interestingly, studies have shown that the proteinaceous coat of the MFGM of human and bovine milk have similar carbohydrate and amino acid compositions (Freudenstein, Keenan et al. 1979). Cytoplasmic crescents often form in the outer bilayer due to the trapping of cytoplasm between the two layers. This occurs in the milk of many mammals (Huston and Patton 1990, Dewettinck, Rombaut et al. 2008). The size
of the crescents can vary and studies by Evers et al. (2008) have shown that these crescents can reach a volume exceeding that of the fat globule itself. Both the inner and outer layers of the MFGM are composed of polar lipids and proteins. The bilayer contains peripheral and transmembrane proteins as well as lipid components such as cholesterol. In addition, many of the lipids and proteins are heavily glycosylated (Dewettinck, Rombaut et al. 2008).

MFGM is heterogenous, with some areas of the membrane containing proteins which are evenly distributed, while other areas lack proteins (Robenek, Hofnagel et al. 2006). In addition, distinct domains of the MFGM are characterised by different shapes such as ridges and elongated or rounded bumps (Robenek, Hofnagel et al. 2006). The external surface of human and horse MFGM is rough due to the presence of mucin-1 (MUC1)-containing filaments whereas the absence of these filaments on the surface of the MFG of mammals such as sheep, cow and goat result in a smooth surface (El-Loly 2011). These filaments extend outwards from the human MFGM to a distance of 0.5 μm, with some extending as far as 1 μm and have been shown to bind to E. coli, neutralising the threat of bacterial disease (Buchheim, Welsch et al. 1988). In addition, portions of the outer bilayer appear to be absent in some MFG, (Evers et al. 2008) as visualised where neither lipophilic dyes nor the lectin wheat germ agglutinin (WGA) fluorescent probes, which bind to the carbohydrate residues sialic acid, were incorporated in some areas of human, bovine and ovine milk fat globules.

Fong et al. (2007) recently reported on the concentrations of MFGM in bovine milk and found 3.6±0.3 g/L of MFGM in cream, with the protein and lipid fractions making up an estimated 22.3±1.5% and 71.8±1.7%, respectively. The polar lipids found in MFGM are glycerophospholipids - such as phosphatidylethanolamine, phosphatidylcholine, phosphatidylserine and phosphatidylinositol - and sphingolipids - such as sphingomyelin (SM). There are also small quantities of glycosphingolipids in the form of lactosylceramide (LacCer), glucosylceramide and gangliosides (Christie, Noble et al. 1987, Sprong, Hulstein et al. 2002). MFGM lipids are arranged asymmetrically with
Chapter I

SM, phosphatidylcholine (which contains the polar choline) and the glycolipids present on the outside of the globule membrane, while phosphatidylethanolamine, phosphatidylerine and phosphatidylinositol are found inside the membrane (Dewettinck, Rombaut et al. 2008). Milk glycolipids are almost exclusively located in the outer layer of the MFGM (Georgi, Bartke et al. 2013).

MFGM proteins contribute 1-2% of the total protein content in milk (Riccio 2004). These proteins, like the lipids, are arranged asymmetrically in the membrane. Such proteins include adipophilin (ADPH), which can be found in the inner monolayer as it has a high affinity for triglycerides, and xanthine oxidoreductase (XOR), which is found between the inner monolayer and outer bilayer. Butyrophilin (BTN) is a transmembrane protein found in the outer bilayer with its cytoplasmic tail in the proteinaceous coat (Heid and Keenan 2005, Dewettinck, Rombaut et al. 2008). It is thought that these three proteins (ADPH, XOR and BTN) interact to form a complex that connects the inner and outer membrane and aids in the binding of the bilayer to the MFG during the extrusion of the globule from the alveolar epithelial cell (Mather and Keenan 1998, Mather 2000, Cavaletto, Giuffrida et al. 2004). An overview of the MFGM structure is given in Figure 1.

Many different functions are associated with MFGM proteins including protein synthesis/folding, transport, and metabolism. Other functions include cell signalling and membrane/protein trafficking and have been attributed to almost half of the bovine MFGM proteins (Reinhardt and Lippolis 2006). Differences exist in the type and abundance of proteins in the MFGM of different species (Mather 2000, Cavaletto, Giuffrida et al. 2004, Fong, Norris et al. 2007, Zamora, Guamis et al. 2009, Cebo, Caillat et al. 2010). Six of the minor bovine MFGM proteins, including apolipoprotein E, apolipoprotein A1 and heat shock proteins are homologous to the human MFGM proteins (Fong, Norris et al. 2007).
Chapter I

It is evident that bovine MFGM is a rich source of proteins and lipids, many of which are glycosylated. Indeed, MFGM contains lipids and proteins in a 1:1 ratio (Spitsberg 2005) and their diverse glycosylation could indicate that these milk components may possess many health promoting properties. Overall, bovine MFGM glycoproteins differ from their human counterparts in their glycosylation. For instance, bovine MFGM contains primarily core-1 O-linked oligosaccharides while those of human MFGM are predominantly core-2-type structures. Furthermore, human MFGM N-linked oligosaccharides contain terminal fucose (Fuc) residues which are not present in that of bovine MFGM (Wilson, Robinson et al. 2008). In addition, bovine MFGM glycoproteins contain a high degree of N-linked oligosaccharides terminating in N-acetylgalactosamine (GalNAc), in the form of N, N’-diacetyllactosamine (LacdiNAc, GalNAc-\(\beta-(1\rightarrow4)\)-GlcNAc). For example, the oligosaccharide chains of bovine BTN and cluster of differentiation 36 (CD36) contain 37% and 28% LacdiNAc, respectively (Sato, Furukawa et al. 1993). Similarities also exist between bovine and human glycoprotein oligosaccharides. For example, both contain bi-, tri-, and tetra-antennary complex N-linked structures which are sialylated (Wilson, Robinson et al. 2008). In the next section, we discuss the structure and function of MFGM glycolipids and glycoproteins and highlight the potential of these bioactives for inclusion in functional foods.
Chapter I

Glycoconjugates of the MFGM

Glycolipids

Approximately 70% of the total glycolipid content of bovine milk is found to be associated with the MFGM (Newburg and Chaturvedi 1992). These glycolipids can be divided into two groups, neutral and acidic (or gangliosides) (Jensen 2002), and are summarised in Table 1. The neutral glycolipids of bovine MFGM are composed mainly of LacCer (65%) and its precursor glucosylceramide (35%), both of which contain non-hydroxylated fatty acids. The neutral glycolipids of human MFGM are mainly galactosylceramide, glucosylceramide and LacCer and 20% of the fatty acid residues of these glycolipids are hydroxylated (Newburg and Chaturvedi 1992). Additionally, ovine and human MFGM contain globotriosylceramide and globotetraosylceramide but bovine MFGM does not (Zancada, Sanchez-Juanes et al. 2010). The neutral glycolipids consist of one or more carbohydrate residues linked to the lipid moiety ceramide (a sphingosine linked to a fatty acid through an amide linkage). For example, galactosylceramide consists of a galactose (Gal) residue connected to a ceramide via a glycosidic linkage while addition of a \(\beta\)-linked Gal to a hydroxyl group of the glucose (Glc) residue forms LacCer (Schnaar, Suzuki et al. 2009). Acidic glycolipids (gangliosides) are composed of a ceramide attached via a glycosidic linkage to an oligosaccharide with at least one sialic acid and various other residues attached (Khatun, Gayen et al. 2013). Approximately 90% of gangliosides in milk are associated with the MFGM (Colarow, Turini et al. 2003). The gangliosides of mammalian MFGM consist of monosialoganglioside (GM) 1, GM2, GM3, disialoganglioside (GD) 1A, GD1B, GD2, GD3, and trisialoganglioside GT (Jensen 2002). These structures are mainly composed of long chain fatty acids but have also been found to contain medium, monounsaturated and polyunsaturated fatty acids. Most research performed to date has concentrated on human and bovine glycolipids but some analyses have been carried out on other mammalian milk glycolipids. For instance, Zancada et al. (2010) determined the neutral glycosphingolipid content of ovine whole milk. The
profile was found to be similar to that of human milk neutral glycosphingolipids. In addition, Iwamori et al. (2008) demonstrated that gangliosides belonging to the biosynthetic pathway of the ganglio series of gangliosides for neural tissues were absent in human and bovine milk but present in goat milk. Interestingly, the overall concentration of glycolipids in human milk decreases during the transition from colostrum to mature milk, with GD3 being the major ganglioside found in colostrum (Uchiyama, Sekiguchi et al. 2011, Georgi, Bartke et al. 2013), which could suggest a role in protecting the newborn from infection. Additionally, ganglioside concentrations were found to change in the murine mammary gland over the course of lactation with the concentration of acidic glycosphingolipids reducing over this time. Furthermore, it was found that GD1a was the predominant ganglioside found in the MFGM of colostrum (Momoeda, Momoeda et al. 1995).

LacCer and GD3 from bovine whole milk and MFGM was shown to bind to four strains of enterotoxigenic E. coli (Sanchez-Juanes et al., 2009). These binding activities may play a role in prevention of bacterial adherence to human gastrointestinal cells. The glycoconjugates act as soluble receptors and prevent bacterial binding to host cell receptors. Interestingly, the glycolipids isolated from whole milk demonstrated a greater binding capacity than those isolated from MFGM, which was most likely due to conformational differences as a result of the difference in fatty acid composition (Sanchez-Juanes, Alonso et al. 2009). Human milk GM1, GM3 and GD3 inhibit adhesion of pathogenic bacteria such as enterotoxigenic E. coli, Campylobacter jejuni, and Listeria monocytogenes to human intestinal Caco-2 cells (Idota and Kawakami 1995, Salcedo, Barbera et al. 2013). Ovine milk neutral glycosphingolipids have also demonstrated binding ability to certain EPEC and UPEC strains (Zancada, Sanchez-Juanes et al. 2010).

Milk glycolipids also confer protection to the host by binding to bacterial toxins and preventing outcomes such as cell membrane disruption, inhibition of protein synthesis, fever and diarrhoea. Human milk GM1 had inhibitory activity against cholera toxin in vivo in a rabbit intestine model and
against *E. coli* heat labile enterotoxin *in vitro* (Otnaess, Laegreid et al. 1983). In addition, the bovine brain gangliosides GM1, GM2, GM3, GD1a, GD1b, GD3 and GT have been shown to neutralise the vacuolating cytotoxin (VacA) of *H. pylori*. Internalisation of the toxin by human gastric cells was inhibited in the presence of GM1. Interestingly, lyso-gangliosides (lyso-GM1, lyso-GM2 and lyso-GM3), which lack fatty acids, bound VacA while 3’sialyllactose, a carbohydrate moiety of GM3, alone did not inhibit the activity of VacA. These results suggest that the carbohydrate portion of the gangliosides alone cannot inhibit VacA but is an important overall factor in neutralisation of the toxin (Wada, Hasegawa et al. 2010). Although the gangliosides mentioned in the latter study were brain-derived, the same effect may result using the carbohydrate moiety of milk-derived gangliosides. Further studies are required to determine the value of milk-derived glycolipids.

Protection of the host from viral infection has also been linked to milk glycolipids. For example, human milk GM3 and GD3 inhibit adhesion of reovirus to L cells (murine areolar and adipose cells) and HeLa cells (human cervical epithelial cells) (Iskarpatyoti, Morse et al. 2012). In this study, GM3 and GD3 were preincubated with reovirus before inoculation of the cell line with the virus-ganglioside mixture and infectivity was subsequently determined by immunofluorescence. In addition, 3’sialyllactose inhibited infection of reovirus with L cells and HeLa cells, when preincubated with reovirus. However, the anti-adhesive capability was lower than that of GM3 and GD3, indicating the carbohydrates play a role in viral inhibition but that the molecular conformation of membrane-bound ganglioside may be an important factor in the anti-viral activity of the carbohydrate (Iskarpatyoti, Morse et al. 2012). Furthermore, human milk GM2 binds respiratory syncytial virus and inhibits adhesion of the virus to Hep-2 (human larynx) cells (Portelli, Gordon et al. 1998). These studies highlight the potential use of MFGM glycolipids, however, it is obvious that further studies are required to determine the anti-viral activity associated with these compounds. Indeed, the limited knowledge
on the *in vivo* bioactive potential of glycolipids associated with MFGM warrants further investigation to realise their commercial potential.

**Glycoproteins**

More information exists regarding the glycoprotein fraction of MFGM compared to the glycolipid fraction. Some of the most researched glycosylated proteins of the MFGM are the mucins, BTN, CD36 and LDH (PAS6/7). The “PAS” name given to the MFGM glycoproteins relates to their positive reaction with periodic acid-schiff (PAS) stain as a result of their glycosylation. An overview of the major MFGM glycoproteins is presented in Table 2, while their structural properties and beneficial effects are discussed individually below.

**Mucins**

Mucins of the MFGM are type 1 integral-membrane proteins (Jonckheere, Skrypek et al. 2012). They contain a variable number of tandem repeat (VNTR) domains that are rich in the amino acids serine, threonine and proline. The hydroxyl groups of serine and threonine function as glycosylation sites while the proline confers rigidity to the structure of the protein and also extends its structure from the cell surface (Patton, Gendler et al. 1995, Moran, Gupta et al. 2011). The extent of glycosylation of each protein differs and depends on the number of repeats in the VNTR. Thus, mucins are polymorphic proteins in species such as human and cow (Huott, Josephson et al. 1995). Other features that are characteristic of the membrane spanning mucins are a cytoplasmic C-terminus and one membrane spanning domain (Pallesen, Pedersen et al. 2007). On average human milk contains 729 ± 75 µg/mL of mucin glycoprotein (Peterson, Hamosh et al. 1998). The levels of mucins in milk samples are dynamic and change depending on the stage of lactation. For example, MUC1 and mucin-15 (MUC15) concentration increased 7.7- and 7.4-fold, respectively, in bovine milk samples taken 7 days post-parturition compared to colostrum (Reinhardt and Lippolis 2008). Degradation of the
protein is avoided due to the high degree of glycosylation of the proteins (Peterson, Hamosh et al. 1998).

The three main mucins which have been characterised in bovine MFGM are MUC1, MUC15 and mucin X (MUCX) (Pallesen, Andersen et al. 2001, Liu, Erickson et al. 2005, Pallesen, Pedersen et al. 2007). There are similarities between the mucins of different species (e.g., bovine and human MUC1 share 76 and 93% amino acid similarity at the cytoplasmic and transmembrane domains, respectively). Interestingly, the least conserved identity occurs at the VNTR domains (Sando, Pearson et al. 2009). Differences also exist, for instance, fucosylated oligosaccharides are found on human mucins but fucosylation is less prevalent in bovine mucins (Wilson, Robinson et al. 2008). Mucins are highly O-glycosylated proteins, as shown by high levels of binding of the lectin Jacalin, which prefers binding to Gal residues, to mucins such as MUC1 and MUCX (Liu, Erickson et al. 2005), and contain N-linked oligosaccharides to a lesser extent. The N-linked glycosylation of MUC1 has been demonstrated by successful cleavage of N-linked oligosaccharides from MUC1 using the enzyme PNGase F (Pallesen, Andersen et al. 2001).

The five major bovine MUC1 isoforms vary in molecular mass, ranging from 156 to 193 kDa (Huott, Josephson et al. 1995, Pallesen, Andersen et al. 2001), and account for 1.21% (w/w) of the total bovine MFGM proteins (Kvistgaard, Pallesen et al. 2004). MUC1 is extensively glycosylated with bovine and human MUC1 glycans contributing ≥50% of the total molecular mass of the protein (Huott, Josephson et al. 1995, Patton, Gendler et al. 1995, Pallesen, Andersen et al. 2001). It is mainly O-glycosylated but five potential N-glycosylation sites have been identified for bovine MUC1 (Pallesen, Andersen et al. 2001). One of the most common carbohydrate residue associated with the MUC1 glycoprotein is sialic acid, followed by Gal. Sialic acid accounts for approximately 30% (molar %) of the total carbohydrate content of the bovine glycoprotein (Huott, Josephson et al. 1995, Patton, Gendler et al. 1995, Pallesen, Andersen et al. 2001) and 11% (molar %) of the total carbohydrate content of the human milk glycoprotein (Shimizu and
Chapter I

Yamauchi 1982, Patton, Gendler et al. 1995). Other carbohydrate residues associated with bovine and human MUC1 include Gal, mannose (Man), N-acetylglucosamine (GlcNAc) and GalNAc, while fucosylation of N-linked oligosaccharides has been detected at low levels (Shimizu and Yamauchi 1982, Pallesen, Andersen et al. 2001, Liu, Erickson et al. 2005, Sando, Pearson et al. 2009).

In contrast to MUC1, MUCX is more loosely associated with the MFGM, leading to its presence in milk fractions such as skim milk. It contains fewer N-linked oligosaccharides compared to MUC1 (Liu, Erickson et al. 2005). The use of Morrisey’s silver staining method (which stains glycoproteins more rapidly when compared to non-glycosylated proteins), demonstrated that bovine MUCX stained more rapidly than bovine MUC1 indicating that MUCX may have a higher carbohydrate content compared to MUC1 (Liu, Erickson et al. 2005). A mucin termed mucin 4 has also been identified in human MFGM but it is likely that this mucin is MUCX (Zhang, Perez et al. 2005).

MUC15 has been identified in mammalian milks including bovine, human, ovine and caprine milk (Pallesen, Pedersen et al. 2008) and it makes up 1.5% (w/w) of the total bovine MFGM protein content. It too is highly O-glycosylated, and N-glycosylated to a lesser extent (Pallesen, Pedersen et al. 2007). Fifteen potential N-linked glycosylation sites have been identified for this glycoprotein (Pallesen, Berglund et al. 2002). It has a molecular weight of 94 kDa in bovine MFGM, 65% of which is attributed to the carbohydrate content of the glycoprotein (Pallesen, Pedersen et al. 2007), and approximately 150 kDa in human MFGM, which implies the presence of even greater quantities of carbohydrate. Human and bovine MUC15 display 59% and 87% similarity in the extracellular and cytoplasmic/transmembrane domains, respectively (Pallesen, Pedersen et al. 2008). Fuc, GalNAc, GlcNAc, Gal, Man and sialic acid are present in the ratio 1:4:6:5:4:5. The O-linked oligosaccharides of bovine MFGM are largely mucin core type 1 structures, with the extended O-linked structures containing high concentrations of GlcNAc, which implies long structures (Pallesen, Pedersen et al. 2007).
Glycosylation patterns of mucins are subject to change depending on several factors including location in the host and infection or disease status. For example, prolonged infection with \textit{H. pylori} is associated with increased sialic acid and decreased Fuc levels of mucin. The glycosylation returns to normal after elimination of the bacteria from the system (Ota, Nakayama et al. 1998).

The protection mucins confer through inhibition of pathogenic infection has been well documented (Schroten, Hanisch et al. 1992, Yolken, Peterson et al. 1992). MUC1 binds fimbriated \textit{E. coli}, preventing the bacteria from binding to buccal epithelial cells (Schroten, Hanisch et al. 1992). Bovine MUC1 also displays anti-viral properties and has been shown to decrease neuraminidase-sensitive rotavirus infection in MA104 cells (embryonic monkey kidney cells) (Kvistgaard, Pallesen et al. 2004). The human mucins, MUC1 and MUC4, have anti-adhesive and anti-invasive activity against \textit{S. enterica} serovar Typhimurium (Liu, Yu et al. 2012) The oligosaccharides are known to contribute to the anti-adherent activities associated with the mucins. Binding of Norwalk virus to gut epithelial cells was shown to be inhibited by $\alpha$-(1→2)-linked Fuc of MUC1 from human milk (Ruven-Cloquet, Mas et al. 2006). Sialic acid on MUC1 from bovine milk has been shown to play a role in prevention of adhesion of bacteria to Caco-2 cells \textit{in vitro} (Parker, Sando et al. 2010) and mucins also play a role in protection against human immunodeficiency virus (HIV). The Lewis x moieties of human milk MUC1 bind to dendritic cell receptors and block HIV attachment to the cell, preventing transmission of HIV to T cells (Saeland, de Jong et al. 2009). However, further \textit{in vivo} studies involving the mucins of MFGM are required to demonstrate the efficacy of the effects described above.

\textit{Butyrophilin}

BTN in milk is often called BTN1A1 to differentiate it from the other BTN genes, of which there are seven (Rhodes, Stammers et al. 2001). Bovine BTN1A1 is an acidic and hydrophobic protein (Ishii, Aoki et al. 1995) that
Chapter I

contains 526 amino acids (Jack and Mather 1990). It is a member of the immunoglobulin superfamily (Gardinier, Amiguet et al. 1992, Banghart, Chamberlain et al. 1998), containing two immunoglobulin-like domains and a cytoplasmic C terminal tail (Robenek, Hofnagel et al. 2006); in human milk, it is expressed predominantly as protein variants of approximately 66 kDa (Quaranta, Giuffrida et al. 2001). It is a type 1 transmembrane protein (Jack and Mather 1990) found in a number of membranes such as the plasma membrane of insect cells (Banghart, Chamberlain et al. 1998) and the apical membrane of epithelial cells of the mammary glands (Robenek, Hofnagel et al. 2006), as well as the MFGM. While associated with the MFGM, it undergoes little proteolysis during and after secretion of milk (Banghart, Chamberlain et al. 1998). Studies by Ishii et al. (1995) demonstrated that BTN1A1 associates with XOR through its C terminal domain or, more specifically, through its B30.2 domain. The interaction between BTN1A1 and XOR is deemed important as it stabilises XOR binding to the MFGM (Jeong, Rao et al. 2009). It has been suggested that the BTN1A1-XOR complex associates with ADPH, further aiding in the formation of complete MFGM (Mather and Keenan 1998).

BTN differs slightly between species (e.g., bovine BTN has a molecular mass of approximately 64 kDa while those of horse, goat and camel are approximately 70 kDa, 67 kDa and 63 kDa, respectively) (Cebo and Martin 2012, Cebo, Rebours et al. 2012). The differences in molecular mass are directly associated with glycosylation levels. BTN1A1 makes up 20-40% of the total protein content of bovine MFGM (Banghart, Chamberlain et al. 1998). In human milk, BTN concentration was found to be 41±3 µg/mL (Peterson, Hamosh et al. 1998). Interestingly, in humans there is no common trend in BTN concentrations as lactation progresses, which is highlighted by the fact that BTN levels in human milk decrease in some women and increase in others over the course of lactation (Peterson, Hamosh et al. 1998). In contrast, BTN in bovine milk increases 3.2-fold in milk samples taken on day 7 compared to colostrum samples (Reinhardt and Lippolis 2008).
Chapter I

Bovine BTN contains three potential N-linked glycosylation sites, two in the N terminal domain and one in the C terminal domain, but does not contain O-linked oligosaccharides (Valivullah and Keenan 1989). The N-linked oligosaccharides are found only at the two N terminal sites at Asn55 and Asn215. (Sato, Takio et al. 1995). The glycoprotein displays site-specific glycosylation (e.g., LacdiNAc is present in approximately 50% of the oligosaccharide structures present at the Asn55). In contrast, this moiety is not found at Asn215 (Sato, Takio et al. 1995).

BTN displays a variety of functions, although little is known of the role glycosylation plays in these functions. Sequencing of human and bovine BTN identified 84% sequence similarity and revealed a possible receptor function by sequence correlation (Taylor, Peterson et al. 1996). BTN also plays an important role in regulation of milk secretion (Ogg, Weldon et al. 2004) and has been found on the inner monolayer of the MFGM. A second model of milk lipid secretion has been proposed where BTN1A1-BTN1A1 interactions between the proteins of the two layers can aid in formation of the complete MFGM (Robenek, Hofnagel et al. 2006). It is through such interactions that the milk lipid droplets become coated in membrane and are eventually extruded from the cell (Jack and Mather 1990, Jeong, Rao et al. 2009).

BTN has been associated with a variety of health benefits. For instance, bovine MFGM BTN has been linked to the protection and amelioration of symptoms of experimental autoimmune encephalomyelitis (EAE) in mice - the animal model for multiple sclerosis (Mana, Goodyear et al. 2004). However, conflicting reports exist as BTN has also been shown to induce EAE (Stefferl, Schubart et al. 2000). To date, BTN has not been shown to possess anti-infective activities and the importance of the oligosaccharides and the role they play in bioactivity has yet to be established.
Chapter I

Lactadherin

Lactadherin, also known as PAS6/7, together with BTN is one of the most abundant glycoproteins of the human MFGM (Quaranta, Giuffrida et al. 2001). All mammalian MFGM LDH contain two epidermal growth factor (EGF)-like domains in the amino terminus. The second EGF-domain contains an Arg-Gly-Asp (RGD) sequence motif. In contrast to all other mammals, human MFGM LDH contains only the second EGF domain (Kvistgaard, Pallesen et al. 2004). LDH also contains a sequence in the C terminus that is homologous to the C1 and C2 domains of the blood clotting factors V and VIII (Andersen, Berglund et al. 1997, Silvestre, Thery et al. 2005). Human MFGM LDH is found at a concentration of 93±10 µg/mL (Peterson, Hamosh et al. 1998) with a molecular mass of 46 kDa (Liu and Newburg 2013). Bovine milk LDH consists of two variants known as PAS6 and PAS7, which have molecular masses of 50 and 47 kDa, respectively. PAS6 and PAS7 share a similar polypeptide core but differ in their glycosylation (Hvarregaard, Andersen et al. 1996).

LDH shows a high level of structural variation between different species. Goat and ovine LDH consist of a single polypeptide chain of approximately 54 and 56 kDa, respectively, whereas bovine LDH consists of two polypeptide chains. Camel milk LDH is also found as two variants of 49 and 55 kDa (Cebo, Caillat et al. 2010, Cebo and Martin 2012) while equine milk contains four variants of LDH, ranging from 45 to 60 kDa (Cebo, Rebours et al. 2012). Multiple isoforms of LDH have been identified in human MFGM and three truncated forms of the protein were also identified (Cavaletto, Giuffrida et al. 1999). The levels of LDH in human milk decrease over the course of lactation and human milk contains significantly higher levels of LDH in early compared to late lactation (Peterson, Hamosh et al. 1998). In bovine milk, it was found that LDH levels remain relatively unchanged in samples taken one week post-parturition compared to colostrum (Reinhardt and Lippolis 2008). Dickow et al. (2011) found that LDH association with the MFGM is affected by temperature, in particular, cooling Holstein-Friesian milk to 4 °C increased the association.
Chapter I

Human milk LDH contains five N-linked glycosylation sites in total (Picariello, Ferranti et al. 2008). In bovine milk, the PAS6 variant of the protein contains two N-linked glycosylation sites at Asn41 and Asn209 and one O-linked glycosylation site at Ser9. PAS7 contains one N-linked glycosylation site at Asn41 and one O-linked glycosylation site at Thr16 (Hvarregaard, Andersen et al. 1996). PAS6 and 7 contain approximately 10.6% and 16.5% (w/w) carbohydrate, respectively (Seok, Shimoda et al. 2001). The oligosaccharides of the bovine glycoprotein variants have both Fuc and sialic acid in the form of N-acetylneuraminic acid (Neu5Ac). Man was present in greater concentrations in PAS6 when compared to PAS7 which correlates with a greater number of N-linked oligosaccharides present (i.e. two N-linked sites for PAS6 compared to one for PAS7) (Hvarregaard, Andersen et al. 1996, Seok, Shimoda et al. 2001).

LDH has displayed a number of biological activities. For example, human milk LDH plays a role in cell adhesion in an RGD-dependent manner (Taylor, Couto et al. 1997). It has also been shown that it mediates vascular endothelial growth factor (VEGF)-dependent angiogenesis both in vitro and in vivo (Silvestre, Thery et al. 2005). In addition, recombinant LDH suppresses both the apoptosis and inflammation associated with ischemic stroke in the rat model (Cheyuo, Jacob et al. 2012).

LDH is resistant to degradation in the infant stomach, most likely due to its high degree of glycosylation (Peterson, Hamosh et al. 1998). Human MFGM LDH has been shown to protect breast fed infants from rotavirus infection, acting as a homologue for the rotavirus receptor and binding to the virus thus preventing it from binding host cells (Newburg, Peterson et al. 1998). It also provides protection from other infectious diseases such as bacterial-associated diarrhoea (Quaranta, Giuffrida et al. 2001, El-Loly 2011). The sialic acid content of human milk LDH plays an important role in its anti-rotavirus activity as removal of sialic acid by chemical hydrolysis reduces the binding of LDH to the virus (Yolkken, Peterson et al. 1992). Interestingly, bovine LDH does not possess these anti-rotavirus properties, which is most likely due to differences in glycosylation (Kvistgaard, Pallesen et al. 2004).
CD59

CD59, also known as Protectin, is an 18-20 kDa GPI-linked glycoprotein found on cell membranes (Watson, Durrant et al. 2006) and has been shown to be present in human MFGM due to its GPI-linked anchor (Hakulinen and Meri 1995). It is also found in the human central nervous system (Akatsu, Yamada et al. 1997) and leukocytes and epithelial cells (Nose, Katoh et al. 1990). To date, only limited research has focused on the CD59 content in other mammalian milks. In a study by Hakulinen and Meri (1995), CD59 was purified from human milk with an average concentration of 1.75 µg/mL (Hakulinen and Meri 1995). Others have demonstrated that the levels of the protein remain stable throughout lactation in human (Bjorge, Jensen et al. 1996) and bovine (Reinhardt and Lippolis 2008) milk.

While limited research has been conducted on the glycosylation of milk CD59, structural information does exist for CD59 isolated from other sources. For example, human erythrocyte CD59 contains eight potential O-linked and one N-linked glycosylation site (Rudd, Morgan et al. 1997) at Asn18 (Ninomiya, Stewart et al. 1992). Wheeler et al. (2002) demonstrated that the glycosylation of GPI-anchored and soluble CD59 are differentially glycosylated, likely due to differing lengths of exposure to glycosyltransferases in the Golgi.

While little information is available on the role of milk-derived CD59, the roles of CD59 isolated from other areas have been defined. CD59 is involved in the regulation of T cell responses (Longhi, Harris et al. 2006) and increases the killing activity of natural killer cells, and a murine variant of CD59 has demonstrated a possible role in B cell proliferation in the humoral response (Kimberley, Sivasankar et al. 2007). It is also involved in lymphocyte activation (Bohana-Kashtan, Ziporen et al. 2004) and functions in cell death. It has been demonstrated that cross-linking of CD59 expressed on T cells, with antibodies, induces apoptosis (Monleon, Martiinez-Lorenzo et al. 2000). It is likely that milk-derived CD59 could also possess similar functions in aiding the immune response. The most well studied function of CD59 is its role in
inhibition of the formation of the membrane attack complex in the complement cascade, where it prevents the C9 molecules from incorporating into the C5b-8 complex (Morgan 1999, Morgan, Chamberlain-Banoub et al. 2006).

Menu et al. (1994) demonstrated that the N-linked glycosylation of human recombinant CD59 is required for efficient CD58-dependent T-cell responses. Conflicting reports exist on the role the oligosaccharides play in complement inhibition by CD59. Ninomiya et al. (1992) showed PNGase F treatment of CD59 resulted in a reduction in complement-inhibitory activity which demonstrated the importance of N-linked glycosylation in this activity. In contrast, other studies have reported that N-linked oligosaccharides do not appear to play a role in complement cascade inhibition (Rother, Zhao et al. 1996, Rushmere, Tomlinson et al. 1997). Further studies are required to determine the functions of milk CD59 and the activities ascribed to its glycosylation.

CD36

CD36 is a 78 kDa glycoprotein which has been found in bovine, human (Greenwalt, Lipsky et al. 1992, Spitsberg 2005), camel (Saadaoui, Henry et al. 2013), caprine, ovine and equine (Cebo and Martin 2012) MFGM. It forms 2-5% of the total protein content of bovine MFGM (Rasmussen, Berglund et al. 1998). CD36 has four free sulfhydryl groups at the cytoplasmic tail (Spitsberg, Matitashvili et al. 1995) and contains at least one intra-chain disulphide bridge which is necessary for its correct transport and processing (Rasmussen, Berglund et al. 1998). CD36 gene expression increases during bovine lactation (Bionaz and Loor 2008) but protein expression remains unchanged in day seven bovine milk compared to colostrum (Reinhardt and Lippolis 2008). However, little is known about how lactation influences the attached oligosaccharides.

Bovine MFGM CD36 contains approximately 24% (w/w) carbohydrate content and eight N-linked glycosylation sites have been characterised for the glycoprotein, namely Asn78, Asn101, Asn171, Asn204, Asn234, Asn246,
Chapter I

Asn320 and Asn416 (Berglund, Petersen et al. 1996). The oligosaccharides are high-mannose, hybrid and complex type, including bi-, tri-, and tetra-antennary structures (Nakata, Furukawa et al. 1993). The oligosaccharides are fucosylated and include, GalNAc, i.e. the LacdiNAc moiety is present, and more than one oligosaccharide type can be found at each site (Berglund, Petersen et al. 1996). The high degree of glycosylation most likely confers resistance to the membrane bound protein from degradation (Greenwalt, Lipsky et al. 1992). However, to date few biological activities have been attributed to the glycosylation of this protein.

Proteose peptone component 3

Proteose peptone component 3 (PP3), also known as lactophorin (LP), is a 28 kDa phosphoglycoprotein (Zhu and Damodaran 2011). It is a substrate of the milk enzyme plasmin that is proteolytically degraded to form a second glycopeptide (18 kDa) called lactophorin 18 (LP18). PP3 has been detected in caprine (Lister, Rasmussen et al. 1998), camel (El-Hatmi, Girardet et al. 2007), bovine and ovine milk but has not been identified in human milk (Sørensen, Rasmussen et al. 1997). It is present at a high concentration in bovine milk at 300 mg/L (Sørensen and Petersen 1993). El-Hatmi et al. (2007) reported that camel milk PP3 concentration varies from 2.7-6.8 g/L during the first eight days of lactation.

Girardet et al. (1995) demonstrated that bovine MFGM PP3 contains two \(O\)-linked glycosylation sites at Thr16 and Thr86 and one \(N\)-linked glycosylation site at Asn77 while LP18 only contains one of the \(O\)-linked (Thr86) and the \(N\)-linked glycosylation site (Girardet, Coddeville et al. 1995, Inagaki, Nakaya et al. 2010). Three neutral oligosaccharides have been identified for the \(O\)-linked glycosylation site of PP3, namely GalNAc (Tn antigen), T antigen and Gal-\(\beta\)-(1→4)-GlcNAc\(\beta\)-(1→6)-[Gal-\(\beta\)-(1→3)]-GalNAc (extended mucin core type-2) (Coddeville, Girardet et al. 1998). The \(N\)-linked structures have been identified as bi-, tri- and tetra-antennary, mono-sialylated
Chapter I

oligosaccharides containing \( N \)-acetyllactosamine (LacNAc) and LacdiNAc, which also may or may not be fucosylated (Inagaki, Nakaya et al. 2010). Furthermore, caprine MFGM PP3 was found to contain one \( N \)-linked and one \( O \)-linked glycosylation site (Lister, Rasmussen et al. 1998). There is debate as to whether PP3 is a membrane bound protein or whether it is loosely associated with the MFGM. It has been found in the bovine MFGM fraction but this soluble glycoprotein has also been sourced from bovine whey (Sørensen, Rasmussen et al. 1997). Interestingly, Fong et al. (2007) hypothesised that the presence of PP3 in the bovine MFGM could be due to the binding of the protein to the membrane during processing. The incorporation of free milk proteins into the MFGM during milk processing has been demonstrated previously however, more studies are required to determine if the presence of PP3 in the MFGM is due to this phenomenon.

PP3 has demonstrated anti-lipase activity in milk (Girardet, Linden et al. 1993) which likely functions to protect the MFG from spontaneous lipolysis (Sørensen, Rasmussen et al. 1997). Stable emulsifying properties have also been identified for the glycoprotein in bovine milk (Innocente, Corradini et al. 1998), which could be an important factor in milk processing. A component of LP28 called lactophorin 16 (LP16), thought to be formed due to proteolytic degradation of LP28, was shown to be a potent inhibitor of the replication of human rotavirus and was shown to play a role in prevention of human rotavirus infection in mice. Although the importance of the glycosylation in this activity has yet to be established, the authors predicted that glycosylation may play an important role as oligosaccharides confer heat stability and LP16 is remarkably heat-resistant (Inagaki, Nagai et al. 2010). In addition, a synthetic peptide consisting of a portion of the C terminus of LP28 demonstrated anti-bacterial activities against pathogenic \textit{Staphylococcus aureus, Salmonella St Paul} and \textit{Pseudomonas aeruginosa}. This demonstrates the potential of the C terminal peptide backbone of milk-derived LP28 glycoprotein for protection against pathogenic infection (Campagna, Mathot et al. 2004). Furthermore, PP3 derived from bovine, ovine and caprine milk has been linked to mitosis - a PP3-
containing fraction of the mammalian milks increased mitosis of MARK3 hybrid cells (murine myeloma-B cell fusion). Interestingly, desialylation of the PP3 fraction lead to decreased activity, indicating the glycosylation plays a role in mitotic stimulation. It was hypothesised that the negative charge attributed by the carboxyl group of the sialic acid residues could enhance stability and solubility of the proteins (Mati, Moulti-Mati et al. 1993). Shida et al. (1994) also demonstrated the ability of glycoproteins sourced from the proteose peptone fraction from bovine milk to bind *E. coli* heat-labile enterotoxin, demonstrating a possible protective function against bacterial-associated disease. Removal of the carbohydrates from these glycoproteins resulted in an inability to bind the toxin, highlighting the important role the oligosaccharides play in this activity.

*Xanthine oxidoreductase*

XOR has been identified in mammalian milks including camel (Baghiani, Harrison et al. 2003), human (Godber, Schwarz et al. 2005), goat (Cebo, Caillat et al. 2010) bovine, sheep and horse (Cebo and Martin 2012). Bovine XOR is found as a peripheral protein on the inner monolayer of the MFGM (Heid and Keenan 2005, Robenek, Hofnagel et al. 2006). It exists as a 300 kDa homodimer with each subunit consisting of a molybdenum cofactor, two Fe2S2 redox centres and one FAD redox centre. In contrast, human milk XOR lacks molybdenum cofactors and is deficient in Fe2S2 centres (Godber, Schwarz et al. 2005). XOR exists in the cell as two inter-convertible forms – xanthine oxidase (XO), which reduces molecular oxygen with the generation of superoxide leading to purine degradation, and xanthine dehydrogenase (XDH), which reduces NAD+ without generation of the reactive oxygen species superoxide (Chen, Park et al. 2012). XDH is the major form which is found in the cell, and can be converted to XO by proteolysis or alternatively, by oxidation of sulfhydryl residues (Enroth, Eger et al. 2000). Whether or not XOR is a glycosylated protein remains unclear. Indeed, glycosylation of this protein could be dependent on its source. Human XO has been found to be *N*-
glycosylated (Picariello, Ferranti et al. 2008) and goat MFGM XO has recently been shown to contain sialic acid in an O-linked oligosaccharide. O-linked glycosylation sites have been identified at Thr207, Thr1069, and Thr1071 in the caprine XO sequence (Cebo, Caillat et al. 2010). However, to date the glycosylation of bovine MFGM XOR has not been confirmed and it has even been suggested that bovine XOR is not glycosylated (Mather 2000). XOR has been isolated from sheep’s milk at a concentration of 22.6±3.3 mg/L (Benboubetra, Baghiani et al. 2004) and bovine MFGM XOR has been shown to be present at a concentration of 16-33 mg/g (Chatterton et al. 2013). Human milk XOR levels in term milk decreased significantly over the course of lactation. However, pre-term milk only had a slight decrease in XOR levels over time (Molinari, Casadio et al. 2013). Interestingly, in both cases, the XOR levels are higher in colostrum compared to mature milk, indicating the important function it serves in the neonatal gut. In contrast, bovine milk levels of XDH are 2.6-fold higher in MFGM samples taken seven days post-parturition compared to colostrum samples (Reinhardt and Lippolis 2008).

XOR has demonstrated anti-microbial properties. Human milk XOR inhibits E. coli and Salmonella enteritidis growth in vitro through nitric oxide production (Stevens, Millar et al. 2000). This activity has been further validated by Hancock et al. (2002) who demonstrated that nitric oxide produced by bovine and human MFGM XOR can protect against E. coli infection in vitro. In addition, hydrogen peroxide produced by bovine XOR MFGM can lead to protection from S. aureus infection (Martin, Hancock et al. 2004). Unfortunately, none of these studies have linked the anti-infective activity of XOR to its glycosylation, the role of which has yet to be fully understood.

**Carbonic anhydrase**

The 40 kDa glycoprotein carbonic anhydrase (CA) has been identified in the human MFGM (Quaranta, Giuffrida et al. 2001) as well as in bovine milk (Ichihara, Asari et al. 2003). CA VI is the CA variant which is found in MFGM
and is a member of a family of zinc-containing enzymes (Karhumaa, Leinonen et al. 2001). It is the only member that is secreted, and has been identified in the saliva or salivary glands and the milk of mammals including that of sheep, human, cow and dog (Fernley, Darling et al. 1989, Parkkila, Kaunisto et al. 1990, Karhumaa, Leinonen et al. 2001, Ogawa, Matsumoto et al. 2002). Both human and bovine colostral milk contain higher levels of CA VI compared to mature milk. Interestingly, the saliva of newborn infants contains lower concentrations of CA VI compared to adult saliva, most likely due to the slower saliva secretion rate of newborns compared to adults. As the human gastrointestinal tract (GIT) is supplied with CA VI through saliva, the high levels of CA VI in human and bovine colostrum could be an example of an evolutionary measure to ensure supply of adequate levels of CA VI to newborns (Karhumaa, Leinonen et al. 2001, Kitade, Nishita et al. 2003). The exact physiological role of CA VI in the GIT has yet to be established; however, the increased levels of CA VI in colostrum suggest a regulatory or protective role in the neonatal gut. Bovine milk CA VI levels decrease over the length of lactation. Studies by Nishita et al. (2007) demonstrated that the colostral concentration of CA VI is on average 119 ng/mL while levels in mature milk decrease to 7.9 ± 12.1 ng/mL (Nishita, Tanaka et al. 2007).

The glycosylated structures of milk CA VI are not well established. However, CA VI isolated from bovine submaxillary gland and parotid gland (Hooper, Beranek et al. 1995, Jiang, Woitach et al. 1996) have been shown to have different glycosylation depending on their site of post-translational modification. This is due to the presence of different glycosyltransferases in each of the sites of isolation. Bovine CA VI isolated from submaxillary gland and parotid gland contain oligosaccharides modified with sulfated and non-sulfated GalNAc, respectively (Hooper, Beranek et al. 1995). The protein isolated from bovine submaxillary gland contains two possible \(N\)-linked glycosylation sites (Jiang, Woitach et al. 1996). In addition, human and ovine secreted CA VI possesses three potential \(N\)-linked sites, two of which are glycosylated (Aldred, Fu et al. 1991, Thatcher, Doherty et al. 1998). Studies on
Chapter I

the glycosylation of mammalian milk CA VI are required in order to fully understand the structure and function of the protein including its glycosylation pattern.

Clusterin

Clusterin, also known as apolipoprotein J, is a 70 kDa (Kounnas, Loukinova et al. 1995) glycoprotein composed of an α- and β- chain connected by disulfide links. Human serum clusterin α- and β-chains can contain 0-30% and 27-30% carbohydrate (w/w), respectively (Kapron, Hilliard et al. 1997). Thus, there are many isoforms of these chains with differing extents of glycosylation. Charlwood et al. (2002) first established the presence of human MFGM clusterin and since then, it has been shown that the human MFGM contains higher levels of the protein when compared to bovine MFGM (Hettinga, van Valenberg et al. 2011). The protein can be found in many fluid types including breast milk, seminal fluid and urine, as well as many cells such as certain cells of the heart, brain and stomach (Rosenberg and Silkensen 1995). Its levels are regulated by the endocytic receptor glycoprotein 330 (gp330), which causes endocytosis and subsequent degradation of clusterin (Kounnas, Loukinova et al. 1995). Human serum clusterin contains three N-linked glycosylation sites in each subunit, while it does not possess any O-linked glycans (Kapron, Hilliard et al. 1997). It has been demonstrated that human MFGM clusterin contains ten possible N-linked glycans, with an abundance of biantennary fucosylated structures. The glycans contain Gal, Man, sialic acid and GlcNAc (Charlwood, Hanrahan et al. 2002). No evidence to date exists on bovine milk clusterin glycosylation. It is known that levels of bovine milk clusterin are not static over the course of lactation. MFGM clusterin levels decline in day seven MFGM samples when compared to colostrum MFGM (Reinhardt and Lippolis 2008).

The role of milk clusterin has yet to be established; however, the functions of clusterin from other sources have been investigated. For example,
human seminal clusterin plays a role in male fertility (Obryan, Baker et al. 1990), human serum clusterin plays a role in inhibition of the complement system (Hochgrebe, Humphreys et al. 1999), and human clusterin also acts as an extracellular chaperone protein, targeting proteins for degradation (Hammad, Ranganathan et al. 1997). Human serum clusterin can bind S. aureus \textit{in vitro} (Partridge, Baker et al. 1996). However, the clinical relevance of this interaction has not been established. It is possible that MFGM-derived clusterin could also display similar bacterial binding and thus could act as a decoy receptor \textit{in vivo}.

Further studies on the glycosylation, and the glycosylation-dependent health promoting properties, of milk-derived clusterin are required. Deglycosylation of the human serum clusterin has led to increased binding to ligands compared to the wild type protein, indicating the oligosaccharides could hinder the optimal binding of the protein to various ligands. However, it is likely that the glycosylation of clusterin mediates the binding to carbohydrate receptors, as demonstrated when Gal significantly inhibited binding of wild type clusterin to liver cells (Stewart, Aquilina et al. 2007). It is likely that the glycosylation of clusterin sourced from MFGM may behave in a similar manner to human serum clusterin.
Chapter I

Commercial potential

Over the past number of decades, manufacturers have attempted to improve foods such as infant formula by adding health promoting ingredients including long chain fatty acids, inulin and fructo- and galacto-oligosaccharides. More recently, manufacturers have recognised domestic animal milks as a potential source of such ingredients as they are readily available and relatively low in cost. Bovine milk is seen as one of the most attractive sources as it is produced in high quantities in countries all over the world and consumed by millions of individuals daily. This nutritional food is also an ideal source of biologically active MFGM glycoconjugates which can be easily harvested for use as functional food ingredients (Haug, Hostmark et al. 2007). Indeed, the isolation and purification of bovine MFGM proteins has already successfully been established at laboratory scale for CD36 (Greenwalt 1993), PAS6/7 (Kim, Kanno et al. 1992), ADPH, BTN (Nielsen, Andersen et al. 1999), MUC1 (Sando, Pearson et al. 2009) and XOR (Sanders, Eisenthal et al. 1997). In addition the laboratory scale extraction of human and camel MFGM XOR (Sanders, Eisenthal et al. 1997) and human protectin (Hakulinen and Meri 1995) has been successful. Interestingly, the purification of the whole bovine MFGM fraction has also been established (Vanderghem, Blecker et al. 2008). The use of intact MFGM containing all the membrane bound and associated proteins may be advantageous in that the full complement of glycolipids and glycoproteins would be available to confer their beneficial properties.

MFGM can be purified from whole milk through a four stage process (Singh 2006), as illustrated in Figure 2. This process involves separation of fat globules, washing of cream, MFGM release from globules, and MFGM collection. In more detail, the first step involves separation of the fat globules from whole milk which is achieved by centrifugation or the use of a cream separator. Next, the cream is washed in buffers such as deionised water or sucrose solutions, a requirement for removal of free milk components such as caseins and whey proteins (Dewettinck, Rombaut et al. 2008). It has been demonstrated that the use of large wash volumes, while keeping washing steps
to a minimum, maximises the yield of MFGM components. Additionally, the use of deionised water as the washing buffer negates the need for salt removal from the purified MFGM (Le, Van Camp et al. 2009). MFGM is separated from the triglyceride core through agitation, churning or freeze-thaw (Dewettinck, Rombaut et al. 2008). It enters the aqueous phase and can be collected from buttermilk or butter serum by methods such as microfiltration (Le, Debyser et al. 2013), ultracentrifugation or freeze-drying. In the past, the large scale production of MFGM was limited by the need to further develop isolation technologies. However, recent advancements in such technologies suggest that whole MFGM is likely to become a common functional food ingredient in the future. An example of such advancement is the use of microfiltration to remove micellar caseins, which are of a similar size to MFGM components (Singh 2006). This method involves the treatment of buttermilk with sodium citrate, which disrupts the casein micelles, and subsequent microfiltration through a 0.1 μm membrane produces an MFGM fraction that contains a high ratio of MFGM proteins (70% of the total protein fraction) compared to skim milk proteins (6% caseins and 24% whey) (Corredig, Roesch et al. 2003). Furthermore, Morin et al. (2007) demonstrated that the use of skim milk ultrafiltrate to wash cream prior to buttermilk generation can increase the ratio of MFGM proteins to skim milk proteins, although there was still some loss of MFGM components. Interestingly, microfiltration of buttermilk produced from washed cream displayed a two-fold increase in permeation flow. Therefore, washing followed by filtration proves to be another viable option for MFGM isolation (Morin, Britten et al. 2007). Indeed, a number of patents exist describing whole MFGM isolation (e.g., WO2001035760 A1 - bulk preparation of milk fat globule membranes) and application (e.g., EP 2509429 A1 - use of MFGM components in infant formula by Morinaga). Moreover, methodologies for the isolation of particular components of MFGM have also been patented (e.g., US 6800739 B2 - a method developed by Davisco for isolation of MFGM glycoproteins from bovine milk). These patents have led many dairy processing companies to isolate glycosylated ingredients from MFGM for use as functional food
Chapter I

ingredients. Such companies include Synlait, Arla Foods, and Büllinger SA, who produce Lipidex, Lacprodan MFGM-10/Lacprodan PL20, and INPULSE, respectively. Lipidex contains many biologically active ingredients including phospholipids (5-7%), gangliosides (0.45%) and sialic acid (0.45%). Lacprodan MFGM-10 contains a number of oligosaccharides and glycoconjugates including sialylated structures, lactoferrin and IgG as well as MUC15 (0.69%) (6.86 µg/mg of total protein) (Pallesen, Pedersen et al. 2007), LDH (0.17%) and MUC1 (0.32%) (Kvistgaard, Pallesen et al. 2004). This product has demonstrated health promoting activities both in vitro and in vivo. For instance, Lacprodan MFGM-10 inhibited the infection of rotavirus in human intestinal cells in vitro (Kvistgaard, Pallesen et al. 2004) and increased oral levels of probiotic Lactobacillus gasseri in infants compared to non-supplemented infant formula (Vestman, Timby et al. 2013). Health promoting properties have also been associated with Lacprodan PL20. Indeed, clinical trials demonstrated that fortification of a drink with Lacprodan PL20 led to reduced stress levels in patients when compared to the non-fortified drink. This clinical trial and other health benefits of Lacprodan PL20 are outlined in a review by Burling and Graverholt (2008). Interestingly, Lacprodan MFGM-10 and Lacprodan PL20 have also been identified as ideal sources of milk polar lipids as outlined in the patent application (US 20130071446 A1). Here, these products have been identified as sources of milk polar lipids for inclusion in infant formulae to increase brain development and improve cognitive function in infants. INPULSE is an MFGM-enriched product with scientifically proven health-promoting activity. In fact, when incorporated into milk and fed to healthy children for four months, the numbers of short febrile episodes were significantly decreased compared to a control group (Veereman-Wauters, Staelens et al. 2012). Although this was attributed to the phospholipid fraction of MFGM, this study highlights the potential use of MFGM in everyday products as a health promoting supplement.

Some issues in MFGM isolation still require resolution. For example, there is a need to establish milk processing conditions that do not result in a loss
of activity and/or concentration of MFGM bioactive compounds. Conditions which adversely affect MFGM isolation include the incorporation of air bubbles into milk by milking machinery, which destabilises the MFGs and causes changes to the MFGM. Cooling, heating and pressurising of milk can lead to a loss of both glycoproteins and glycosphingolipids (Evers 2004). Pasteurisation of cream can lead to an increase in whey protein binding to MFGM and spray drying can decrease the phospholipid content of MFGM (Morin, Jimenez-Flores et al. 2007). Although washing of cream can be used to produce an MFGM fraction free of contaminants, such as free milk proteins, caseins and lactose, it can also increase the loss of loosely bound proteins from MFGM when compared to the integral proteins. Therefore, to maximise MFGM recovery without compromising composition, further investigation is required to determine the exact methods of pre- and post-churning processing which minimise loss of MFGM bioactivity. Additionally, carbohydrate analysis after processing is necessary in order to establish the method which best preserves glycosylation structure and function to capitalise fully on the beneficial health promoting activities they possess.
Chapter I

Future prospectives

The bioactivities demonstrated by MFGM glycoconjugates highlight their potential use as functional ingredients that could be incorporated into an array of products. Glycosylated MFGM components show enormous potential in acting as decoy receptors which bind pathogens and prevent adhesion to host cells. The pathogens are then flushed out of the body through the natural defences, leaving the host unharmed (Sharon and Ofek 2000). MFGM glycoconjugates could also be used in the prevention of impending bacterial infection. Their use as supplements in food and beverages could increase the resistance of consumers against the possible threat of infection. Furthermore, MFGM glycosylation could be exploited by immuno-compromised patients to fight infection and some in vivo studies suggest they may ameliorate the symptoms of previous infection (Mouricout, Petit et al. 1990). There are a number of products in which MFGM glycoconjugates could be incorporated, including infant formula, which currently lacks some protective ingredients that are present in maternal breast milk. Other food matrices which could be supplemented with these bioactives include water-based beverages and cereals (Sørensen, Rasmussen et al. 1997). Indeed, laboratory-based research has already begun to investigate the possibility of including MFGM in foods such as yoghurt (Le, Van Camp et al. 2009).

A novel approach to produce new bioactives from MFGM may be to ferment MFGM-containing milk fractions. Lactobacillus helveticus strains are widely used for production of dairy products such as yoghurt and cheese and grow well on dairy sources (Matar, Amiot et al. 1996), and could be used to produce fermented MFGM products. Their glycolytic activity could successfully hydrolyse MFGM glycans to produce free oligosaccharides (Adamberg, Antonsson et al. 2005) and they possess proteolytic and peptidase activities (Matar, Valdez et al. 2001) which could result in the production of bioactive glycopeptides. Indeed, the production of bioactives with health promoting benefits, such as anti-cancer (Rachid, Matar et al. 2006),
immunomodulatory (Meena, Kapila et al. 2008) and anti-hypertension (Aihara, Kajimoto et al. 2005) activities by fermentation with *L. helveticus*, is well documented.

Although research to date has mainly focused on MFGM glycoconjugates and their health promoting activities in the prevention of infection and disease, MFGM could also be used for other health promoting properties. For example, it shows potential to act as a delivery system for drugs and to enhance their intestinal absorption (Sato, Liu et al. 1994). MFGM may also be used as an alternative source of phospholipids for liposome preparation to prevent the gastric digestion of protein-based prescriptions. A recent study has shown that it can protect lactoferrin from hydrolysis *in vitro* (Liu, Ye et al. 2013). Ultimately, there are many possible uses of MFGM for improved human health applications.
Chapter I

Conclusions

It is our belief that MFGM has been undervalued in the food industry; however, this perspective has changed in recent years as the knowledge of MFGM composition and bioactivities has increased significantly. The potential it holds commercially as a functional ingredient is being uncovered due to the increased number of studies that have characterised its physical and chemical properties, particularly those of the glycoconjugates. In order to capitalise on all the health-promoting bioactivities MFGM offers, further studies are required to determine the most suitable methods of MFGM isolation and purification that would most preserve its integrity and bioactivities. Further work is required to fully characterize the physical properties of MFGM glycoconjugates and their health-promoting roles in order to truly understand the potential this milk component holds.
Chapter I

References


Chapter I


Chapter I


Chapter I


Chapter I


Chapter I


Khatun, U. L., A. Gayen and C. Mukhopadhyay (2013). "Gangliosides containing different numbers of sialic acids affect the morphology and
Chapter I

structural organization of isotropic phospholipid bicelles." Chemistry and Physics of Lipids 170: 8-18.


Chapter I


suppression of experimental autoimmune encephalomyelitis with the milk protein butyrophilin." International Immunology 16(3): 489-499.


Chapter I


Chapter I


Chapter I


Chapter I


Chapter I


Chapter I


Chapter I


Chapter I

binding of gangliosides to *Helicobacter pylori* vacuolating cytotoxin (VacA) neutralizes its toxin activity." Glycobiology 20(6): 668-678.


Chapter I

milk and at the luminal surfaces of blood vessels." Journal of Cellular Physiology 204(1): 166-177.

Figure 1. Structure of the MFGM. The sizes of the membrane components are not proportional. The MFGM consists of an inner phospholipid monolayer and outer bilayer derived from the endoplasmic reticulum and apical membrane, respectively. A proteinaceous coat exists between both layers. Glycoproteins present in the MFGM include XOR (xanthine oxidoreductase) and CLUS (clusterin), BTN (butyrophilin), CD36 (cluster of differentiation 36), MUC1 (mucin 1), CD59, LDH (lactadherin) and PP3. Adipophilin (ADPH) is not a glycoprotein however is involved in MFGM formation with XOR and BTN. Neutral and acidic glycolipids are also present in the MFGM.
Step 1: Separation

Step 2: Washing of cream
2-3 wash steps

Step 3: MFGM release
-Churn
-Freeze/thaw

Step 4: MFGM collection
-Ultracentrifugation
-Microfiltration
-Freeze Drying

Figure 2. Outline of MFGM processing. Isolation of MFGM consists of 4 steps. 1) Separation of raw milk into cream and skim milk. 2) Washing of cream to remove free milk proteins. 3) Release of MFGM through methods including churning or freeze-thaw. 4) Collection of MFGM through the use of ultracentrifugation, microfiltration or freeze drying. *GC – glycoconjugate fraction
Table 1. Overview of the mammalian MFGM glycolipids and their properties

<table>
<thead>
<tr>
<th>Glycolipid</th>
<th>Group</th>
<th>Glycosylation</th>
<th>Bioactivity</th>
<th>Function of glycan</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lactosylceramide (LacCer)</td>
<td>Neutral glycolipid</td>
<td>Gal-β-(1,4)-Glc-β-Cer</td>
<td><em>E. coli</em></td>
<td>Unknown</td>
<td>Schnaar et al. 2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sanchez-Juanes et al. 2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zancada et al. 2010</td>
</tr>
<tr>
<td>Glucosylceramide</td>
<td>Neutral glycolipid</td>
<td>Glc residue connected to ceramide</td>
<td><em>E. coli</em> binding</td>
<td>Unknown</td>
<td>Schnaar et al. 2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zancada et al. 2010</td>
</tr>
<tr>
<td>Galactosylceramide</td>
<td>Neutral glycolipid</td>
<td>Gal residue connected to ceramide</td>
<td><em>E. coli</em> binding</td>
<td>Unknown</td>
<td>Schnaar et al. 2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zancada et al. 2010</td>
</tr>
<tr>
<td>Monosialoganglioside 1 (GM1)</td>
<td>Ganglioside</td>
<td>1 sialic acid residue Gal-β-(1,3)-GalNAc-β-(1,4)-Neu5Ac-α-(2,3)-Gal-β-(1,4)-Glc-β-Cer</td>
<td>Anti-adhesive <em>in vitro</em> (bacterial) Bacterial toxin binding</td>
<td>Unknown</td>
<td>Jensen 2002 Idota and Kawakami 1995</td>
</tr>
<tr>
<td>Monosialoganglioside 2 (GM2)</td>
<td>Ganglioside</td>
<td>1 sialic acid residue GalNAc-β-(1,4)-Neu5Ac-α-(2,3)-Gal-β-(1,4)-Glc-β-Cer</td>
<td>Anti-adhesive <em>in vitro</em> (viral)</td>
<td>Unknown</td>
<td>Jensen 2002 Portelli et al. 1998 Smilowitz et al. 2014</td>
</tr>
</tbody>
</table>
| Ganglioside | Monosialoganglioside 3 (GM3) | Cer | 1 sialic acid residue | Neu5Ac-α-(2,3)-Gal-β-(1,4)-Glc-β-Cer | Anti-adhesive in vitro (bacterial and viral) | Unknown | Jensen 2002  
Iota and Kawakami 1995  
Salcedo et al. 2013  
Iskarpayoti et al. 2012  
Takamizawa et al. 1986 |
|---|---|---|---|---|---|---|---|
| Ganglioside | Disialoganglioside 1A (GD1A) | Cer | 2 sialic acid residues | Neu5Ac-α-(2,3)-Gal-β-(1,3)-GalNAc-β-(1,4)-Neu5Ac-α-(2,3)-Gal-β-(1,4)-Glc-β-Cer | Unknown | Unknown | Jensen 2002  
Ito et al. 2012 |
| Ganglioside | Disialoganglioside 1B (GD1B) | Cer | 2 sialic acid residues | Gal-β-(1,3)-GalNAc-β-(1,4)-Neu5Ac-α-(2,8)-Neu5Ac-α-(2,3)-Gal-β-(1,4)-Glc-β-Cer | Unknown | Unknown | Jensen 2002  
Jensen and Newburg 1995 |
| Ganglioside | Disialoganglioside 2 (GD2) | Cer | 2 sialic acid residues | GalNAc-β-(1,4)-Neu5Ac-α-(2,8) | Unknown | Unknown | Jensen 2002  
Jensen and Newburg 1995 |
| Ganglioside                         | Ganglioside                          | Possible protection of neonatal gut E. coli binding Anti-adhesive in vitro (bacterial and viral) | Unknown                      | Jensen 2002  
Sanchez-Juanes et al. 2009  
Idota and Kawakami 1995  
Salcedo et al. 2013  
Iskarpatyoti et al. 2012  
Takamizawa et al. 1986 |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Disialoganglioside 3 (GD3)</td>
<td>Neu5Ac-α-(2,3)-Gal-β-(1,4)-Glc-β-Cer</td>
<td>2 sialic acid residues Neu5Ac-α-(2,8)-Neu5Ac-α-(2,3)-Gal-β-(1,4)-Glc-β-Cer</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
</tbody>
</table>
| Trisialoganglioside (GT)           | Neu5Ac-α-(2,8)-Neu5Ac-α-(2,3)-Gal-β-(1,4)-Glc-β-Cer | 3 sialic acid residues Neu5Ac-α-(2,8)-Neu5Ac-α-(2,3)-Gal-β-(1,4)-Glc-β-Cer                    | Unknown                     | Jensen 2002  
Takamizawa et al. 1986 |

Chapter I
Chapter I
Table 2. Overview of the mammalian MFGM glycoproteins and their properties

<table>
<thead>
<tr>
<th>Glycoprotein</th>
<th>Glycosylation</th>
<th>Bioactivity</th>
<th>Importance of glycans</th>
<th>References</th>
</tr>
</thead>
</table>
Chapter I

| CD59 | Unknown | Unknown | Unknown |
| CD36, GP88, GPIIb, PASIV, FAT, glycoprotein IV | Bovine: 8 N-linked glycosylation sites Bovine residues: Man, Gal, Fuc, GalNAc and GlcNAc | Unknown | Resistance to digestion | Berglund et al. 1996 Greenwalt et al. 1992 |
| Xanthine oxidoreductase | Caprine: O-linked glycosylation at Thr207, Thr1069, and Thr1071 O-linked glycosylation containing sialic acid Human: N-linked glycosylation | Anti-microbial through nitric oxide and hydrogen peroxide production Unknown | Cebo et al. 2010 Picariello et al. 2008 Stevens et al. 2000 Hancock et al. 2002 Martin et al. 2004 |
| Carbonic anhydrase VI | Unknown | Possible protective role in neonatal gut Unknown | Karhumaa et al. 2001 Kitade et al. 2003 |
| Clusterin, Apolipoprotein J | Human residues: Sialic acid, Man, Gal, GlcNAc. High abundance of biantennary fucosylated glycans | Unknown Unknown | Charlwood et al. 2002 |
Chapter II

Defatted bovine milk fat globule membrane inhibits association of enterohaemorrhagic *Escherichia coli* O157:H7 with human HT-29 cells

Sarah A. Ross, Jonathan A. Lane, Michelle Kilcoyne, Lokesh Joshi and Rita M. Hickey

Published in International Dairy Journal (2016), 59, 36-43.

Acknowledgements: Sarah Ross performed all experimental work for this chapter. Jonathan Lane aided with experimental assays, and Michelle Kilcoyne and Jonathan Lane assisted with manuscript preparation.
Chapter II

Abstract

The bovine milk fat globule membrane (MFGM) is a source of food-derived glycans which can offer an approach to prevent *Escherichia coli* O157:H7 infection by inhibiting attachment of the pathogen to host cells. Such glycans may decrease the need for antibiotic treatment by acting as prophylactics. In this study, we generated a defatted bovine MFGM fraction, rich in proteins and glycoproteins, and demonstrated its ability to prevent the association of several enterohaemorrhagic *E. coli* O157:H7 strains with human colonic adenocarcinoma, HT-29 cells. This defatted MFGM fraction reduced bacterial association with HT-29 cells in a concentration dependent and strain specific manner. This study may present a new approach to mitigate the adverse health effects caused by *E. coli* infections in humans.
Chapter II

Introduction

Enterohemorrhagic *Escherichia coli* (EHEC) serogroup O157:H7 is an enteric pathogen commonly associated with food borne illnesses. It has a very low infectious dose, which varies from 10 to a few hundred cells (Strachan, Doyle et al. 2005), and causes an estimated 73,480 illnesses and 2168 hospitalisations annually (Mead, Slutsker et al. 1999). The symptoms of infection include abdominal cramping, nausea, vomiting, abdominal distension and bloody diarrhoea (Symonds 1988). Severe cases of infection can result in haemolytic uremic syndrome (HUS) (Wong, Mooney et al. 2012), haemorrhagic colitis (Sang, Saidi et al. 1996) and even death. Patients who develop HUS have a mortality rate of 3-5% (Feng, 2012). The incidences of death due to *E. coli* O157:H7 infection have reduced in recent years, reaching an average of 10 deaths annually and generally occurring in people aged ≥ 60 years (Gould, Demma et al. 2009, Scallan, Hoekstra et al. 2015). However, large sporadic outbreaks due to contaminated food sources such as beef (Duffy, Butler et al. 2006), continue to emerge (CDC, 2008, 2013) which increases the risk of development of HUS. The costs associated with outbreaks in the USA were estimated to be US$405 million in 2003 (Frenzen, Drake et al. 2005) highlighting the need for *E. coli* O157:H7 infection prevention and treatment.

In developing countries, enteropathogenic *E. coli* (EPEC) is a leading cause of diarrhoea in infants (Trabulsi, Keller et al. 2002). The symptoms of EPEC infections most commonly include acute diarrhoea and may lead to persistent diarrhoea (Ochoa and Contreras 2011). EPEC infection is the cause of thousands of deaths annually, with the majority of these deaths occurring in children and infants (Ochoa, Barletta et al. 2008). EPEC differs in its virulence to EHEC in that infection rates are higher in developing countries and infectious doses are also much higher (10^8-10^10 bacteria to infect adults). Furthermore, EPEC colonise the small intestine as opposed to the colon – the area commonly targeted by EHEC (Mellies, Barron et al. 2007).
During infection, *E. coli* spp. overcome the harsh acidic environment of the gut to colonize the gastrointestinal tract (Foster 2004). Adhesion is an important early step in *E. coli* pathogenicity. It enables the bacteria to gain access to essential nutrients, deliver bacterial toxins into the microenvironment of the host tissue, and triggers host cell invasion (Sharon and Ofek 2000, Schlumberger and Hardt 2006). It also increases the chance of bacterial survival as their resistance to cleansing mechanisms, immune factors and antibiotics is higher (Ofek, Hasty et al. 2003). *E. coli* adhesion mechanisms have been extensively reviewed in recent years (Croxen, Law et al. 2013, Pereira and Giugliano 2013) and some of the carbohydrate binding sites of EPEC and EHEC have been determined. These include intestinal glycosphingolipids (Teneberg, Ångström et al. 2004), asialolactosamine (Vanmaele, Finlayson et al. 1995), N-acetylgalactosamine on the surface of HeLa cells (Scaletsky, Milani et al. 1988), fucosylated human milk oligosaccharides (Cravioto, Tello et al. 1991), monosialogangliosides (GM) (Idota and Kawakami 1995) and mucin-type core 2 O-glycans (Ye, Song et al. 2015).

As pathogen adhesion and colonization is often a prelude to infection, intervention at the adhesion stage can reduce or prevent disease. Intervention can occur at birth via bioactive breast milk components such as free and bound glycans. These structurally mimic epithelial cell surface glycans and thus function as decoys which pathogens can bind to instead of the host and thereby prevent infection (Sharon and Ofek 2000). Indeed, human milk oligosaccharides (HMO) have been shown to inhibit *E. coli*, *Vibrio cholerae* and *Salmonella typhimurium* adhesion to human epithelial colorectal adenocarcinoma Caco-2 cells *in vitro* (Coppa, Zampini et al. 2006) and fucosylated HMO are thought to decrease *Campylobacter jejuni* associated diarrhoea in breast fed infants (Morrow, Ruiz-Palacios et al. 2004). Human milk fat globule membrane (MFGM) glycoproteins have also been shown to possess anti-infective properties. For instance, mucins isolated from human MFGM prevented S-fimbriated *E. coli* adhesion to buccal epithelial cells (Schroten, Hanisch et al.
Chapter II

1992) and also decreased *Salmonella enterica* serovar Typhimurium invasion of human intestinal epithelial cells (Liu, Yu et al. 2012).

However, the limited availability of human milk makes it difficult to produce large quantities of these bioactives at a commercial scale for human consumption. As a result, scientists have begun to focus on domestic animal milks, particularly bovine milk, as a source of potential anti-infective agents such as glycans and glycoconjugates that are structurally similar to those found in human breast milk. For example, bovine milk oligosaccharides (BMO) were shown to prevent the cellular invasion of *C. jejuni* in vitro (Lane, Marino et al. 2012) and to display anti-infective activity against *Helicobacter pylori*, *Neisseria meningitides* and influenza virus (Matrosovich, Gambaryan et al. 1993, Hakkarainen, Toivanen et al. 2005). Bovine MFGM may represent an alternative source of anti-infective glycans, which can be harvested from whole milk through conventional technologies. Indeed, MFGM is a rich source of glycoproteins which have complex glycan components (Ross, Lane et al. 2015). Furthermore, recent studies have demonstrated the anti-infective potential of certain purified bovine MFGM glycoproteins and whey fractions against a range of pathogens including rotavirus and enteric bacteria (Kvistgaard, Pallesen et al. 2004, Inagaki, Nagai et al. 2010, Parker, Sando et al. 2010). However, these studies used various forms of purified MFGM glycoproteins rather than a glycoprotein-enriched MFGM fraction.

In this study, a glycoprotein and protein enriched MFGM fraction from bovine milk was prepared and characterised and then assessed for its ability to prevent EHEC and EPEC association with human colonic intestinal epithelial cells. The main objectives of this study were to continue to explore bovine milk as a source of human health promoting bioactives and to potentially identify an alternative preventative therapy for *E. coli* infection.
Chapter II

Materials and methods

Preparation of a defatted MFGM-enriched fraction (dMFGM) from bovine milk

Milk was collected from the bulk tank of the Holstein-Friesian cattle milking parlour at the Teagasc Food Research Centre, Moorepark (Fermoy, Co. Cork, Ireland). Cream and milk fat globules were separated from the milk using a FT15 disc bowl centrifuge (Armfield Ltd., Ringwood, England) and the cream was stored at 4 °C for 12-24 h. After chilling, the cream was churned using a mixer to produce butter and buttermilk. The buttermilk was then passed through glass wool (Sigma-Aldrich, Dublin, Ireland) to remove minute butter granules. Washed cream was produced as described previously (Le, Van Camp et al. 2009, Struijs, Van de Wiele et al. 2013). Briefly, warm deionized water (37 °C) was added to the cream at a ratio of 1:10 and the mixture was added to the FT15 disc bowl centrifuge. This process was repeated twice and the washed cream was then used to produce buttermilk as described above.

Modifications were made to the conventional Rose Gottlieb method (IDF 2008) to delipidate the produced buttermilks. Briefly, 10 mL ethanol, 25 mL diethyl ether and 25 mL of petroleum spirit were added to 10 mL of buttermilk. The mixture was shaken for 30 s and then centrifuged (600 x g, 5 min) at 20 °C. The solvent layer was then removed and the aqueous layer was treated with 5 mL ethanol, 15 mL diethyl ether and 15 mL of petroleum ether prior to centrifugation as described above. The solvent was removed and the aqueous layer was treated as per the second extraction. After the third extraction, the aqueous layer was collected and all solvents were removed by rotary evaporation under reduced pressure (Buchi Rotavapor R-210, Mason Technology Ltd, Dublin, Ireland). The sample was then lyophilised and the dry powder stored in a desiccator at room temperature for long-term storage.
Sodium dodecyl sulfate polyacrylamide gel electrophoresis analysis

Sample preparation and reduction for sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis was performed as per manufacturer’s instructions (NuPAGE system, Life Technologies, Thermo Fisher Scientific Inc., Dublin, Ireland). Briefly, 2.5 µL of sample buffer and 1 µL of reducing agent were added to 6.5 µL of sample 10 mg mL\(^{-1}\) (by mass of powder). The sample was then centrifuged and heated to 70 °C for 10 min. 10 µL of sample was then added to each well of a 4-12% Bis-Tris gel (1.00 mm x 15 well, Life Technologies). A molecular weight standard solution (Invitrogen Mark12 Unstained Standard, Thermo Fisher Scientific Inc.), prepared as per manufacturer’s instructions (diluted 1:10 with Invitrogen LDS Sample Buffer), was also loaded onto the gels. Electrophoresis was performed at 200 V for 50 min using MOPS buffer with MOPS buffer supplemented with 0.25% NuPAGE Antioxidant (Life Technologies) in the upper chamber. Protein bands were visualized on the gels using Coomassie blue stain (Invitrogen SimplyBlue SafeStain) following the manufacturer’s procedure. MFGM proteins and contaminating proteins were identified through the use of bovine milk protein standards β-lactoglobulin A, β-lactoglobulin B, α-lactalbumin, κ-casein, α-casein, β-casein, lactoferrin, and immunoglobulin G (IgG) and by comparative analysis with previously published SDS-PAGE data (Ye, Singh et al. 2002, Rombaut, Dejonckheere et al. 2007).

Bacterial strains and culture conditions

EHEC serogroup O157:H7 strains (NCTC 12900*, DAF454†, DPC 6055†, P1432†, DPC 6054†, ATCC 43888§) and EPEC serogroups O125:H19 (strain NCTC 8623*) and O111:H2 (strain NCTC 8007*) were obtained from the National Collection of Type Cultures* (NCTC; London, UK), the Dairy Products Research Centre culture collection at Teagasc Moorepark (DPC)† and the American Type Culture Collection (ATCC, Rockville, MD)‡. All strains were cultured directly from storage into brain heart infusion (BHI) (Oxoid®
Chapter II

Lactobacillus plantarum (ATCC Ltd, Basingstoke, England) broth and incubated under aerobic conditions at 37 °C. Strains were stocked in BHI broth containing 50% glycerol (v/v) and stored at -20 °C.

Mammalian cell culture

The human colonic adenocarcinoma cell line, HT-29, was purchased from the ATCC. HT-29 cells were routinely grown in McCoy’s 5A modified medium (Sigma-Aldrich®) supplemented with 10% fetal bovine serum (FBS). All cells were routinely maintained in 75 cm² tissue culture flasks and incubated at 37 °C in 5% (v/v) CO₂ in a humidified atmosphere. Cells were passaged when the confluency of the flask was approximately 90% as previously described (Lane, Marino et al. 2012).

In vitro total association assays

A series of total association assays, adapted from (Horemans, Kerstens et al. 2012, Salcedo, Barbera et al. 2013), were performed with HT-29 cells and all EPEC and EHEC strains in the absence (control) and presence of the generated dMFGM fraction. HT-29 cells were seeded into 12 well PVDF plates (Corning, Sigma Aldrich) at a density of 1 x 10⁵ cells well⁻¹. Cells were cultured for 48 h and the media was changed to McCoy’s 5A modified medium supplemented with 2% FBS at least 24 h prior to total association assays. E. coli strains in the early stationary phase were harvested from BHI broth after overnight growth, washed three times in phosphate buffered saline, pH 7.2 (PBS) and re-suspended in McCoy’s 5A modified medium supplemented with 2% FBS. Standard competition assay, concentration dependency assay, bacterial interaction assay, host cell interaction assay and no pre-incubation assay were then performed as described below.
Chapter II

Standard competition assay

Prior to infecting the HT-29 cell line, the bacteria (1 x 10^8 CFU mL⁻¹) were pre-incubated for 1 h at 37 °C (5% CO₂) with the dMFGM fraction (5 mg mL⁻¹) in McCoy’s 5A media (2% FBS). Confluent monolayers of HT-29 cells were washed twice in PBS and infected with 500 µl of the pre-incubated bacterial mix. HT-29 cells were then incubated for 1 h at 37 °C in a humidified atmosphere (5% CO₂). To determine the number of cell-associated bacteria, non-adherent bacteria were removed by washing the HT-29 cells five times with PBS and cells were lysed with 500 µl of 0.1% Triton X-100 solution in PBS. Cell lysates were then serially diluted in maximum recovery diluant (Oxoid®) and plated onto BHI agar (Oxoid®). Plates were incubated at 37 °C overnight and CFU mL⁻¹ was determined.

Concentration dependency assay

Prior to infecting the HT-29 cell line, the bacteria (1 x 10^8 CFU mL⁻¹) were pre-incubated for 1 h at 37 °C (5% CO₂) with various concentrations of the dMFGM fraction (5, 2.5, 1.25 and 0.625 mg mL⁻¹) in McCoy’s 5A media (2% FBS). Confluent monolayers of HT-29 cells were washed twice in PBS and infected with 500 µl of the pre-incubated bacterial mix. HT-29 cells were then incubated for 1 h at 37 °C in a humidified atmosphere (5% CO₂). To determine the number of cell-associated bacteria, non-adherent bacteria were removed by washing the HT-29 cells five times with PBS and cells were lysed with 500 µl of 0.1% Triton X-100 solution in PBS. Cell lysates were then serially diluted in maximum recovery diluant (Oxoid®) and plated onto BHI agar (Oxoid®). Plates were incubated at 37 °C overnight and CFU mL⁻¹ was determined.
Chapter II

Bacterial interaction assay

Prior to infecting the HT-29 cell line, the bacteria \(1 \times 10^8\) CFU mL\(^{-1}\) were pre-incubated for 1 h at 37 °C (5% CO\(_2\)) with the dMFGM fraction (5 mg mL\(^{-1}\)) in McCoy’s 5A media (2% FBS). The bacterial mix was then centrifuged at 3920 x g for 7 min to pellet the bacterial cells. The supernatant containing unbound dMFGM was removed and the bacterial pellet was then re-suspended in McCoy’s 5A media (2% FBS). Confluent monolayers of HT-29 cells were washed twice in PBS and infected with 500 µl of the resuspended bacterial mix. HT-29 cells were then incubated for 1 h at 37 °C in a humidified atmosphere (5% CO\(_2\)). To determine the number of cell-associated bacteria, non-adherent bacteria were removed by washing the HT-29 cells five times with PBS and cells were lysed with 500 µl of 0.1% Triton X-100 solution in PBS. Cell lysates were then serially diluted in maximum recovery dilutant (Oxoid®) and plated onto BHI agar (Oxoid®). Plates were incubated at 37 °C overnight and CFU mL\(^{-1}\) was determined.

Host cell interaction assay

Prior to infecting the HT-29 cell line, the confluent monolayer of HT-29 cells was washed twice in PBS and pre-incubated for 1 h at 37 °C (5% CO\(_2\)) with 500 µl of the dMFGM fraction (5 mg mL\(^{-1}\)) in McCoy’s 5A media (2% FBS). Unbound dMFGM components were removed by washing the mammalian cells five times in PBS prior to bacterial infection. The confluent monolayers of HT-29 cells were infected with 500 µl of the bacterial mix. HT-29 cells were then incubated for 1 h at 37 °C in a humidified atmosphere (5% CO\(_2\)). To determine the number of cell-associated bacteria, non-adherent bacteria were removed by washing the HT-29 cells five times with PBS and cells were lysed with 500 µl of 0.1% Triton X-100 solution in PBS. Cell lysates were then serially diluted in maximum recovery dilutant (Oxoid®) and plated onto BHI agar (Oxoid®). Plates were incubated at 37 °C overnight and CFU mL\(^{-1}\) was determined.
No pre-incubation assay

Confluent monolayers of HT-29 cells were washed twice in PBS and infected with 500 µl of the non-preincubated bacterial mix. HT-29 cells were then incubated for 1 h at 37 °C in a humidified atmosphere (5% CO₂). To determine the number of cell-associated bacteria, non-adherent bacteria were removed by washing the HT-29 cells five times with PBS and cells were lysed with 500 µl of 0.1% Triton X-100 solution in PBS. Cell lysates were then serially diluted in maximum recovery dilutant (Oxoid®) and plated onto BHI agar (Oxoid®). Plates were incubated at 37 °C overnight and CFU mL⁻¹ was determined.

**Bacterial growth assays**

To assess whether dMFGM had an effect on the growth of the *E. coli* strains used in this study, the bacteria were grown in optimum growth media BHI (Oxoid®) broth, and colonisation media, McCoy’s 5A media modified with L-glutamine and NaHCO₃ (Sigma-Aldrich) (2% FBS) and supplemented with 5 mg mL⁻¹ dMFGM. Unsupplemented media was used as a negative control. EPEC and EHEC strains harvested from BHI broth were used to inoculate (1%) the test media. Solutions were then incubated at 37 °C under aerobic conditions and bacterial growth was monitored by making serial dilutions of the inoculated media at the following time points: 0 h (T₀), 3 h (T₃), 6 h (T₆) and 24 h (T₂₄) post-inoculation. Colony forming units were determined by plating sample aliquots in triplicate on BHI agar (Oxoid®). Plates were incubated at 37 °C overnight and CFU mL⁻¹ was determined.

**Statistical analysis**

All inhibition studies were carried out three times in triplicate. Data were graphed in Microsoft Excel and the unpaired student t-test was used to determine statistically significant results; P<0.05 was considered significant.
Chapter II

Results and discussion

Isolation and characterisation of the dMFGM fraction

Approximately 16 well separated protein bands were identified on the Coomassie-stained SDS gels (Figure 1A). High intensity bands were observed for the MFGM glycoproteins XOR, CD36, butyrophilin (BTN), Pas 6 and Pas 7 (approximately 130, 68, 60, 50 and 46 kDa, respectively). High intensity bands were also observed for the milk glycoprotein κ-casein (approximately 26 kDa) and for the proteins α-casein, β-casein (approximately 34 kDa and 31 kDa, respectively) and the whey proteins α-lactalbumin and β-lactoglobulin (approximately 10 kDa and 19 kDa, respectively). Low intensity bands were observed for glycoproteins IgG (heavy chain 55 kDa and light chain 20 kDa approximately) and proteose peptone 3 (PP3) (approximately 28 kDa) and for the protein adipophilin (approximately 40 kDa). These findings were comparable with previous publications (Ye, Singh et al. 2002, Rombaut, Dejonckheere et al. 2007) and suggested that MFGM was successfully isolated from whole milk but was contaminated with non-MFGM proteins such as caseins (Figure 1A). Washing removed the majority of the contaminating proteins including α-lactalbumin, β-lactoglobulin, α-casein, β-casein and κ-casein, while enriching for MFGM proteins and glycoproteins (Figure 1B). This approach was adapted from Le et al. (2009) where the use of deionized water to wash cream for the removal of contaminating non-MFGM proteins without the loss of MFGM proteins was demonstrated. In this study, the anti-infective potential of defatted bovine MFGM-enriched fraction against E. coli was under investigation. Therefore, as the bovine MFGM is composed of membrane proteins and lipids, the MFGM-enriched fraction was delipidated and the subsequent dMFGM fraction was used for the inhibition assays.
dMFGM reduced the association of *E. coli* O157:H7 with HT-29 cells

Previously, Wang et al. (2001) and Horemans et al. (2012) demonstrated that a concentration of ≥5 mg mL\(^{-1}\) of defatted MFGM was required to cause 50%-80% inhibition of *H. pylori* adherence to HeLa S3 cell monolayers and NCI-N87 cells, respectively. Therefore, an initial concentration of 5 mg mL\(^{-1}\) dMFGM was selected for this study. Prior to mammalian cell infection, *E. coli* strains were pre-incubated with dMFGM. Subsequently, dMFGM was shown not to kill the bacteria, did not influence bacterial growth over the course of the assay and did not affect the viability of the HT-29 cells as confirmed by bacterial growth assays and real-time analysis of cell viability using an xCELLigence system (Roche), respectively (data not shown). The standard competition assay demonstrated the dMFGM inhibited the association of the EHEC strains only, namely *E. coli* NCTC 12900, *E. coli* DAF454 and *E. coli* DPC 6055 to HT-29 cells by 92%, 88% and 69%, respectively (*P*-value < 0.05) when compared with untreated controls (Figure 2). No reduction in bacterial association with HT-29 cells was evident for *E. coli* P1432, *E. coli* DPC 6054, *E. coli* ATCC 43888, *E. coli* NCTC 8623 and *E. coli* NCTC 8007. Interestingly, bovine MFGM has been shown to influence the expression of genes involved in the motility of *E. coli* O157:H7 (Tellez, Corredig et al. 2012). For instance, it was demonstrated that an increase in the expression of fliC gene, which encodes the major bacterial flagellar protein, occurs after 4 h treatment of *E. coli* with bovine MFGM. Thus, an increase in fliC expression could have occurred in some strains in this study, increasing their motility and aiding them in subverting the anti-adhesive activity of the dMFGM fraction. This may explain why no anti-infective activity was observed for some of the EHEC strains screened. Previously, bacterial flagella were shown to impact greatly on bacterial ability to adhere and infect host cells (Mahajan, Currie et al. 2009). EHEC O157:H7 possess H7 type flagella, which have been shown to have adhesive properties and have demonstrated binding abilities to mucins. The EPEC strains, *E. coli* NCTC 8623 and *E. coli* NCTC 8007, possess H19 and H2 type flagella, respectively (Erdem, Avelino et al. 2007). This difference in
adhesive properties might explain the difference in cellular association observed for the EPEC strains when compared to the EHEC strains. It is also possible that since the EPEC strains bind to different ligands they may not interact with dMFGM in a similar manner to that of *E. coli* O157:H7. Moreover, EPEC and EHEC differ in several other ways. The main virulence property of EPEC is the production of attaching and effacing (A/E) lesions while the main EHEC virulence properties include production of A/E lesions and verotoxin (Bardiau, Szalo et al. 2010). Both serotypes contain the bacterial receptor Tir which is inserted into the host cell membrane prior to binding bacterial intimin protein which is necessary for pedestal formation. EPEC relies on tyrosine phosphorylation of its Tir protein for pedestal formation while EHEC does not (DeVinney, Puente et al. 2001). Additionally it is known that attachment of EPEC to host cells relies on the presence of intimin-α and bundle forming pili which are also important in the formation of microcolonies. EHEC largely requires intimin-γ and long polar fimbriae for initial host cell adhesion (Croxen, Law et al. 2013, Pereira and Giugliano 2013). These differences may contribute to the differences seen in the anti-infective activity of the dMFGM fraction against the pathogens.

The anti-infective activity of the dMFGM fraction was demonstrated to be concentration dependent (Figure 3A) with an 80%, 79%, 70% and 44% reduction of bacterial association with HT-29 cells at a concentration of 5, 2.5, 1.25 and 0.625 mg/mL of dMFGM, respectively. A plateau effect was observed where maximum inhibition was reached at 2.5 mg mL\(^{-1}\). This would suggest that studies specifically targeting *E. coli* O157:H7 should use a dMFGM concentration of at least 2.5 mg mL\(^{-1}\).

The mechanism by which the dMFGM fraction causes the anti-infective effect may have occurred in a number of ways: dMFGM may be interacting with bacteria, preventing HT-29 cell association; dMFGM may be interacting with HT-29 cells, preventing bacterial interaction with the cell line; or a dual method which incorporates dMFGM interacting with both bacteria and HT-29
cells and reducing the association of bacteria with the cell line. To investigate these possible modes of action(s), two additional assays were performed. Firstly, a bacterial interaction assay was performed. Inhibition of *E. coli* NCTC 12900, *E. coli* DAF454 and *E. coli* DPC 6055 association with the HT-29 cells was observed at 28%, 30% and 25% respectively compared with the control (Figure 3B). This indicated that dMFGM likely interacted with the bacteria, thereby reducing the interaction of the bacteria with the HT-29 cells. However, the magnitude of inhibitory activity of dMFGM was reduced under these conditions. As the unbound material was removed in this assay, components of the dMFGM fraction binding directly to host cell surface receptors were potentially at least partially responsible for reducing the number of bacteria associating with the HT-29 cells. In addition, the removal of unbound dMFGM reduced the occurrence of competitive binding, thus increasing the likelihood of bacteria associating with HT-29 cells during natural association-dissociation. To determine whether the dMFGM interacted with HT-29 cells, a host cell interaction assay was performed. No reduction in *E. coli* O157:H7 association with the HT-29 cells was observed (Figure 3C). Similar findings were observed by other researchers when investigating the mode of action of their compounds of interest. For example, Manthey et al. (2014) pre-incubated HMO with HeLa, Hep-2 and T84 epithelial cell lines, resulting in a loss of anti-adhesive activity against EPEC. Inagaki et al. (2014) pre-incubated MA104 Rhesus monkey kidney cells with bovine κ-casein glycopeptides prior to rotaviral infection and demonstrated no viral inhibition, suggesting that bovine κ-casein glycopeptides do not bind host cell receptors. Furthermore, Marotta et al. (2014) demonstrated the loss of anti-infective activity of the milk oligosaccharide 3’-sialyllactose (3’SIL) against *Pseudomonas aeruginosa* association with human pneumocytes when the human cell line was pre-incubated with 3’SIL. Overall, the results presented here suggest that dMFGM does not target host cell receptors for *E. coli* O157:H7 and instead a direct dMFGM-bacterial interaction is likely responsible for the anti-infective activity. Bacterial environmental sampling and competitive binding may be an important factor here and may explain why
removal of unbound dMFGM led to a reduction in the anti-infective activity of the fraction. Bacterial association-dissociation which occurs *in vivo* may have led to the increased bacterial interactions with HT-29 cells due to the removal of unbound dMFGM.

Initially, screening studies were performed with the inclusion of a pre-incubation step which encouraged maximum exposure of the dMFGM to the bacteria. However, as this may not be a good representation of an *in vivo* situation, HT-29 cells were also infected with *E. coli* O157:H7 in the presence and absence of dMFGM without a pre-incubation step (Figure 3D). Overall, inhibition of *E. coli* adherence was still observed although at reduced magnitude compared with dMFGM pre-incubated with the bacteria. The inhibition observed was 19%, 47% and 5% for NCTC 12900, DAF454 and DPC 6055, respectively. These results suggest that the dMFGM fraction requires a certain period of time to exert its maximal inhibitory effect on *E. coli* cellular association. Indeed, a previous study indicated that optimum binding of a particular *E. coli* O157:H7 strain, CL-49, to mucins occurs at 37 °C for 2 h at pH 6.5 (Sajjan and Forstner 1990). Thus, the result observed in this experiment is likely due to the reduced time period in which dMFGM and *E. coli* strains were in contact, compared with the standard competition assay. This would lead to reduced competitive binding of dMFGM and HT-29 cell adhesins for *E. coli* O157:H7 cell receptors. The lack of a pre-incubation step reduces the ability of the dMFGM fraction to inhibit bacterial binding to host cells. However, it is interesting that a reduction was still evident instantaneously.

It is likely that the anti-infective role of dMFGM is due to its glycans, as highlighted by many other studies demonstrating the anti-infective potential of MFGM glycoconjugates such as mucins, Pas6, Pas7 and acidic glycolipids, with much bioactivity being attributed to sialic acid in particular (Yolken, Peterson et al. 1992, Shida, Takamizawa et al. 1994, Parker, Sando et al. 2010). The dMFGM fraction had the ability to prevent the association of EHEC strains, but not EPEC strains, with HT-29 cells. This could reflect the different
mechanisms and adhesins used by EHEC and EPEC strains in their colonisation of host cells. Further investigations are required to determine more information on the different glycan receptors used by each serotype which would allow the subsequent tailoring of anti-adhesives to target a wide variety of pathogens.

Ingestion of dMFGM fraction would lead to digestion of many of the dMFGM components prior to its interaction with intestinal epithelial cells that could alter its components. In fact, defatted bovine MFGM has previously been digested in vitro and SDS gels have been used to analyse the protein and glycoprotein composition post-digestion. Coomassie-stained gels demonstrated hydrolysis of MFGM proteins after 20 min of pepsin digestion; however, many bands corresponding to MFGM glycoproteins were observed on the PAS stained gel indicating incomplete hydrolysis and partial resistance to the action of pepsin (Le, Van de Wiele et al. 2012). In vitro intestinal digestion led to hydrolysis of all remaining MFGM glycoproteins with the exception of Muc1 (Le, Van de Wiele et al. 2012). Similarly, PAS staining of whole MFGM indicated resistance of CD36 and Muc1 to in vivo gastric digestion (Gallier, Cui et al. 2013) and potential resistance of Muc1 to in vivo intestinal digestion (Gallier, Zhu et al. 2013). Thus, the high molecular weight Muc1 appears to be most resistant to digestion. The conformation of MFGM glycoproteins may result in some glycans being unavailable to bind E. coli strains. Digestion of dMFGM fraction may release bioactive components, allowing access of bioactive glycans to bacterial receptors, contributing to further anti-infective activity of the fraction. Further studies are required to investigate the anti-infective potential of dMFGM after digestion has occurred.

In addition to the anti-infective activities evidenced in this study, MFGM glycoproteins have been shown to possess other health promoting activities. For example, such components have been shown to affect the metabolism of colonic microbiota, causing an increase in butyrate production, thus potentially reducing the risk of colon cancer (Struijs, Van de Wiele et al. 2013). Furthermore, Ito et al. (1993) demonstrated the ability of MFGM to
inhibit β-glucuronidase intestinal activity in mice which could also contribute to the anti-cancer activities of MFGM. The anti-infective potential of glycans and MFGM fractions in vivo has been studied. For instance, bovine plasma glycans have demonstrated the ability to reduce E. coli adhesion to calf intestine in vivo (Mouricout, Petit et al. 1990) and bovine milk 3’SL can reduce H. pylori infection in rhesus monkeys (Mysore, Wigginton et al. 1999). Furthermore, a HMO pool has been shown to reduce EPEC colonisation in mice (Manthey, Autran et al. 2014). Interestingly, a bovine MFGM fraction has been shown to reduce Listeria monocytogenes colonisation in a rat model (Sprong, Hulstein et al. 2012) and H. pylori colonisation in a mouse model (Wang, Hirmo et al. 2001). Although these studies indicate the anti-infective activity of the dMFGM fraction against EHEC could also be viable in vivo, further studies are required to validate this. This hypothesis requires further investigation.
Chapter II

Conclusions

Bovine MFGM is a source of glycoproteins with the potential for improving human health. Of the 8 *E. coli* strains screened in this study, anti-infective activity of dMFGM was demonstrated against 3 strains of EHEC. The assays performed suggested that dMFGM was most likely associating with the infecting bacteria and thereby inhibiting bacteria from associating with HT-29 cells. This activity was also shown to be concentration dependent and required pre-incubation in order for the dMFGM fraction to exert its maximum protective activity. The inclusion of this bioactive fraction in functional foods may be of great benefit to the general population, as well as immune-compromised individuals, including infants and the elderly and may offer an alternative approach to the more technically challenging isolation of pure MFGM components. As demonstrated for the first time in this study, dMFGM may offer an alternative approach to reduce *E. coli* O157:H7 infection in humans.
Chapter II

References


CDC. (2013). Multistate outbreak of Shiga toxin-producing Escherichia coli O157: H7 infections linked to organic spinach and spring mix blend. Atlanta, GA, USA: Centers for Disease Control.


Chapter II


Inagaki, M., S. Nagai, T. Yabe, S. Nagaoka, N. Minamoto, T. Takahashi, T. Matsuda, O. Nakagomi, T. Nakagomi, T. Ebina and Y. Kanamaru (2010). "The bovine lactophorin C-terminal fragment and PAS6/7 were both potent in the inhibition of human rotavirus replication in cultured epithelial cells and the
Chapter II

prevention of experimental gastroenteritis." Bioscience Biotechnology and Biochemistry 74(7): 1386-1390.


Chapter II


Chapter II


Chapter II


Figure 1. Electrophoresis profile of bovine milk fractions. Reducing 4-12% SDS-PAGE profiles of the dairy streams made from raw unpasteurised milk for unwashed total protein (A) and washed total protein (B). Lane 1 – skim milk; Lane 2 – cream; Lane 3 – buttermilk; M, molecular mass marker. Gels were stained with coomassie blue. Results are representative of duplicate experiments.
Figure 2. The effect of the dMFGM fraction at a concentration of 5 mg/mL on association of *E. coli* strains with human HT-29 cells. Control – no dMFGM. Data represent mean +/- standard deviation of 3 replicates. Error bars represent standard deviation. * indicates values are significant (P-value < 0.05).
Chapter II

A

![Graph A]

B

![Graph B]
Figure 3. The effect of the dMFGM fraction on association of *E. coli* O157:H7 with HT-29 cells. (A) The effect of different dMFGM concentrations on the anti-infective activity against *E. coli* NCTC 12900. (B) The effect of removing unbound dMFGM subsequent to pre-incubation of *E. coli* O157:H7 with 5 mg/mL dMFGM for 1 hr, on the association of *E. coli* O157:H7 with HT-29 cells. (C) The effect of removing unbound dMFGM subsequent to pre-incubating HT-29 cells with 5 mg/mL dMFGM for 1 hr, on the association of *E. coli* O157:H7 with HT-29 cells. (D) The effect of no pre-incubation of *E. coli* O157:H7 with 5 mg/mL dMFGM on the association of *E. coli* O157:H7 with HT-29 cells. Data represent mean +/- standard deviation of 3 replicates. Error bars represent standard deviation. * indicates values are significant (P-value < 0.05).
Chapter III

Temporal alterations in the bovine buttermilk glycome from parturition to milk maturation

Sarah A. Ross, Jared Q. Gerlach, Satbir K. Gill, Jonathan A. Lane, Michelle Kilcoyne, Rita M. Hickey, and Lokesh Joshi

Published in Food Chemistry Journal (2016)

Acknowledgements: Sarah Ross generated and fluorescently labeled all buttermilk samples, incubated samples on lectin microarrays and performed lectin blots. Lectin microarrays were synthesized by John Glavin and Michelle Kilcoyne and data analysis was performed by Sarah Ross, John Glavin and Jared Gerlach. Statistical analysis was performed by Satbir Gill and Jared Gerlach. Sarah Ross and Jared Gerlach performed lectin blotting analysis.
Chapter III

Abstract
The bovine milk fat globule membrane (MFGM) has many associated biological activities, many of which are linked with specific carbohydrate structures of MFGM glycoconjugates. Bovine buttermilk is a commercially viable source of MFGM and is an under-valued by-product of butter making. However, the changes in buttermilk glycosylation over the course of lactation have not been extensively investigated. In this study, buttermilk was generated from three individual multiparous cows at 13 time points over the first three months of lactation. Buttermilk glycosylation was profiled using lectin microarrays and lectin blotting. Suggested differences in glycosylation, including N-glycosylation, sialylation and fucosylation, were observed between early and late time points and between individual animals. Overall, these data suggest temporal changes in the glycosylation of buttermilk proteins which may have an important impact on commercial isolation of glycosylated ingredients.
Chapter III

Introduction

The milk fat globule membrane (MFGM) is a heterogeneous membrane which surrounds and stabilises milk fat droplets (Evers, Haverkamp et al., 2008) and is a rich source of proteins and lipids, many of which are glycosylated. The glycoconjugates of MFGM include glycoproteins such as mucins (Muc), butyrophilin (BTN), and Pas6/7, and glycolipids such as monosialogangliosides (GM), disialogangliosides (GD) and trisialogangliosides (GT) (Berglund, Petersen et al., 1996; Hvarregaard, Andersen et al., 1996; Pallesen, Andersen et al., 2001; Ross, Lane et al., 2015; Seok, Shimoda et al., 2001).

Many potential health benefits imparted by MFGM glycoconjugates have been demonstrated in vitro and in vivo. For example, MFGM can inhibit pathogenic colonisation and subsequent infection in the gut and MFGM glycoconjugates are likely to contribute to this anti-infectivity (Ross, Lane et al., 2016). Muc1 from human MFGM demonstrated prevention of Escherichia coli adhesion to buccal epithelial cells in vitro (Schroten, Hanisch et al., 1992) and Muc1 from bovine MFGM demonstrated inhibition of neuraminidase-sensitive rotavirus infection in MA104 cells (Kvistgaard, Pallesen et al., 2004). Specific carbohydrate structures on MFGM glycoconjugates have been demonstrated to play a role in the inhibition of microbial attachment. For example, sialylation of bovine Muc1 was suggested to reduce binding of E. coli and Salmonella enterica serovar Typhimurium to Caco-2 cells in vitro (Parker, Sando et al., 2010). In addition, deglycosylation of a complex of human milk mucin led to a reduction in rotavirus inhibitory activity which indicated the potential importance of glycosylation in imparting the anti-rotaviral activity (Yolkken, Peterson et al., 1992).

Changes in milk glycosylation and protein abundance occur as lactation progresses, with concentrations and specific structures differing between colostrum and mature milks (Takimori, Shimaoka et al., 2011; Wilson, Robinson et al., 2008). For instance, Muc1, Muc15, adipophilin and BTN are upregulated in bovine MFGM 7 days post-partum compared to
colostrum (Reinhardt and Lippolis 2008) while the abundance of immunoglobulin G (IgG) and lactoferrin are higher in milk colostrum than in samples taken later in lactation. Additionally, concentrations of sialylated and highly fucosylated glycoproteins in bovine milk are highest in colostrum compared to milk sampled at later time points and the ratio of $N$-glycolylneuraminic acid (Neu5Gc) to $N$-acetylneuraminic acid (Neu5Ac) is significantly higher in colostrum, and decreases gradually thereafter (Takimori, Shimaoka et al., 2011). Glycosylation profiles of individual bovine milk glycoproteins have also demonstrated changes in glycosylation over the course of lactation (O’Riordan, Gerlach et al., 2014; Ujita, Furukawa et al., 1993) as have MFGM glycoprotein components. For example, MFGM glycoproteins from one Holstein cow’s milk at days 0, 1, 3 and 5 post-parturition was analysed for glycosylation changes by soybean agglutinin (SBA) blotting, which binds to $N$-acetylgalactosamine (GalNAc), which revealed that binding to a glycoprotein, believed to be CD36, increased over the six day period assessed (Ujita, Furukawa et al., 1993).

While changes in MFGM glycosylation over the course of lactation are expected, few investigations have been undertaken to date. Previous studies have focused on MFGM from milk sampled at just a few time points (Reinhardt and Lippolis 2008) or from a single animal using a small number of lectins (Ujita, Furukawa et al., 1993). Other studies have compared bovine milk from early time points and late time points but did not monitor the potentially critical changes in between (Wilson, Robinson et al., 2008). Since glycosylation of bovine MFGM components is an important factor contributing to their health-promoting activities, further knowledge on how lactation affects this glycosylation may be of commercial importance. Identification of lactation time points associated with the most glycosylated components may aid in generation of bovine milk fractions best suited for use as functional ingredients. For instance, since glycoproteins such as BTN are known to be conserved across MFGM from human and bovine milk (Lu, Wang et al., 2016), it may be of benefit to utilise highly glycosylated bovine milk ingredients to enrich infant formula since infant formula often has less free oligosaccharide content than human milk. This may help to narrow the
gap between formula composition and breast milk. Buttermilk is a viable commercial source of MFGM and its associated functional glycoconjugates which have the potential to further benefit human health if used as nutritional food additives. To the best of our knowledge, changes in bovine buttermilk glycosylation over an extended lactation period in multiple animals have not been investigated.

In this study, buttermilk was generated from milk sampled from three animals at 13 time points (days 1 to 10 (D1 to D10) and days 30, 70 and 90 (D30, D70 and D90) post-partum) from colostrum to mature milk. The glycosylation of the individual buttermilk samples was profiled using lectin microarrays featuring a panel of 43 lectins, electrophoretic analysis and lectin blotting.
Chapter III

Materials and Methods

Materials

The bicinchoninic acid (BCA) Protein Assay Kit and the SuperSignal Pico kit were from Pierce Biotechnology (Thermo Fisher Scientific Inc., Dublin, Ireland). NuPAGE 4-12% Bis-Tris gels, MOPS buffer, and carboxylic acid succinimidyl ester Alexa Fluor® 647 (AF647) were purchased from Life Technologies (Carlsbad, CA). Molecular mass ladder (Mark12 Unstained Standard), LDS sample buffer, antioxidant and Coomassie Brilliant Blue (SimplyBlue SafeStain) were from Thermo Fisher Scientific (Carlsbad, CA). Pure, unlabelled lectins were acquired from EY Laboratories, Inc. (San Mateo, CA) or Vector Laboratories, Ltd. (Orton Southgate, UK) (Table S1). Biotinylated lectins (RCA-I, MPA, LTA, WGA, WFA, LEL and AIA) were from EY Laboratories, Inc. Avidin-D horseradish peroxidase (HRP) conjugate was from Vector Laboratories, Ltd. Nexterion® Slide H microarray slides were from Schott AG (Mainz, Germany). Glycoprotein Detection Kit, streptavidin horseradish peroxidase conjugate, bovine serum albumin (BSA), bovine asialofetuin (ASF), bromochloroindophosphate nitroblue tetrazolium (BCIP/NBT) were from Sigma-Aldrich Co. (Dublin, Ireland). Neu5Ac monosaccharide was obtained from Dextra Ltd. (Reading, U.K.). Polyvinylidene fluoride (PVDF) membranes (0.2 µm) were from Merck-Millipore Corp., (Dublin, Ireland). Unless otherwise noted, all other reagents were from Sigma-Aldrich Co. and were of the highest grade available.

Sample collection

Morning milk was collected daily from three multiparous Holstein-Friesian cows (animals 1, 2 and 3) at Teagasc Research Centre, Moorepark, Fermoy, Co. Cork. Samples were collected from D1 to D10 and at D30, D70, and D90 post-parturition per animal. Milk from D9 for animal 1 and D2 for animal 3 was not collected due to scheduling conflicts.
Chapter III

Buttermilk generation

Buttermilk generation was adapted from Morin et al. (2007). Briefly, samples were incubated at 45 °C for 1 h immediately after collection, followed by cream separation from whole milk using an FT15 disc bowl centrifuge (Armfield Ltd., Ringwood, England). The fat content was determined using Milkoscan FT120 (FOSS, Denmark) and fat content was adjusted to 40% using skim milk. Cream was stored at 4 °C for 24 h. Buttermilk and butter were generated by agitating the cream using a food mixer. The buttermilk was passed through glass wool twice to remove butter granules. Samples were frozen at -20 °C, freeze-dried and stored in a desiccator at room temperature (RT) until further use.

Characterisation of buttermilk samples

Protein concentration was determined using the BCA Protein Assay Kit (Smith, Krohn et al., 1985), with BSA as the standard. Carbohydrate content was assayed by the Monsigny method (Monsigny, Petit and Roche, 1988) using glucose (Glc) as the standard. Total sialic acid content was determined using the periodate-resorcinol assay (Jourdian, Dean and Roseman, 1971), using Neu5Ac as the standard. Assays for all samples were carried out in triplicate and the mean value is reported.

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis

Dithiothreitol-reduced buttermilk samples were electrophoresed in 4-12% Bis-Tris gels. Briefly, 65 µg protein was loaded for each sample with a molecular mass ladder in a single lane and all samples were diluted 1:10 with Lithium dodecyl sulfate sample buffer. The gel was resolved at 200 V for 50 min using NuPAGE MOPS buffer. NuPAGE MOPS buffer containing 0.25% NuPAGE antioxidant was used in the upper chamber. Protein bands were visualized on the gels using Coomassie Brilliant Blue following the manufacturer’s procedure. Glycoprotein bands were visualised using a Glycoprotein Detection Kit. ImageJ software
Chapter III

(http://rsb.info.nih.gov/ij/index.html) was used for relative quantitation of gel bands by densitometric analysis.

**Fluorescent labeling of MFGM and glycoproteins**

The bovine MFGM samples and the ASF standard for lectin microarrays were labelled with AF647 ($\lambda_{ex}$ 650 nm, $\lambda_{em}$ 665 nm) in 100 mM sodium bicarbonate, pH 8.3. Briefly, 10 µL of AF647, dissolved in DMSO, was added to 500 µL of sample (2 mg/mL) and incubated for 1 h in the dark at RT. Samples were kept in the dark after this point. Excess dye was removed from the labelled bovine buttermilk samples on a Bio-Gel P-6 column (1 x 12 cm) (Bio-Rad Laboratories, Ltd., Hertfordshire, U.K.) eluted with phosphate buffered saline (PBS), pH 7.4. Absorbance at 647 and 280 nm for each sample was measured and the protein concentration and degree of substitution was calculated according to manufacturer’s instructions. Extinction coefficients of 100,000 and 19,844 M$^{-1}$ cm$^{-1}$ and molar masses of 100,000 and 48,400 g/mol were used for the calculations for the samples and ASF (based on fetuin) (Spiro 1960), respectively.

**Construction of lectin microarrays and MFGM profiling**

Microarrays consisting of 43 lectins with reported carbohydrate binding specificities (Table S1) were constructed as previously described (Gerlach, Kilcoyne and Joshi, 2014; O'Riordan, Gerlach et al., 2014). Each feature was printed on Nexterion® Slide H microarray slides in replicates of six and eight replicate subarrays were printed per slide. Lectin performance after printing was tested by incubation with fluorescently-labelled glycoprotein standards (Gerlach, Kilcoyne and Joshi, 2014). The lectin microarrays were stored at 4 °C with desiccant until use.

Microarray slides were incubated in the dark using an eight-well gasket slide and incubation cassette system (Agilent Technologies, Cork, Ireland) as previously described (Gerlach, Kilcoyne and Joshi, 2014; O'Riordan, Gerlach et al., 2014). Initially, four randomly selected fluorescently labelled MFGM samples were titrated (5 to 20 µg/mL) on the lectin microarrays for optimal signal to noise ratio and a concentration of 10
µg/mL in Tris-buffered saline (TBS; 20 mM Tris-HCl, 100 mM NaCl, 1 mM CaCl₂, 1 mM MgCl₂, pH 7.2) with 0.05% Tween® 20 (TBS-T) was selected for all fluorescently labelled MFGM samples. Three replicate slides were incubated per experiment with appropriate haptenic carbohydrates (100 mM) co-incubated in parallel for a subset of samples to verify carbohydrate-mediated binding (Gerlach, Kilcoyne et al., 2011). ASF labelled with AF647 (0.5 µg/mL TBS-T) was included in one subarray per experiment to verify the presence and retained function of the printed lectins. Dried microarrays were scanned immediately in an Agilent G2505B microarray scanner using the Cy5 channel (633 nm excitation, 80% PMT, 5 µm resolution) (Figure S1).

Data extraction, hierarchical clustering and statistical analysis

Data extraction was performed essentially as previously described (Gerlach, Kilcoyne and Joshi, 2014, O'Riordan, Gerlach et al., 2014). For graphical representation, data were normalized to the mean total intensity value of three replicate microarray slides and binding data was presented in histogram form of average intensity with one standard deviation of experimental replicates (n=3; 18 total data points per probe). Heat mapping and unsupervised hierarchical clustering of normalized data was performed with Hierarchical Clustering Explorer v3.0 (http://www.cs.umd.edu/hcil/hce/hce3.html) using the following parameters: no pre-filtering, complete linkage, Euclidean distance.

The normalized, unfiltered lectin microarray data also was subjected to multi-scale bootstrap re-sampling prior to hierarchical clustering analysis in the R ‘pvclust’ environment (http://cran.r-project.org/web/packages/pvclust/index.html). Estimation of certainty for each of the p-value calculations for clusters was obtained from the bootstrapped data. The matrix was generated using the method.dist = ‘euclidean’, cluster method using helust function = ‘average’ and 10,000 bootstraps where with relative sample size was fixed from 0.5 to 1.4,
incrementing in steps of 0.1. Using this matrix, approximately unbiased p-values were calculated for every node.

To further refine the hierarchical clustering, a threshold cut-off filter was also used to identify the most critical subsets of lectin data for clustering and statistical analysis similar to that previously reported (Gerlach, Kilcoyne and Joshi, 2014, 2014). For these, a minimum deviation value (1000 relative fluorescence units (RFU)) was applied resulting in the selection of 12 lectins (AIA, SNA-II, WFA, GSL-II, LEL, GNA, HHA, WGA, PCA, LTA, RCA-I, MPA) which most strongly supported the hierarchy. These data were again statistically evaluated within the R environment using the same parameters listed above.

**Lectin blotting analysis**

After SDS-PAGE, proteins were transferred to 0.2 µm PVDF membranes using semi-dry transfer for 120 min at 15 V. Membranes were blocked overnight in 1% BSA in TBS-T at 4 °C and all following procedures were carried out at RT. The membranes were washed three times in TBS-T for 5 min per wash. All subsequent washes were 5 min per wash. All biotinylated lectins were diluted in TBS-T with 0.05% BSA (final concentrations were 2.5 µg/mL for RCA-I and 2 µg/mL for MPA, LTA, WFA, AIA, LEL and WGA). Each membrane was incubated in 10 mL biotinylated lectin with gentle shaking for 90 min. Membranes were washed three times in TBS-T. For RCA-I, MPA, LTA and WFA, the streptavidin alkaline phosphatase conjugate (1 mg/mL stock concentration) was incubated with the membrane at a dilution of 1:15,000 with gentle shaking for 1 h. The membrane was washed three times in TBS-T and once in TBS. BCIP/NBT solution (1 tablet in 10 mL water) was added to the membrane and incubated until colour enhancement was adequate for visualization (approximately 2 min). Colour development was halted by rinsing in distilled H₂O. Membranes were allowed to fully dry prior to imaging.

For AIA, LEL and WGA, previously probed membranes were stripped by incubating membranes twice for 5 min in 100 mM glycine supplemented with 1% (v/v) Tween® 20 and 3.5 mM SDS, pH 2.2.
Chapter III

Membranes were then washed three times in TBS before blocking again in 1% BSA in TBS-T for 1 h. Each membrane was washed and probed again with one biotinylated lectin as described above. Membranes were washed as above before incubation in 1:20,000 Avidin-D HRP conjugate in TBS-T for 1 h. Lectin-binding glycoprotein bands on the membranes were visualised by chemiluminescence using ECL substrates diluted in a 1:1 ratio. All membranes were imaged using an Image Station 4000 MM imaging system (Carestream Health, USA). ImageJ software (http://rsb.info.nih.gov/ij/index.html) was used for lectin blot band quantitative densitometric analysis.
Results and discussion

Characterisation of bovine buttermilk over 3 month lactation

The total protein content of all buttermilk samples from three animals from D1 to D90 post-parturition was estimated by the BCA method (Figure 1A) and the distribution of protein was analysed by SDS PAGE (Figure 1B and Figure S2A). For all three animals, protein content was highest at D1 post-partum (760-893 µg/mL) but decreased by approximately 46% by D2, and steadily decreased further until D10. Similarly, Ujita et al., (1993) previously reported a bovine MFGM protein concentration decrease over the first 5 days of lactation. After D10, protein concentration was fairly constant (approximately 310 µg/mg) until the end of the sampling period (D90) (Figure 1A).

Coomassie Brilliant Blue staining of the SDS-PAGE gels revealed approximately 16 protein bands in the buttermilk samples (Figure 1B and Figure S2A). High intensity bands were visible at approximately 34, 31, 26, 19 kDa. Protein bands demonstrating lower intensity were observed at 10, 45, 50, 54, 60 and 130 kDa. Low intensity bands were detected at approximately 40, 50, 64, 66 80 and 150 kDa. Rombaut et al. (2007) demonstrated a similar variety of protein bands by Coomassie-stained SDS-PAGE analysis of acidified buttermilk. Interestingly, the relative band intensity changed as lactation progressed highlighting differences in the buttermilk composition as lactation progressed. For example, the relative abundance of the 130 kDa band increased over time (Figure 1B and Figure S2A). Changes in relative abundance of bovine milk proteins over lactation have been previously demonstrated. For example, bovine milk xanthine dehydrogenase concentrations were higher in milk taken 7 days post-parturition compared to colostrum (Reinhardt and Lippolis 2008).

Densitometric analysis of the Coomassie-stained gel from animal 1 suggested that the relative abundances of the proteins at 45 and 50 kDa in the buttermilk samples remained relatively unchanged over the post-parturition timeline (Figure S3), with the exception of D8 and D90 where
relative protein abundances were highest and lowest, respectively. Interestingly, differences in relative band intensities were evident between individual animals (Figure S2A). For example, the increase in intensity of the 60 kDa protein band at D7 and D9 for animal 3 was not observed for animal 1. However, overall relative protein band intensities for animal 1 were most similar to those of animal 3. In contrast, animal 2 had a greater number of band intensities which differed from the other two animals. For instance, band intensities for 130 kDa protein band increased over the course of lactation for animals 1 and 3. However the relative intensity of the 130 kDa band for animal 2 decreased at D3 and D4, increased from D6 to D8 and decreased again at D9. Thus the relative compositions of the buttermilk samples varied from animal to animal over the course of lactation. Studies by Ye et al. (2002) also observed differences in bovine MFGM protein abundance throughout the course of lactation through the use of Coomassie-stained gels. The authors noted differences in the abundance of proteins including BTN, Pas 7 and xanthine oxidase in milk sourced from early, mid and late season milks.

PAS staining of SDS-PAGE gels (Figure 2A and Figure S2B) demonstrated the presence of numerous glycoproteins in buttermilk throughout lactation. Similar to the Coomassie-stained gels, the relative band intensity changed throughout the progression of lactation, highlighting changes in glycoprotein abundance may have occurred as milk matured. The presence of glycoproteins in buttermilk was not surprising as it has been documented previously (Hvarregaard, Andersen et al. 1996, Pallesen, Pedersen et al. 2007, Ross, Lane et al. 2016, Takamizawa, Iwamori et al. 1986), however it was interesting to note the change in intensity as lactation progressed.

Furthermore, comparison of the PAS-stained gels suggested the relative abundances of individual glycoproteins in the buttermilk samples varied between individual animals as lactation progressed (Figure S2B). Similar findings have been shown for human milk oligosaccharides which incorporate fucose and sialic acid (Miller, Bull et al., 1994). In addition, the acidic and neutral sugar content of buttermilk was assayed based on mg of
lyophilised buttermilk. The sialic acid concentration in buttermilk was highest for all three animals on D1 at 192-237 nmoles/mg and decreased gradually until D10 when it levelled off between 119-151 nmoles/mg and maintained this concentration into milk maturity. Overall the data suggested a time-dependent fluctuation of sialic acid in the buttermilk samples.

A rapid increase in neutral sugar concentration was observed from D1 to D2 (1857 and 1880 nmoles/mg for animals) in animals 1 and 2, respectively, over the first 10 days and then increased more gradually up to D10 for all three animals. Neutral sugar concentration plateaued by D30 and remained constant through milk maturation (2166 – 2533 nmoles/mg) (Figure 2C). It should be noted that the increase in neutral sugar concentration observed could have resulted due to the known increase in lactose which occurs throughout milk maturation (Ontsouka, Bruckmaier and Blum 2003). The decline in sialic acid concentration over the course of lactation may reflect the role that sialylation plays in protection of the early postnatal gut from pathogen adhesion (Kelly and Coutts 2000). The rise in concentration of neutral sugars after parturition may reflect a role for nutrition rather than protection and may also contribute to the growth of beneficial bacteria such as Bifidobacteria and Lactobacilli in the gut (Lönnerdal 2003).

**MFGM glycoprofiling on lectin microarray**

To establish the optimum concentration for lectin microarray profiling, four randomly selected, labelled buttermilk samples were titrated on the lectin microarray consisting of a panel of 43 immobilised lectins of established binding specificities (Table S1). A sample concentration was chosen such that the response of the majority of the lectins was within the dynamic range of the microarray scanner (0-65,000 RFU, approximately). Subsequently, all 37 fluorescently labelled samples were profiled on lectin microarrays at the optimal concentration in triplicate (Figure 4).

Lectin microarray data was examined for similarities between animals and time points. Unsupervised hierarchical analysis of the complete
lectin microarray data set was performed without (Figure 3A) and with (Figure 3B) bootstrapping to obtain $p$-values for the different cluster relationships calculated in the R environment using the pvclust package. Animals 1 and 3 displayed the greatest similarity to one another at the greatest number of time points and animal 2 was indicated to be less similar to the others (Figure 3B). After the initial hierarchical clustering, filtering was applied to select only data from lectins that significantly contributed towards the hierarchical relationship. Increasing the minimum intensity cut-off to exclude all lectins demonstrating <1000 RFU deviation resulted in the selection of just 12 lectins (AIA, SNA-II, WFA, GSL-II, LEL, GNA, HHA, WGA, PCA, RCA-I, LTA, MPA). Including only data from these in the clustering calculations greatly strengthened the certainty of the dendrogram (Figure 3C). The D1 data for all animals clustered together on one branch and were distant from the remaining data, which suggested that these profile differences were most significant compared to profiles of buttermilk MFGM derived from the following days’ samples. Profiles generated for samples later in lactation did not cluster with any strict pattern, thus demonstrating the heterogeneity of the profiles at later time points. As well as more firmly positioning D1 MFGM as being quite different than later samples, filtering the lectin profile data also tightened the overall relationship between the animals, which resulted in them appearing to be more similar overall than the unfiltered data suggested (Figure 3C).

On the lectin microarrays, buttermilk samples bound to mannose (Man)-binding lectins Con A, GNA and HHA (Table S1) (Figure 4). Man residues are attached to the chitobiose (GlcNAc-β-(1,4)-GlcNAc) core of all N-linked glycans and Man residues are more abundant in hybrid and high-Man type N-linked structures. Goat MFGM proteins including BTN, Pas6/7 and mucins have also been demonstrated to bind ConA lectin (Cebo, Caillat et al., 2010). Colostrum (D1-3) buttermilk samples had relatively higher binding to GNA and HHA compared to mature (D30-90) buttermilk samples. The data suggested that animal 2 may have a greater proportion of high Man type N-linked structures in its colostrum buttermilk than animals 1 and 3.
PCA lectin predominantly binds to biantennary N-linked glycans (Table S1). Based on PCA binding, the proportion of biantennary structures was much greater in colostrum than in transitional and mature buttermilk samples for all three animals (Figure 4). The binding intensities of buttermilk samples with PCA did not differ much between the later time points with the exception of samples from animal 2, which increased in binding intensity at D5 and D6 of lactation (Figure S5B). No binding was observed for PHA-E and PHA-L lectins which have an affinity for extended biantennary and bisecting GlcNAc motifs, and tri- and tetraantennary N-linked complex structures, respectively (Table S1).

The lectins GSL-II and STA have an affinity for GlcNAc residues and GlcNAc oligomers, respectively while LEL has been reported to have greatest affinity for the chitobiose core of N-linked structures (Table S1). LEL bound to all buttermilk samples with low to intermediate intensity (Figure 4). The binding intensity of GSL-II for all samples from all animals was reduced in mature buttermilk compared to colostrum (Figure 4 and Figure S4), which suggested a lower prevalence of terminating GlcNAc structures as lactation progressed. The binding intensities of LEL with buttermilk samples from animals 1 and 3 changed very little over lactation. However the D1 sample from animal 2 bound intensely to LEL and samples from later timepoints had decreased binding intensity (Figure 4 and Figure S4). The binding intensities of samples with the lectins PCA and RCA-I (both with affinity to Type II LacNAc (Table S1)) were also highest for the D1 sample from animal 2 compared to animals 1 and 3 (Figure 4 and Figure S4). Binding of bovine MFGM glycoproteins to lectins such as RCA-I has been reported previously (Ujita, Furukawa et al., 1993), and thus it was not surprising to see binding in this study also.

WGA binding indicated the presence of GlcNAc and/or sialic acid (Neu5Ac and/or Neu5Gc) (Table S1) and all samples exhibited binding intensity with WGA. In this study, a sinusoidal binding pattern, i.e. the binding intensity successively increased and decreased repetitively, was observed for all samples with WGA lectin as lactation progressed (Figure S4 and Figure S5A). Relatively low binding intensities for all samples with
SNA-I indicated the presence of α-(2,6)-linked sialic acid (Table S1 and Figure 4). SNA binding to MFGM components has been observed previously for goat MFGM glycoproteins (Cebo, Caillat et al., 2010). It has previously been demonstrated that bovine whole milk sialylation decreased during the transition from colostrum to mature milk (Takimori, Shimaoka et al., 2011) and bovine MFGM sialylation decreased over the first 5 days of lactation (Ujita, Furukawa et al., 1993).

Binding of samples to Lch-B and PSA lectins (Table S1) indicated the presence of fucosylated high Man structures in the buttermilk (Figure 4). Very intense LTA binding by all samples for all animals and lower intensity AAL binding (Figure 4) indicated the presence of Fuc in α-(1,3)- and α-(1,6)-linkages, respectively (Table S1). Although only low binding was observed for all samples with Lch-B and PSA, binding intensities were highest for samples from early lactation and decreased thereafter. Binding intensities of samples with both LTA (Figure S5C) and AAL also decreased as lactation progressed (Figure 4 and Figure S4) which indicated a lower proportion of Fuc in mature buttermilk compared to colostrum buttermilk. Similarly, concentrations of fucosylated N-linked glycans were reported to be most abundant in bovine colostrum and decreased thereafter (Takimori, Shimaoka et al., 2011). This may indicate a possible requirement of Fuc residues in specific linkages for prevention from pathogenic infection in early infancy (Morrow, Ruiz-Palacios et al., 2004).

The presence of non-sialylated, O-linked mucin core type 1 glycans is often indicated by binding of PNA, which has highest affinity for T-antigen (Gal-β-(1,3)-GalNAc), and by the lectins ABL and ACA, which also have an affinity for mucin core-type structures (Table S1). The presence of O-linked glycosylation may also be indicated by the presence of Gal, as observed by low binding by the lectin AIA (Figure 4 and Table S1). For the buttermilk samples, no binding was observed for PNA and ABL. However, very low binding with ACA was observed, with lowest binding in colostrum samples for all three animals (Figure 4). Since ACA lectin binding was evident coupled with no binding by PNA and ABL, this indicated the presence of sialylated T-antigen as PNA and ABL do not bind to the T-
antigen when it is sialylated. AIA binding intensities were lower in mature samples compared to colostrum for animals 1 and 3 (Figure S5D). Animal 2 differed as binding intensities did not change much over the course of lactation, with the exception of buttermilk from D6 which bound to AIA with the greatest intensity for all sample time points for animal 2 (Figure S5D). This binding intensity pattern over the course of lactation indicated an overall lower abundance of terminal Gal residues in the buttermilk samples of animal 2 compared to animals 1 and 3, particularly in colostrum buttermilk.

Low intensity binding was observed for all samples with the lectins VVA-B4 and WFA (Figure 4), which suggested the presence of GalNAc (Table S1), with lowest binding for D1 samples for all animals (Figure S4). WFA binding intensity increased in D2 to D3 samples and binding intensities followed a sinusoidal trend in the days thereafter with a slight increase in the mature milk samples of animals 1 (D70 and D90) and 3 (D30 and D70), but not of animal 2 (Figure S5F). Additionally, low binding of the samples with SNA-II (Figure 4), which also binds to GalNAc residues (Table S1), likely indicated the presence of LacNAc and N,N’-diacetyllactosamine (GalNAc-β-(1-4)-GlcNAc, LacdiNAc). RCA-I binding to the samples was observed with buttermilk at early time points (e.g. D1) with lower binding observed for samples from later time points (e.g. D90) (Figures 4 and S5E). RCA-I typically binds to type II LacNAc (Gal-β-(1,4)-GlcNAc) (Table S1). Interestingly no binding was observed for the samples with ECA (Figure 4), which binds to type II LacNAc oligomers (Table S1). SNA-II binding displayed a sinusoidal binding trend for samples over the course of lactation for all animals (Figure S4). Interestingly, relative binding intensities of the samples with SBA, which has an affinity for GalNAc (Table S1), increased from D1-3 for animal 1 and increased from D1-5 for animals 2 and 3. This trend was broadly in agreement with Ujita et al. (1993) who reported an increase in SBA binding intensities over the first 6 days of lactation (Figure S6).

Very high binding by MPA lectin to all samples (Figure 4) indicated the presence of terminal α-linked Gal (Table S1) and the binding intensity
varied sinusoidally over time. Overall, MPA binding intensities were among the highest lectin binding intensities for samples for animals 1 and 3. Most prominently, the relative binding intensity for MPA with MFGM from animal 2 was markedly lower compared to animals 1 and 3 for much of the early portion (D1-D10) of the lactation time course (Figure S5G). For animal 2, binding to MPA lectin was lowest in colostrum (D1-D3) and highest in late transitional buttermilk (D9) while MPA binding decreased slightly in mature buttermilk (D30-D90). For animals 1 and 3, MPA binding intensity varied sinusoidally over time. These results may suggest that Gal residues play a larger role in nutrition throughout the course of lactation rather than in the protection of the infant gut from infection. For instance, the presence of Gal throughout lactation may be beneficial for establishing a Bifidobacteria-rich flora, given that bacteria such as *B. bifidum* are known to utilise Gal as a carbon source, in the infant gut (Asakuma, Hatakeyama et al., 2011).

The global changes observed for bovine buttermilk throughout lactation were diverse and expected since buttermilk is a source of many glycosylated components (Ross, Lane et al., 2015). For example, N-linked glycosylation has been reported on many MFGM components including mucins (Pallesen, Nielsen et al., 2001), BTN (Sato, Takio et al., 1995), Pas6/7 (Hvarregaard, Andersen et al., 1996) and CD36 (Nakata, Furukawa et al., 1993). Bovine MFGM is also known to contain predominantly mucin core type 1 O-linked structures with core type 2 structures present up to three days post-parturition (Wilson, Robinson et al., 2008). Furthermore, BTN (Sato, Takio et al., 1995) and α-lactalbumin (Takimori, Shimaoka et al., 2011) contain LacdiNAc while PP3 contains LacNAc and LacdiNAc (Inagaki, Nakaya et al., 2010). GM2 and GD2 also contain terminating GalNAc residues (Jensen and Newburg 1995; Smilowitz, Lebrilla et al., 2014). Additionally, sialylated buttermilk glycoproteins include Pas6/7 (Seok, Shimoda et al., 2001), Muc1 (Pallesen, Nielsen et al., 2001), Muc15 (α-(2,6)-linked sialylation) (Pallesen, Pedersen et al. 2007), IgG (α-(2,6)-linked sialylation) (Takimori, Shimaoka et al., 2011) and the gangliosides, including GM (α-(2,3)-linked sialic acid), and GD (α-(2,3)- and α-(2,8)-
linked sialic acid) (Lee, German et al., 2013; Smilowitz, Lebrilla et al., 2014). Fucosylation has also been documented for bovine buttermilk with Pas7 containing α-(1,6)-linked Fuc (Seok, Shimoda et al., 2001). Additionally, CD36 (Berglund, Petersen and Rasmussen, 1996), PP3 fraction (Inagaki Nakaya et al., 2010) and mucins (Pallesen, Pedersen et al. 2007) were reported as potentially modified with Fuc.

Lectin blotting analysis of bovine buttermilk samples

Proteins from buttermilk samples generated from animal 1 were transferred to PVDF membrane and incubated with seven (RCA-I, MPA, LTA, WFA, AIA, LEL and WGA) of the 12 lectins identified by threshold filtering of the lectin microarray data (Figure 5 and Figure S7). Day by day, relative binding patterns for the lectins RCA-I, WFA and LEL with the samples were similar to those for the lectin microarrays. RCA-I demonstrated high binding to bands from early day samples (D1 - D6), with particularly high binding on D1 for most bands. Binding dropped again and increased slightly in later time points (D10 - D90). No binding was evident to the band at approximately 30 kDa (Figure S7A). WFA showed a similar pattern to that of RCA-I but with less binding overall to bands below approximately 40 kDa (Figure S7B). LEL lectin bound more intensely to proteins of transitional milk (D4 - D10) than those from other time points, with high binding intensity at D6 (Figure 5A), in agreement with lectin microarray data. Densitometric quantification of LEL lectin blot supported these observations (Figure S8A) and the relative densities were highest for band 2 and band 3 (approximately 66 and 46 kDa, respectively).

The relative binding intensities for lectins MPA, LTA, AIA and WGA differed over time to those of the lectin microarrays. This is likely due, in part, to the absence of glycolipids in the blots which would be included in samples profiled on the microarray. Binding by MPA increased at the later time points for certain bands, such as at approximately 130 kDa, while binding decreased in later time points for other bands, such as those at approximately 31 and 66 kDa (Figure S7C). To our knowledge, terminal α-
linked Gal is not found on bovine milk glycolipids (Ito, Tominaga et al., 2012; Jensen and Newberg 1995; Martín, Martín-Sosa and Hueso 2001; Takamizawa, Iwamori et al., 1986). Thus, the differences seen between the MPA blots and arrays could be simply a result of differences in the intact MFGM glycosylation accessibility with the lectin microarray approach.

High intensity bands in samples from later time points in the LTA lectin blot, such as those at approximately 80 and 130 kDa, were in contrast to lectin microarray data which showed high LTA binding in colostrum which declined as lactation progressed (Figure 5B). AIA lectin binding indicated no binding at bands below approximately 55 kDa (Figure S7D). High intensity binding was evident at some bands for colostrum, for example D1 at approximately 55 and 116 kDa, and high intensity binding was evident for other bands later in lactation, for example at approximately 97 kDa (Figure S7D). Similarly, densitometric analysis of AIA blot revealed highest binding of AIA at D1 for 2 bands – band 2 and 5 which correspond to the bands at approximately 116 and 55 kDa, respectively (Figure S8B). For all other bands, the analysis demonstrated binding was highest in transitional and mature time points. However this data did not correlate strongly with the lectin microarray data. For instance, the relative binding at D8 and D10 was low in the lectin microarray data but was higher in the lectin blot. WGA in general had increased binding to bands as lactation progressed (Figure S7E). Bands at approximately 45, 50 and 64 kDa displayed the most intense binding overall. For the band at 64 kDa highest binding was evident at the late transitional and mature time points while for the bands at 45 and 50 kDa binding was highest at D10 and 30.
Chapter III

Conclusions

Lectin analysis demonstrated that temporal changes of bovine buttermilk glycosylation occurred during milk maturation which may vary between individuals. Future investigations would benefit from including a greater number of animals in order to fully investigate the potential for variation among individuals. Also examining different breeds of animals and varying feeding systems should provide additional insights into factors affecting glycosylation of buttermilk proteins. In this study, glycosylation differed to such a remarkable extent that exploitation of buttermilk at different stages of lactation could provide functional ingredients which could be tailored to target health promotion in a variety of areas, such as to narrow the gap between infant formula and human breast milk.
Chapter III

References


Chapter III


Chapter III


Figure 1. Protein characterisation of buttermilk samples over the course of lactation. Protein concentration was determined using bicinchoninic acid (BCA) protein assay for all buttermilk samples from all animals (A). Buttermilk samples from animal 1 from D1 to D90 post-partum separated on a NuPAGE 4-12% SDS PAGE gel under reducing conditions and Coomassie-stained for total protein (B).
Figure 2. Acidic and neutral sugar characterisation of buttermilk samples over the course of lactation. Buttermilk samples from animal 1 from D1 to D90 post-partum separated on a NuPAGE 4-12% SDS PAGE gel under reducing conditions and PAS-stained (A). Sialic acid concentration (B) for all buttermilk samples from all animals determined using periodate-
Chapter III

resorcinol assay and carbohydrate content (C) determined using Monsigny method.
Figure 3. Heatmap and hierarchical clustering results for lectin microarray profiles generated for labelled buttermilk samples. Data is represented as X_Y where X = animal (1, 2 or 3) and Y = timepoint post parturition. (A) Total intensity mean normalized data subjected to unsupervised, Euclidean
Chapter III

distance, complete linkage clustering without bootstrapping. (B) Bootstrapped clustering of Euclidean distance by Ward method using the entire lectin microarray data set. (C) Bootstrapped clustering of Euclidean distance by Ward method using only data which qualified after 1000 RFU minimum deviation filtering (AIA, SNA-II, WFA, GSL-II, LEL, GNA, HHA, WGA, PCA, LTA, RCA-I, MPA). Approximately Unbiased (AU) p-value and Bootstrap Probability (BP) value.
Chapter III

A. Colostrum (D1) buttermilk

B. Transitional (D8) buttermilk
Figure 4. Lectin microarray profiles of bovine buttermilk generated for three animals over the course of lactation. Each histogram represents the differences in recognition of lectins by fluorescently labelled buttermilk. Data represents the average of three replicate experiments and error bars represent standard deviation. Results for colostrum buttermilk (A – D1 post-partum), transitional buttermilk (B – D8 post-partum), mature buttermilk (C – D70 post-partum) are given.
Figure 5. Lectin blot profiles of bovine buttermilk generated for animal 1 over the course of lactation for lectins LEL (A) and LTA (B). Samples were separated on 4-12% gradient SDS-PAGE gels and transferred to PVDF membranes using semi-dry transfer. Membranes were blocked overnight in 1% BSA and subsequently washed three times in TBS-T. Blots were then incubated with 10 mL biotinylated lectin (2 μg/mL) for 90 min and washed three times in TBS-T. For the blot incubated with biotinylated LEL, the blot was incubated in 1:20,000 Avidin-D Horseradish peroxidase conjugate and lectin-binding was visualized by chemiluminescence. For the blot incubated with biotinylated LTA the streptavidin alkaline phosphatase conjugate was incubated with the membrane until colour enhancement was adequate for visualization. Colour development was halted by rinsing in distilled H$_2$O.
Table S1. Lectins printed in order, their binding specificities, their simple print sugars (1 mM) and the supplying company.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Source</th>
<th>Species</th>
<th>Common name</th>
<th>General binding specificity*</th>
<th>Print sugar</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIA, Jacalin</td>
<td>Plant</td>
<td><em>Artocarpus integrifolia</em></td>
<td>Jack fruit lectin</td>
<td>Gal, Gal-β-(1,3)-GalNAc (sialylation independent)</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>RPBAlI</td>
<td>Plant</td>
<td><em>Robinia pseudoacacia</em></td>
<td>Black locust lectin</td>
<td>Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>PA-I</td>
<td>Bacteria</td>
<td><em>Pseudomonas aeruginosa</em></td>
<td>Sambucus lectin</td>
<td>Gal, Gal derivatives</td>
<td>Gal</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>SNA-II</td>
<td>Plant</td>
<td><em>Sambucus nigra</em></td>
<td>Sambucus lectin-II</td>
<td>Gal/GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>SJA</td>
<td>Plant</td>
<td><em>Sophora japonica</em></td>
<td>Pagoda tree lectin</td>
<td>β-linked GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>DBA</td>
<td>Plant</td>
<td><em>Dolichos biflorus</em></td>
<td>Horse gram lectin</td>
<td>GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>GHA</td>
<td>Plant</td>
<td><em>Glechoma hederacea</em></td>
<td>Ground ivy lectin</td>
<td>GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>SBA</td>
<td>Plant</td>
<td><em>Glycine max</em></td>
<td>Soy bean lectin</td>
<td>GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>VVA-B4</td>
<td>Plant</td>
<td><em>Vicia villosa</em></td>
<td>Hairy vetch lectin</td>
<td>GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>BPA</td>
<td>Plant</td>
<td><em>Bauhinia purpurea</em></td>
<td>Camels foot tree lectin</td>
<td>GalNAc/Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>WFA</td>
<td>Plant</td>
<td><em>Wisteria floribunda</em></td>
<td>Japanese wisteria lectin</td>
<td>GalNAc/sulfated GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>ACA</td>
<td>Plant</td>
<td><em>Amaranthus caudatus</em></td>
<td>Amaranthin</td>
<td>Sialylated/Gal-β-(1,3)-GalNAc</td>
<td>Lac</td>
<td>Vector Labs</td>
</tr>
<tr>
<td>ABL</td>
<td>Fungus</td>
<td><em>Agaricus bisporus</em></td>
<td>Edible mushroom lectin</td>
<td>Gal-β(1,3)-GalNAc, GlcNAc</td>
<td>Lac</td>
<td>EY Labs</td>
</tr>
<tr>
<td>PNA</td>
<td>Plant</td>
<td><em>Arachis hypogaeae</em></td>
<td>Peanut lectin</td>
<td>Gal-β-(1,3)-GalNAc (sialylation intolerant)</td>
<td>Lac</td>
<td>EY Labs</td>
</tr>
<tr>
<td>GSL-II</td>
<td>Plant</td>
<td><em>Griffonia simplicifolia</em></td>
<td>Griffonia lectin-II</td>
<td>GlcNAc</td>
<td>GlcNAc</td>
<td>EY Labs</td>
</tr>
<tr>
<td>sWGA</td>
<td>Plant</td>
<td><em>Triticum vulgaris</em></td>
<td>Succinyl WGA</td>
<td>GlcNAc</td>
<td>GlcNAc</td>
<td>EY Labs</td>
</tr>
<tr>
<td>Plant</td>
<td>Lectin Name</td>
<td>Sugar Binding Pattern</td>
<td>Supplier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>---------------------------</td>
<td>-------------------------------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSA</td>
<td>Jimson weed lectin</td>
<td>GlcNAc</td>
<td>EY Labs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STA</td>
<td>Potato lectin</td>
<td>GlcNAc oligomers</td>
<td>EY Labs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEL</td>
<td>Tomato lectin</td>
<td>GlcNAc-β-(1,4)-GlcNAc</td>
<td>EY Labs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NPA</td>
<td>Daffodil lectin</td>
<td>α-(1,6)-Man</td>
<td>Man</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNA</td>
<td>Snowdrop lectin</td>
<td>α-(1,3)-Man</td>
<td>Man</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHA</td>
<td>Amaryllis agglutinin</td>
<td>Man-α-(1,3)-Man-α-(1,6)</td>
<td>Man</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Con A</td>
<td>Jack bean lectin</td>
<td>Man, Glc, GlcNAc</td>
<td>Man</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lch-B</td>
<td>Lentil isoelectin B</td>
<td>Man, core fucosylated, agalactosylated biantennary N-linked oligosaccharides</td>
<td>Man</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSA</td>
<td>Pea lectin</td>
<td>Man, core fucosylated trimannosyl N-linked oligosaccharides</td>
<td>Man</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WGA</td>
<td>Wheat germ agglutinin</td>
<td>NeuAc/GlcNAc</td>
<td>EY Labs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAA</td>
<td>Maackia agglutinin</td>
<td>α-(2,3)-Neu5Ac</td>
<td>EY Labs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNA-I</td>
<td>Sambucus lectin-I</td>
<td>α-(2,6)-Neu5Ac</td>
<td>EY Labs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHA-L</td>
<td>Kidney bean leukoagglutinin</td>
<td>Tri- and tetra-antenary β-Gal/Gal-β-(1,4)-GlcNAc</td>
<td>Lac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCA</td>
<td>Scarlet runner bean lectin</td>
<td>GlcNAc in complex oligosaccharides</td>
<td>Sigma Aldrich</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHA-E</td>
<td>Kidney bean erythrogglutinin</td>
<td>Biantennary with bisecting GlcNAc,β-Gal-β-(1,4)-GlcNAc</td>
<td>Lac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCA-I/120</td>
<td>Castor bean lectin I</td>
<td>Gal-β-(1,4)-GlcNAc</td>
<td>Gal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPA</td>
<td>Chickpea lectin</td>
<td>Complex oligosaccharides</td>
<td>EY Labs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAA</td>
<td>Pea tree lectin</td>
<td>Gal-β-(1,4)-GlcNAc</td>
<td>EY Labs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECA</td>
<td>Cocks comb/coral tree lectin</td>
<td>Gal-β-(1,4)-GlcNAc oligomers</td>
<td>Lac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAL</td>
<td>Orange peel fungus lectin</td>
<td>α-(1,6)-Fuc and α-(1,3)-Fuc</td>
<td>Fuc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LTA</td>
<td>Lotus lectin</td>
<td>α-(1,3)-Fuc, α-(1,6)-Fuc and α-(1,2)-Fuc</td>
<td>Fuc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UEA-I</td>
<td>Plant</td>
<td>tetragonolobus</td>
<td>Gorse lectin-I</td>
<td>Fuc</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>---------------</td>
<td>---------------</td>
<td>-----</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>EEA</td>
<td>Plant</td>
<td>Euonymous europaeus</td>
<td>Spindle tree lectin</td>
<td>Terminal α-Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>GSL-I-B4</td>
<td>Plant</td>
<td>Griffonia simplicifolia</td>
<td>Griffonia lectin-I</td>
<td>Terminal α-Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>MPA</td>
<td>Plant</td>
<td>Maclura pomifera</td>
<td>Osage orange lectin</td>
<td>Terminal α-IGal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>VRA</td>
<td>Plant</td>
<td>Vigna radiata</td>
<td>Mung bean lectin</td>
<td>Terminal α-Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>MOA</td>
<td>Fungus</td>
<td>Marasmius oreades</td>
<td>Fairy ring mushroom lectin</td>
<td>Terminal α-Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>Empty</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PBS</td>
<td>-</td>
<td>-</td>
<td>Phosphate buffered saline</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BSA</td>
<td>Animal</td>
<td>-</td>
<td>Bovine serum albumin</td>
<td>-</td>
<td>-</td>
<td>Sigma-Aldrich</td>
</tr>
<tr>
<td>CCA</td>
<td>Animal</td>
<td>Cancer antennarius</td>
<td>California crab lectin</td>
<td>O-acetylated sialic acids</td>
<td>Lac</td>
<td>EY Labs</td>
</tr>
</tbody>
</table>

* Reported recognition based on literature consensus or experimental evidence generated within our laboratory.
Figure S1. Representative image of experimental subarrays
Figure S2. Buttermilk samples from each animal were separated under reducing conditions on a NuPage 4-12% SDS PAGE gels for days 1-90 post-partum, coomassie-stained for total protein (A) and PAS-stained for glycoprotein (B).
Figure S3. Differential gel band intensities for the bands at 45/50 kDa of the coomassie-stained SDS PAGE gel (animal 1) were quantified using ImageJ software. Arbitrary units were used.
Figure S4. Lectin microarray profiles of bovine buttermilk generated for three animals demonstrating all time points over the course of lactation for animal 1 (A), animal 2 (B) and animal 3 (C). Histograms represent the differences in recognition of printed lectins by the buttermilk samples. Data represents the average of three replicate experiments and error bars represent standard deviation.
Chapter III

A

B

C

Fluorescence intensity (RU)
Day
Animal 1
Animal 2
Animal 3
Figure S5. Line graphs representing lectin microarray profiles of bovine buttermilk generated for all animals demonstrating all time points over the course of lactation for lectin WGA (A), PCA (B), LTA (C), AIA (D), RCA-I (E), WFA (F) and MPA (G). Data represents the average of three replicate experiments.
Figure S6. (A) SBA lectin microarray profile of bovine buttermilk generated for all animals over the course of lactation (Days 1-10 and days 30, 70 and 90) in comparison to (B) densitometry of the SBA binding to bovine milk fat globule membrane glycoproteins (Days 1, 2, 4 and 6) as published by Ujita et al., (1993).
Figure S7. Lectin blot profiles of bovine buttermilk generated for animal 1 over the course of lactation for lectins RCA-I (A), WFA (B), MPA (C), AIA (D) and WGA (E). Samples were separated on 4-12% gradient SDS-PAGE gels and transferred to PVDF membranes using semi-dry transfer. Membranes were blocked overnight in 1% BSA and subsequently washed three times in TBS-T. Blots were then incubated with 10 mL biotinylated lectin (2 μg/mL for MPA, LTA, WFA and AIA and 2.5 μg/mL for RCA-I) for 90 min and washed three times in TBS-T. For the blots incubated with biotinylated AIA and WGA, the blot was incubated in 1:20,000 Avidin-D Horseradish peroxidase conjugate and lectin-binding was visualized by chemiluminescence. For the blot incubated with biotinylated RCA-I, MPA and WFA the streptavidin alkaline phosphatase conjugate was incubated with the membrane until colour enhancement was adequate for visualization. Colour development was halted by rinsing in distilled H₂O.
Figure S8. Differential gel band intensities were quantified using ImageJ software for LEL (A) and AIA (B) lectin blots. Arbitrary units were used.
Chapter IV

Lectin and neoglycoconjugate microarray profiling of enterohemorrhagic and enteropathogenic *Escherichia coli* strains indicate serotype-dependent glycan presentation and glycan binding affinity

Sarah A. Ross, Jared Q. Gerlach, Stephen Cunningham, Michelle Kilcoyne, Rita M. Hickey, and Lokesh Joshi

Manuscript in preparation

Acknowledgements: Sarah Ross performed all incubations of bacterial strains on the lectin microarrays and neoglycoconjugate arrays. Lectin and neoglycoconjugate microarrays were synthesized by John Glavin and Michelle Kilcoyne. Bacteria were fluorescently labeled by Stephen Cunningham and data analysis was performed by John Glavin, Jared Gerlach and Sarah Ross.

Statistical analysis was performed by Jared Gerlach.
Chapter IV

Abstract

The surface glycosylation of bacteria plays important roles in pathogenesis, cell-cell interactions and evasion of host immune responses. Profiling of surface glycosylation is important in order to fully understand the mechanisms of colonization and pathogenesis. Similarly, the host’s glycan presentation and the molecules presented by colonizing bacteria which are capable of binding to them have been shown to be vital to pathogenicity and therefore defining the glycan binding preferences of pathogenic species is also paramount. The use of microarrays designed specifically for glycomics applications allows high-throughput analysis of bacterial cell surface glycosylation and glycan binding molecules which may also aid in the prediction or identification of similarly virulent strains. In this study, the cell surface glycosylation and carbohydrate binding profiles of 5 enterohemorrhagic Escherichia coli (EHEC) O157:H7 strains (NCTC 12900, DAF454, DPC 6055, P1432, ATCC 43888) and 2 enteropathogenic E. coli (EPEC) strains (NCTC 8007, NCTC 8623) were characterized. Lectin microarray profiles suggested presentation of a broad range of glycan structures with different apparent abundance or accessibility for each strain screened. Neoglycoconjugate (NGC) microarray profiling demonstrated bacterial binding to an assortment of typically mammalian oligosaccharide structures. Both the lectin and NGC microarray analyses highlighted that the differences in surface glycosylation and glycan binding preferences of these EHEC and EPEC strains were different enough to allow the distinction of at least two groups and also revealed that EHEC strain DAF454 presented carbohydrates more consistent in profile with EPEC strains than the other EHEC strains profiled. While groups of strains displayed similar lectin and NGC microarray profiles, it was possible to distinguish between EHEC and EPEC strains through their specific microarray fingerprints.
Chapter IV

Introduction

Particular *Escherichia coli* strains are known for their ability to cause infection. Enterohaemorrhagic *E. coli* (EHEC), including serotype O157:H7, and enteropathogenic *E. coli* (EPEC) are commonly associated with contaminated food and water sources (Feng 2012, Dobrowsky, Van Deventer et al. 2014) and can cause watery diarrhoea (EPEC strains), bloody diarrhoea or severe systemic disease including haemolytic uremic syndrome (HUS) and haemorrhagic colitis (both resulting from EHEC strains). Long-term renal damage often requiring dialysis or kidney transplant can occur in patients who develop HUS (Mellies, Barron et al. 2007) and the disease is associated with a mortality rate of 3-5% (Feng 2012). Large-scale sporadic outbreaks of pathogenic *E. coli* infection continue to occur worldwide. Recently in the US, 33 people were infected with an outbreak of O157:H7 during October-November 2013 due to contaminated ready to eat salad. Of those infected, 32% were hospitalized while two patients developed HUS (CDC, 2013). Previous severe outbreaks include one caused by an O104:H4 strain in Germany between May and September 2011 which grew to approximately 3,000 reported EHEC cases resulting in 18 deaths (Burger 2012). In 1993, O157:H7-contaminated hamburgers led to 477 confirmed infections in the USA (CDC, 1993) Of the 477 cases, there were 144 hospitalizations (30 of whom developed HUS) and three deaths (CDC, 1993). In addition, a large number of less severe cases go unreported (Michel, Wilson et al. 2000). The high number of EHEC and EPEC infection cases worldwide can be of significant economic expense as well; outbreak-associated costs were estimated to be $405 million in the USA in 2003 alone (Frenzen, Drake et al. 2005).

A clear understanding of both the surface glycosylation and glycan binding habits of highly pathogenic bacteria is advantageous in order to better characterize their modes of pathogenicity and to find new ways to combat infection. Various components of the bacterial cell surface are composed of or incorporate monosaccharide, oligosaccharide and polysaccharide moieties.
Bacterial glycosylation is important for pathogenic processes such as biofilm formation, colonization, and immune evasion (Schmidt, Riley et al. 2003, Itoh, Wang et al. 2005). Gram-negative bacteria, including *E. coli*, present a relatively diverse range of cell surface polysaccharides as components of their lipopolysaccharide (LPS) (Caroff and Karibian 2003) and extracellular capsule (Whitfield and Roberts 1999) while the peptidoglycan layer of bacterial cells also contains oligosaccharides which include β-(1,4)-N-acetylglucosamine (GlcNAc) and β-(1,4)-N-acetylneuraminic acid (sialic acid) constituents (Adibekian, Stallforth et al. 2011). Furthermore, bacterial appendages including flagella and adhesins can also be glycosylated in many cases (Reid, Fulton et al. 2010, Iwashkiw, Vozza et al. 2013). Moreover, bacterial glycosylation profiles are known to include moieties which are not found in mammals such as 3-deoxy-D-manno-octulosonate (Kdo) and β-D-glucose polysaccharides (β-glucans) (Adibekian, Stallforth et al. 2011). Sialic acid can aid in host immune evasion during pathogenesis (Severi, Hood et al. 2007) while biofilm formation is mediated in part by GlcNAc polymers (Itoh, Wang et al. 2005). Glycosylation also plays a role in adhesion and invasion, for instance, glycosylation of *Helicobacter pylori* LPS is an important factor for *in vivo* colonization of mice (Moran, Sturegard et al. 2000).

Emerging evidence suggests bacteria may be able to adapt to changing environments by phase variation (Hallet 2001). Phase variation has been observed for *E. coli* in certain circumstances. For instance variation in *ahpC* gene leads to interconversion of the AhpC protein between a peroxidase and a disulfide reductase under stress conditions (Ritz, Lim et al. 2001). In addition, Ag43 phase variation has been observed which is regulated by DNA methylation. Transcription occurs when the regulatory region of the gene is methylated, however OxyR, an oxidative stress response protein, binds unmethylated regulatory region and prevents transcription of the adhesin (Owen, Meehan et al. 1996). Interestingly, phase variation may also lead to changes in expression of cell surface glycosylation. Lectin binding affinities for *E. coli* RS218 indicated a decrease in glycosylation, including GalNAc, sialic
Chapter IV

acid and Fuc, as a function of growth. Interestingly, this change in glycosylation was shown to be independent of the medium in which bacterial cells were grown (Hsu, Pilobello et al. 2006). The bacterial glycome may also change over the course of infection (Erwin and Smith 2007).

Surprisingly, especially in light of the severity and prevalence of EHEC and EPEC outbreaks, few links between the pathogenicity of strains and specific glycosylation have been established. There are at least 167 different specific polysaccharide O-antigens found in the LPS of E. coli (Hsu, Pilobello et al. 2006). Some of these glycans, including tetrasaccharides such as [-2-D-Rha4NAc-α-(1,3)-L-Fuc-α-(1,4)-D-Glc-β-(1,3)-D-GalNAc-α-1-] (Gao, Liu et al. 2012) and [-4)-β-Glc-(1,3)-α-PerNAc-(1,4)-α-GalNAc-(1,3)-α-Fuc-(1-], are not commonly found in mammals (Nishiuchi, Doe et al. 2000). Such unusual glycans could trigger an immune response in the mammalian host, however modification of these structures by polymorphisms or phase variation of glycosylation enzymes may actually aid in protecting invading bacteria from immune detection (Schmidt, Riley et al. 2003). E. coli cells can also display monosaccharides including α-mannose (Man), α-galactose (Gal) and N-acetylgalactosamine (GalNAc) (Ertl, Wagner et al. 2003). The E. coli adhesins, autoaggregation factor antigen 43 (Ag43), adhesin involved in diffuse adherence (AIDA-I), and TibA, are also glycosylated. Such glycosylated moieties may aid in adhesion of E. coli to host epithelial cells and thus play a role in the initial steps required for colonization and infection of the host (Iwashkiw, Vozza et al. 2013). Interestingly, EHEC and EPEC colonize different parts of the intestine; EHEC colonizes the colon while EPEH infects the small intestine (Mellies, Barron et al. 2007) and this may be due, in part, to their different surface glycosylation patterns (Georgia-Korea, Ghigo et al. 2011).

Similarly, the glycan presentation of the mammalian host and the molecules capable of binding to them presented by colonizing bacteria have been shown to be vital to pathogenicity and defining the glycan binding
preferences of pathogenic species is also of paramount importance. The increasing frequency of bacterial antibiotic resistance makes treating bacterial infections more difficult. Since adherence is an important step to establish infection, compounds blocking adhesion (anti-infective compounds) could be used to reduce dependence upon traditional antibiotic strategies in some situations (Sharon and Ofek 2000). Preventing infection by interfering with the interaction between glycans and their receptors requires knowledge of bacterial glycan binding preferences. Given their different preferred localization during infection, it seems likely that the glycan binding preferences of EHEC and EPEC may differ also, a possibility which has been supported in research of other pathogenic microorganism binding to gastrointestinal mucins (Naughton, Mariño et al. 2013). Some studies have already begun to elucidate the glycan receptors of various *E. coli* strains. For example, the importance of GalNAc-β-(1,4)-Gal (Jagannatha, Sharma et al. 1991) and sialic acid (Angeloni, Ridet et al. 2005) as mediators of EPEC adherence and the dependence upon Gal and Man moieties for *E. coli* O157:H7 (Coutiño-Rodríguez, Hernández-Cruz et al. 2001) adherence to cells has been demonstrated. However, studies comparing the glycan binding preferences of individual EHEC and EPEC strains are limited.

Various analytical approaches including mass spectrometry, nuclear magnetic resonance, high performance liquid chromatography and electrophoresis have previously been used to investigate bacterial surface glycosylation (Reid, Fulton et al. 2010). Over the last decade, lectin microarrays have also proven to be an efficient and high throughput method of bacterial glycome analysis which allows the profiling of whole-cell surface glycosylation and this approach makes it possible to screen a high number of carbohydrate-lectin binding affinities at once (Hsu, Pilobello et al. 2006, Reid, Fulton et al. 2010, Kilcoyne, Twomey et al. 2014). In addition, natural lectins, isolated mostly from plants and fungi, are relatively inexpensive and readily available. Neoglycoconjugates (NGC), glycans conjugated to proteins such as bovine serum albumin (BSA) or streptavidin alkaline phosphatase, can be
utilized for purposes including anti-bacterial and anti-viral vaccines, and recognition vehicles for drug targeting (Wong 1995). However, they are also useful tools for the characterization of carbohydrate ligands of a wide array of compounds. For instance neoglycoconjugates have previously been used to assess the binding of human dendritic cells to carbohydrate ligands (Stambach and Taylor 2003) and interestingly, they have also been used to investigate the binding of norovirus to human milk glycoconjugates and oligosaccharides (Shang, Piskarev et al. 2013). In particular, NGC can provide information on bacterial binding to carbohydrate receptors (Boren, Falk et al. 1993, Scharfman, Degroote et al. 1999, Ramphal and Arora 2001). The use of NGC microarrays is beneficial as they are relatively stable over long periods of time and a large number of NGC can be displayed on a single array platform (Wang 2003).

In this study, lectin microarrays have been employed to investigate the surface glycosylation of 5 EHEC and 2 EPEC strains. Furthermore, a proprietary NGC microarray was used to screen the same strains for binding preferences across a defined panel of natural and synthetic glycoproteins with homogenous or heterogeneous carbohydrates covalently linked to them (Kilcoyne, Gerlach et al. 2012). Taken together, this information is important to better understand the pathogenicity of these bacterial strains and may be useful for assessing the risk of pathogenicity for unidentified or newly-isolated bacteria.
Material and methods

Materials

Nexterion® Slide H microarray slides were purchased from Schott AG (Germany). Pure, unlabelled lectins (Table S1) were acquired from EY Laboratories, Inc. (San Mateo, CA) or Vector Laboratories, Ltd. (Orton Southgate, UK). Pure tetramethylrhodamine-(TRITC-) labelled lectins were from EY Laboratories, Inc. Carboxylic acid succinimidyl ester AlexaFluor® 555 fluorescent label was purchased from Life Technologies (Carlsbad, CA). Syto-82 label was purchased from Invitrogen (Paisley, UK). All other reagents were from Sigma-Aldrich Co. unless otherwise noted and were of the highest grade available.

Bacterial strains and culture conditions

Bacterial cultures were obtained from the Dairy Products Research Centre culture collection at Teagasc Moorepark (strains DPC 6055 and P1432, Moorepark, Ireland), the National Collection of Type Cultures (all NCTC strains; London, UK), and the American Type Culture Collection (all ATCC strains, Rockville, MD, USA). Bacteria included in this study include EHEC serotype O157:H7 strains NCTC 12900, DAF454, DPC 6055, P1432, ATCC 43888, and EPEC serogroup O125:H19 strain NCTC 8623 and serotype O111:H2 strain NCTC 8007. All strains were cultured directly from storage into brain heart infusion (BHI) (Oxoid® Ltd, Basingstoke, Hampshire, England) broth and incubated under aerobic conditions at 37 °C. Strains were stocked in BHI broth containing 50% glycerol (v/v) and stored at -20 °C.

Fluorescent labeling of bacterial strains

Internalized fluorescent labeling of bacteria was carried out essentially as previously described with slight variation (Kilcoyne, Twomey et al. 2014). Bacteria were grown overnight in 5 mL BHI broth at 37 °C with shaking.
Chapter IV

Subsequently, bacteria were centrifuged at 6000 x g for 7.5 min, and re-suspended in 5 mL TBS. This wash step was repeated and the pellet was re-suspended in 2 mL TBS. OD600 nm was adjusted to 0.1 prior to staining of the cells with Syto-82 nucleic acid dye. The dye was added to the bacteria to achieve a final concentration of 15 µM dye and incubated at 37 °C with shaking at 180 rpm for 1 h. After incubation samples were centrifuged at 5000 x g for 5 min and the pellet was re-suspended in Tris buffered saline (TBS). The centrifugation and re-suspension step was repeated 3 times. The final pellet was re-suspended in 500 µL of TBS modified with Tween 20 (0.05%) (TBS-T) in an amber microcentrifuge tube.

Construction of lectin and NGC microarrays

Lectin microarrays consisting of a panel of 48 lectins with reported carbohydrate binding specificities (Table S1) was constructed as previously described (Gerlach, Kilcoyne et al. 2014, Kilcoyne, Twomey et al. 2014). Each feature was printed in replicates of six and eight identical replicate subarrays were printed per slide. Lectin performance after printing was tested by incubation with fluorescently-labelled glycoprotein standards (Gerlach, Kilcoyne et al. 2014). The printed microarrays were stored at 4 °C with desiccant until use.

NGC microarrays were prepared essentially as described previously (Kilcoyne, Gerlach et al. 2012, Wang, Cummings et al. 2014). NGCs and natural glycoproteins (see Table S2) were prepared at a concentration of 1 mg/mL in PBS, pH 7.4 based on their protein content measured by the BCA assay and were printed, approximately 1 nL per feature, on Nexterion® Slide H glass microarray substrates in a humid environment (62% +/- 2%) using a piezoelectric microarray printer equipped with an uncoated glass nozzle with a 90 µm orifice (Scienion). Six replicate subarrays were printed per slide with each glycoprotein or neoglycoconjugate (probe) spotted six times. Slides were incubated in a humidity chamber overnight after printing to ensure complete
conjugation. The functional groups on the Nexterion® slide H slides were then deactivated by incubation with 100 mM ethanolamine in 50 mM sodium borate, pH 8, for 1 h at room temperature. The slides were washed with PBS, pH 7.4 with 0.05% Tween-20 (PBS-T) three times and once with PBS. Slides were then centrifuged dry (1,500 rpm, 5 min) and stored dry at 4 °C with desiccant until use.

**Bacterial profiling with glycomic microarrays**

Microarray slides (lectin and NGC) were incubated in the dark using an eight-well gasket slide and incubation cassette system (Agilent Technologies). Initially, bacterial samples were titrated for optimal concentration on the microarrays (ranging from 30 to 70 µL/mL). 70 µL of each fluorescently labelled bacterial sample at 40 µg/mL (NCTC 8007, NCTC 8623), 60 µg/mL (NCTC 12900, DAF454) and 70 µg/mL (DPC 6055, P1432, ATCC 43888) for lectin microarrays or 30 µg/mL (NCTC 8007) and 40 µg/mL (NCTC 12900, DAF454, DPC 6055, P1432, ATCC 43888, NCTC 8623) for NGC microarrays in TBS (20 mM Tris-HCl, 100 mM NaCl, 1 mM CaCl2, 1 mM MgCl2, pH 7.2) with 0.05% Tween-20 (TBS-T) was applied to each well of the gasket. Three replicate slides were incubated per experiment with appropriate haptenic carbohydrates. During the lectin microarray experiments, bovine asialofetuin labelled with AF555 was included in one subarray per experiment (0.5 µg/mL) to verify the performance of the printed lectin microarrays. Similarly, during the NGC microarray experiments, TRITC-conjugated *Maackia amurensis* agglutinin (MAA-TRITC, 2 µg/mL) was included in one subarray per slide. The microarray slide was sandwiched with the gasket, the cassette assembled and placed in a rotating incubation oven (23 °C, approximately 4 rpm) for 1 h. Incubation chambers were disassembled under TBS-T, washed three times in TBS-T for 2 min each in a Coplin jar and once with TBS. Microarrays were dried by centrifugation (473 x g, 5 min) and scanned immediately using an
Chapter IV

Agilent G2505B microarray scanner using the Cy3 channel (532 nm excitation, 90% PMT, 5 µm resolution).

**Data extraction and processing**

Data extraction was performed essentially as previously described (Gerlach, Kilcoyne et al. 2014). In brief, raw intensity values were extracted from the image tag information files (TIF) using GenePix Pro v6.1.0.4 (Molecular Devices, Berkshire, UK) and a proprietary file containing feature address and identity (*.gal) using adaptive diameter (70-130%) circular alignment based on 230 µm features and were exported as text to Excel (version 2007, Microsoft). Local background-corrected median feature intensity data (F532median-B532) was selected for analysis. The median of six replicate spots per subarray was handled as a single data point for graphical and statistical analysis. For graphical representation, data was normalized to the mean total intensity value of three replicate microarray slides and binding data was presented in histogram form of average intensity with one standard deviation (SD) of microarray experimental replicates (n=3; 18 total data points per probe).

**Statistical analysis**

Prior to all comparative and multivariate analyses, lectin and NGC microarray data for each slide replicate was normalized based on a correction factor generated by division of each subarray total intensity value by the median total intensity value for all subarrays compared. The significance of inhibition at each lectin was evaluated using a standard Student’s T-test (paired, two-tailed). Normalized binding intensity data and binding intensity data under interfering conditions from a single slide were paired, and the p value determined based on three pairings.

Unsupervised clustering of normalized lectin binding data was performed with Hierarchical Clustering Explorer v3.5 (Human-Computer
Interaction Lab, University of Maryland, http://www.cs.umd.edu/hcil/hce/hce3.html). Median normalized data was initially clustered with the following parameters: no pre-filtering, complete linkage and Euclidean distance. To evaluate contribution of individual printed features (lectins or NGCs) to the clustering outcomes and establish the minimum number of unique features required to distinguish serotypes, increasingly stringent pre-filtering was applied to previously normalized data (50 relative fluorescence units (RFU) SD steps to previously normalized data sets prior to clustering until a threshold of unambiguous association of serotypes was crossed). A final cut-off value for data filtering was then selected by 20 RFU stepwise reduction of the stringency until glycoform groups were restored. Principal component analysis (PrCA) was performed on select, normalized glycomic microarray data with Minitab 16 software (version 16.2.2, Minitab, Inc., State College, PA). Score plots were generated from the first and second principal component calculations.
Chapter IV

Results

Lectin microarray profiling of EHEC and EPEC strains

All bacterial cell populations were grown with matched inoculum, conditions, volume and duration of growth and used at an identical OD600. Following the method previously reported (Kilcoyne, Twomey et al. 2014), the optimum concentration of Syto-82 dye required for efficient bacterial staining was determined to be 15 µM for all bacterial strains screened in this study (Figure 1).

The glycosylation of Syto-82 labeled bacterial strains was profiled on lectin microarrays consisting of a panel of 48 lectins (Table S1). Lectin microarray data was examined for similarities between bacterial strains. Unsupervised hierarchical clustering of the normalized data generated broad clades with a subtle range of biological and technical variation that was within expectation (Figure 3A). Individual replicates 3A for EHEC strains ATCC 43888 and DPC 6055 demonstrated mild outlier behavior while replicate 3B for EPEC strain NCTC 8007 demonstrated very strong outlier behavior relative to the rest of the data (Figure 3A). Employing a previously reported pre-filtering approach (Gerlach, Kilcoyne et al. 2014), the data set was reduced to include only the most substantial contributors to the grouping of the lectin microarray profiles. Ultimately, the application of a 1000 RFU deviation cutoff resulted in the retention of eight lectins (PA-I, HPA, ABL, sWGA, WGA, VRA, MPA and MOA) which, when used with the same hierarchical clustering algorithm, maintained the split of 4/5 EHEC strains into one group and 1/5 EHEC + 2/2 EPEC strains in the other group but reduced the minimum similarity to just 4% (Figure 3B). The most obvious differences between the two groups were the nearly binary responses at lectins WGA and MOA at which EHEC strains were preferentially bound by MOA whereas EPEC strains were preferentially bound by WGA (Figure 3B). Principal component analysis performed utilizing data from only the pre-filtered subset of lectins (PA-I, HPA, ABL, sWGA, WGA,
Chapter IV

VRA, MPA and MOA) identified three distinct groups of bacteria (Figure 3C). 4 of the 5 EHEC strains screened in this study (NCTC 12900, DPC 6055, ATCC 43888, and P1432) fell into one group while the two EPEC strains (NCTC 8007 and NCTC 8623) fell into another group. Another distinct cluster of replicates contained only the DAF454 EHEC strain (Figure 3C). Examination of the 8-lectin subset for evidence contributing to the isolation of the single EHEC strain showed that the lectin profile of DAF454 was characterized by the strongest relative binding at WGA with a notable absence of the characteristic EHEC response at MOA (Figure 3C).

In general, lectin microarray profiles of EHEC and EPEC strains displayed relatively high responses at only a handful of lectins, most notably AAL, HPA, SNA-II, PA-I, ACA, ABL, WGA, sWGA, MOA, MPA, and VRA; the remaining lectin interactions were substantially lower in relative magnitude. EHEC strains NCTC 12900, DPC 6055, ATCC 43888, and P1432 displayed strongest interactions with MOA, which suggested the presence of terminal α-Gal, and ABL, which suggested the presentation of Gal-β-(1,3)-GalNAc or GlcNAc (Figure 2A). In contrast, EPEC strains NCTC 8623 and NCTC 8007 (Figure 2B) and EHEC strain DAF454 (Figure 2C) displayed only relatively low binding intensities at MOA and no interaction with ABL. In addition, the presence of Gal on all bacterial cell surfaces was suggested by responses at the lectins VRA and MPA, which typically bind α-Gal, SNA II, which has affinity for Gal and GalNAc, and PA-I which binds Gal (Table S1), albeit with lower intensities than with MOA. DAF454 also displayed binding to HPA (Figure 3C) while NCTC 8623 and NCTC 8007 demonstrated relatively low binding to HPA, suggesting the presence of α-GalNAc (Table S1). A similar result was previously documented when binding of K12-derived E. coli strains to HPA lectin was observed (Hsu, Pilobello et al. 2006). The human macrophage C-type lectin has been identified as having affinity for Gal and GalNAc residues and plays a role in filoviral attachment and subsequent infection of human chronic myelogenous leukemia cells (Takada, Fujioka et al. 2004). It is possible that these residues also increase the adhesion of E. coli to host cells.
Incorporation of GalNAc into *E. coli* glycoproteins has previously been shown (Xu, Hanson et al. 2004) and the presence of Gal and GalNAc in LPS O-antigens has been identified previously for EPEC O111:B4 and O55:B5 serotypes (Morales Betanzos, Gonzalez-Moa et al. 2009).

DAF454 displayed highest relative binding to the lectin sWGA (Figure 2C) compared to all other strains (Figure 2), which indicated either a higher proportion of GlcNAc residues or better accessibility to GlcNAc on its cell surface (Table S1). DAF454, and EPEC strains NCTC 8623 and NCTC 8007, displayed the strongest interaction with WGA (Figure 2B and 2C) suggesting the presence of GlcNAc or sialic acid (Table S1). In contrast, NCTC 12900, DPC 6055, ATCC 43888, and P1432 displayed lower binding intensities to WGA (Figure 2A). In particular, P1432 and ATCC 43888 displayed very low relative binding intensities to WGA lectin compared to all other strains. This suggested that DAF454, NCTC 8623 and NCTC 8007 may have a greater proportion of cell surface GlcNAc or sialic acid when compared to the other four EHEC strains, or better accessibility to GlcNAc or sialic acid on their cell surfaces. Fucosylation on the surface of the bacterial strains was also evident as observed by the binding to AAL (Figure 2), a lectin with reported preference for α-(1,6)- and α-(1,3)-fucose (Fuc) (Table S1). Polysaccharide intercellular adhesin (PIA) is an extracellular GlcNAc polysaccharide produced by bacterial strains including *E. coli* (Wang, Preston et al. 2004) which plays an important role in biofilm formation (Itoh, Wang et al. 2005) and evasion of the host innate immune system (Vuong, Voyich et al. 2004). Terminal GlcNAc residues have been identified by lectin microarray analysis previously for *E. coli*, as demonstrated by binding to the lectin GS II (Hsu, Pilobello et al. 2006) and GlcNAc has also been identified for *E. coli* LPS O-antigen by random sequence peptide microarrays (Morales Betanzos, Gonzalez-Moa et al. 2009).

It may be possible to distinguish between closely related EPEC and EHEC strains on the basis of their lectin fingerprints. Similarly, Hsu et al. (2006) demonstrated the ability to differentiate between closely related strains
Chapter IV

through the use of lectin microarray profiling and the researchers also showed that the technique was more sensitive when compared to agglutination assays. The binding intensity of MOA for NCTC 12900 was much higher than that observed for DPC 6055 while the intensities for ACA and PA-I were lower for NCTC 12900 when compared to DPC 6055 (Figure 2A). Furthermore, ATCC 43888 and P1432 could be distinguished from each other as the intensity of binding to MOA was greater for P1432 and binding to ABL was greater for ATCC 43888. The two EPEC strains screened in this study displayed very similar binding profiles (Figure 2B). However, there may be potential to differentiate between these strains also. For instance, the relative binding of NCTC 8623 to WGA was greater when compared to NCTC 8007.

NGC microarray profiling of EHEC and EPEC strains

The glycan binding preferences of Syto-82 labeled bacterial strains were profiled on NGC microarrays consisting of a panel of 52 NGCs (Table S2). In general, NGC microarray profiles of EHEC and EPEC strains displayed binding intensities which were relatively low compared to the lectin microarrays (Figure 4). Additionally, responses from a higher number of NGC interactions were evident compared to the lectin microarrays indicating the acceptance of a wide variety of carbohydrates by the receptors for each of the strains (Figure 4).

Multivariate data analysis approaches were used to examine the relationships between NGC microarray profiles produced by each strain of labelled bacteria. Unsupervised hierarchical clustering of the normalized microarray data generated the tightest clustering behavior for NCTC 12900 and NCTC 8623 replicates while the profiles for the remaining strains were more broadly distributed (Figure 5A). Mean data, derived from all of the like replicates, was subsequently reduced to reinforce the strongest NGC contributors through the application of pre-filtering (Gerlach, Kilcoyne et al. 2014). Pre-filtering was applied to the entire set of NGC responses based on standard deviation (SD) across all strains. The selected SD cut-off point (100
RFU) resulted in the retention of only 22 NGCs which represented the strongest contributors to the clustering outputs (Figure 5B). Hierarchical clustering of this reduced data set resulted in a split (at 22% minimum similarity) of the bacterial strains into two groups with all EPEC strains falling into one group and all EHEC strains falling into another (Figure 5B). Principal component analysis was performed utilizing the reduced data set and this maintained the split of the EHEC strains in one group and the EPEC strains in another (Figure 5C). The individual replicate 2B for NCTC 12900 demonstrated strong outlier behavior relative to the rest of the data for this strain (Figure 5C).

All strains demonstrated a preference for $\alpha$-Man-binding as seen by the binding to M3BSA and Man$\alpha$BSA (Table S2 and Figure 4). Binding to M3BSA was slightly higher for EHEC (Figure 3A) than for EPEC (Figure 3B) strains. A binding preference for Gal in $\beta$-(1,4)-linkage was identified for the EHEC and EPEC strains (Figure 4) as evidenced by binding to NGCs including LNnTHSA, LexBSA, DiLexBSA and LNFPIIIBSA (Table S2). Interestingly, EPEC strains bound with slightly higher binding intensity to LNnTHSA than EHEC strains (Figure 4). Additionally, binding to $\beta$-(1,3)-Gal was identified (Figure 4) through binding to GM1HSA and LeaBSA (Table S2). Some EHEC strains such as NCTC 12900, DPC 6055 and DAF454 (Figure 4A) bound with higher affinity to GM1HSA than the EPEC strains (Figure 4B). Thus it can be seen that Gal moieties are likely important for *E. coli* pathogenesis. A preference for GlcNAc was also observed by bacterial strains binding to GlcNAcBSA (Table S2) which EHEC strains bound with higher binding intensity (Figure 4A) than EPEC strains (Figure 4B). Binding to mannose has previously been shown for *E. coli* strains using carbohydrate microarrays (Disney and Seeberger 2004). Man derivatives also inhibit adhesion of uropathogenic *E. coli* to mammalian intestinal epithelial cells (Firon, Ashkenazi et al. 1987) while Man and Gal are known to be important mediators in *E. coli* O157:H7 adherence to red blood cells (Coutiño-Rodríguez, Hernández-Cruz et al. 2001).
Overall a high binding preference for sialic acid was evident amongst all the profiled bacteria as evidenced by binding to NGCs including 3SLacHSA and 6SLacHSA (Table S2). Indeed, some of the highest relative binding intensities seen for both EHEC and EPEC strains occurred at 3SLacHSA and 6SLacHSA (Figure 4). The ability to bind to sialic acid was further evidenced by the binding seen for MMLNnHHSA (Table S2) for all strains (Figure 4). Thus it is evident the E. coli strains could bind to Neu5Ac in both -α-(2,3) and –α-(2,6) linkages. Binding was also evident for fibrin (Figure 4). Fibrin is a sialylated glycoprotein with Neu5Ac, GlcNAc, Gal and Man moieties (Pabst, Bondili et al. 2007). This strengthens the idea that these glycan moieties are important for adherence of the E. coli strains screened. Interestingly, the EPEC strains (Figure 4B) bound with less relative intensity to fibrin than the EHEC strains (Figure 4A), indicating less preference for this particular glycoprotein. Sialic acid has previously been identified as an erythrocyte receptor for E. coli binding (Parkkinen, Rogers et al. 1986). As it can be seen that sialic acid is an important factor in pathogenic binding the presence of sialic acid in anti-adhesive compounds would be an invaluable glycan moiety in the prevention of bacterial adhesion to host cell surfaces. Indeed, E. coli adherence to human intestinal epithelial cells was reduced in the presence of sialic acid (Parker, Sando et al. 2010) indicating the possibility of using food-derived sialic acid as anti-infective compounds. Bacterial preference for fucosylation was also evident. For example, binding was seen for H2BSA and LebBSA (Table S2) for all strains (Figure 4). This was not surprising as the importance of fucosylation in host cell glycosylation has previously been demonstrated for E. coli strains including EHEC (Ashkenazi, May et al. 1991). Fuc has previously been demonstrated to reduce the adherence of E. coli to Caco-2 cells (Parker, Sando et al. 2010) further highlighting adherence of E. coli to this glycan moiety.
Discussion

Profiling bacterial glycosylation using lectin microarrays identified only the accessible glycans on the bacterial cell surface. As these are the moieties which would interact with the host environment, the lectin microarray profiles presented in this study provide a good indication of the biologically relevant carbohydrates presented by these strains. The different lectin microarray profiles observed for these strains suggested *E. coli* serotypes display specific cell surface glycosylation structural characteristics which may contribute to tissue tropism and may partly explain why EHEC and EPEC colonize different parts of the intestine (Mellies, Barron et al. 2007).

It was interesting to note binding of ACA to DAF454, NCTC 8007 and NCTC 8623 in the lectin microarrays but that there was no binding to other lectins which are known to bind to Gal in a β-(1,3)-linked configuration such as AIA, PNA and ABL (Figure 2B and 2C). The presence of sialic acid on the bacterial surfaces may play an important role in bacterial ability to evade immune function. Sialic acid is a ligand for Factor H, which reduces activation of the complement pathway (Severi, Hood et al. 2007). Furthermore, it can bind with Siglecs expressed on the surface of immune cells, leading to inhibition of immune cell function. Thus the incorporation of sialic acid in bacterial glycomes may play a role in bacterial ability to evade eradication by the host immune system (Severi, Hood et al. 2007). The presence of α-2,3- and α-2,6-sialic acid on bacterial cell surfaces has been shown previously for K12-derived *E. coli* strains JM101 and HB101 and for *E. coli* RS218 (Hsu, Pilobello et al. 2006). Additionally, α-(2,3)-sialic acid moieties were identified as terminal epitopes of O-linked glycoconjugates for *E. coli* O24 and O26 (Gamian, Kenne et al. 1994).

The presence of α-(1,2)-Fuc has previously been identified in another *E. coli* strain, RS218 (Hsu, Pilobello et al. 2006) and the *wca* operon on *E. coli* K-12 contains genes for fucose production (Joseph, Desai et al. 2012). Fuc residues are also an important factor for intestinal colonization. For instance,
surface fucosylation was shown to be an important factor for *Bacteroides fragilis* colonization in mouse models where it was demonstrated to out-compete mutant strains which were unable to incorporate fucose in their cell surface components (Coyne, Reinap et al. 2005). O-antigen side chain fucosylation of *H. pylori* is also an important factor in adhesion to human gastric cells (Edwards, Monteiro et al. 2000). Thus, fucosylation could also give an advantage to the colonization and subsequent infection for the *E. coli* strains screened here.

The inference of bacterial glycan structure from the lectin profiles must be subject to cautious interpretation as most of the lectins on the microarray used here have been most extensively characterized for their binding preferences among structures typically produced by mammalian, if not eukaryotic, cells. The presentation of bacterial oligo- and polysaccharides with monosaccharide components and respective linkages which are distinctly prokaryotic provides the opportunity for interpretation of binding events to be erroneous. For example, while it may be inferred that the strong binding of MOA to the EHEC strains could be directly mediated by exposed terminal α-linked Gal, this may or may not be the case. Additionally, we cannot exclude the possibility that the bacterial adhesins may also be playing some role in determining the particular profiles observed. Many of the lectins on the lectin microarray are glycosylated proteins, mostly displaying motifs typical of plant N-glycosylation. For example, both EHEC and EPEC strains displayed affinity for glycoproteins on the NGC array which have N-linked modifications, including α-crystallin, Fibrin, ASF, Ceruloplasmin and, in the case of EPEC NCTC 8007, RNAse B. Furthermore, NGCs including LNTHSA, LNnTHSA, GlcNAcTHSA, ManaBSA present structures which may be displayed by lectins on the microarray. However, since previous studies have employed lectin microarrays to investigate bacterial glycosylation (Hsu and Mahal 2006, Hsu, Pilobello et al. 2006), it is clear they can successfully be used to identify bacterial species through the generation of unique profiles for individual strains.
Chapter IV

Lectin microarrays could be used as a tool to identify the glycosylation profiles of host cells, particularly from common sites of colonization of particular bacterial strains. This may provide a better understanding of the adhesion process, particularly when compared to the bacterial glycan binding profiles evidenced by the NGC microarrays presented here. Furthermore, lectin microarrays could be used to profile the glycan presentation of anti-adhesive compounds in the research of potential decoy receptor therapies. Decoy receptors, which mimic host cell receptors, provide an alternative binding site for bacteria, thereby preventing bacterial adhesion and subsequent infection in the host (Sharon and Ofek 2000). Knowledge of the glycan presentation of potential decoy receptors would aid the development of anti-infective ingredients to target specific bacterial strains. Indeed, the milk fat globule membrane is a highly glycosylated component of mammalian milk which has been highlighted as a source of anti-adhesive components (Ross, Lane et al. 2015, Ross, Lane et al. 2016). Interestingly, the glycosylation of bovine buttermilk, a source of milk fat globule membrane, has previously been profiled through the use of lectin microarrays (Ross, Gerlach et al. 2016). The comparison of lectin microarray and NGC microarray profiles could aid in advancements in the development of prophylactics targeting common sources of infection. For instance, the presentation of glycans including Man, Fuc and sialic acid in bovine buttermilk (Ross, Gerlach et al. 2016) could target the EHEC and EPEC strains screened in this study based on their NGC microarray profiles.

The glycan receptors of pathogenic and commensal species are of great interest in the quest to understand and control bacterial colonization. Although the differences in binding affinities between EHEC and EPEC strains observed on the NGC arrays was somewhat subtle, the small differences that were detected between serotypes further add to the idea of preferential attachment at specific carbohydrates. Multivariate analysis performed using only data from 22 of the structures on the NGC array was sufficient to divide the data into two distinct groups which highlighted the suggestion that different bacterial strains
use differing bacterial carbohydrate receptors in the selective colonization of the gut and other tissues. This may contribute to the different tissue tropisms of EHEC and EPEC (Mellies, Barron et al. 2007). It has previously been demonstrated that mucins isolated from the gastrointestinal tract of various mammalian species differ in their glycosylation patterns as the mucins progressed through the digestive system (Kilcoyne, Gerlach et al. 2012). For instance, the authors demonstrated greater accessibility to α-(1,2)-Fuc, α-(2,3)-Neu5Ac, complex biantennary N-linked structures and terminal GalNAc epitopes in mucins isolated from the chicken large intestine compared to the small intestine. Given that the cells of the human gastrointestinal tract also differ between the small and large intestines (Mowat and Agace 2014), it is likely that similar regional glycosylation differences in the human gastrointestinal tract would contribute to the tissue tropisms evidenced for EHEC and EPEC strains. It has been demonstrated that the expression of glycans in human cell lines can be profiled through the use of lectin microarrays (Kilcoyne, Gerlach et al. 2012) and thus lectin microarray profiling of cells isolated from various regions of the gastrointestinal tract would be advantageous in order to characterize the potential glycan receptors of these bacterial strains.

The NGC microarray data is of further interest given the broad array of glycan structures the NGC microarray profiles indicated the bacteria could bind. This may indicate that a mixture of glycan structures in anti-adhesive compounds may offer further therapeutic benefit than the inclusion of a single glycan structure in order to maximize the number of decoy receptors the bacteria can bind to reduce the threat of infection from occurring.
Bacterial glycosylation enhances the bacterial interaction with the environment. We have profiled the surface glycosylation of several strains of EHEC and EPEC, through the use of lectin microarrays. The presence of glycans such as α-Gal, GalNAc, sialylated Gal-β-(1,3)-GalNAc and fucosylation in α-(1,6) linkage were suggested as potential cell surface components for all E. coli strains screened. Additionally, DAF454, NCTC 8623 and NCTC 8007 were potentially decorated with either higher proportions of GlcNAc or sialic acid, or had better accessibility to these moieties on their surfaces, than the other strains screened. Furthermore, the carbohydrate binding preferences of the bacterial strains was identified through the use of NGC microarrays. A high preference for sialic acid was identified with bacterial binding to Fuc, Man, Gal and GlcNAc residues also observed. The information presented in this study could be invaluable in order to aid in the identification of strains isolated from sources of infection. It may also contribute to our understanding of bacterial-host interactions and comparison of the lectin microarray profiles of host cells and potential decoy receptors with NGC microarray profiles of bacterial strains, may aid in identifying anti-infective compounds which target and interfere with glycans involved in the colonization and adhesion process.
Chapter IV

References


pathogenic *Escherichia coli* strains isolated from domestically harvested rainwater during low-and high-rainfall periods." Applied and Environmental Microbiology *80*(5): 1633-1638.


Chapter IV


Chapter IV


Chapter IV


Chapter IV


Figure 1. Optimization of *E. coli* NCTC 12900-, ATCC 43888-, DPC 6055-, DPC 6054-, DAF454-, NCTC 8007- and NCTC 8623-staining with Syto-82 at different concentrations of fluorescent dye. To determine the optimum Syto-82 concentration for each strain screened, different concentrations of dye were added to the washed bacterial suspension to give a final range of 5-60 μM. After incubation, fluorescence of bacterial strains was measured and the optimal concentration was based on maximum fluorescent intensity.
Chapter IV
Figure 2. Lectin microarray profile of *E. coli* strains NCTC 12900, ATCC 43888, DPC 6055 and DPC 6054 (A), NCTC 8007 and NCTC 8623 (B), and *E. coli* DAF454 (C). Each histogram represents the differences in recognition of lectins by Syto-82-labelled bacterial cells. Data represent the average of three replicate experiments and error bars represent standard deviation.
Figure 3. Unsupervised hierarchical clustering and principal component analysis results for lectin microarray profiles generated for labelled *E. coli* strains. (A) Median normalized data subjected to no pre-filtering, complete linkage and Euclidean distance. (B) Increasingly stringent pre-filtering was
applied to previously normalized data and a final cut-off value for data filtering was then selected by 20 RFU stepwise reduction of the stringency. (C) Principal component analysis (PCA) utilizing data from only the pre-filtered subset of lectins. Score plots were generated from the first and second principal component calculations.
Figure 4. Neoglycoconjugate microarray profiles of EHEC and EPEC strains. NGC microarray profiles were generated for \textit{E. coli} strains NCTC 12900, ATCC 43888, DPC 6055 and DPC 6054 and \textit{E. coli} DAF454 (A) and NCTC 8007 and NCTC 8623 (B). The histograms represent the differences in recognition of neoglycoconjugates by Syto-82-labelled bacterial cells. Data represent the average of three replicate experiments and error bars represent standard deviation.
Figure 5. Unsupervised hierarchical clustering and principal component analysis results for neoglycoconjugate microarray profiles generated for labelled *E. coli* strains. (A) Median normalized data subjected to no pre-filtering, complete linkage and Euclidean distance. (B) Increasingly stringent
Chapter IV

pre-filtering was applied to previously normalized data and a final cut-off value for data filtering was then selected by 20 RFU stepwise reduction of the stringency. (C) Principal component analysis (PrCA) of select, normalized glycomic microarray data. Score plots were generated from the first and second principal component calculations.
Chapter IV
Supplementary figures and tables

Table S1. Lectins printed in order, their binding specificities, their simple print sugars (1 mM) and the supplying company.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Source</th>
<th>Species</th>
<th>Common name</th>
<th>General binding specificity*</th>
<th>Print sugar</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIA, Jacalin</td>
<td>Plant</td>
<td><em>Artocarpus integrifolia</em></td>
<td>Jack fruit lectin</td>
<td>Gal, Gal-β-(1,3)-GalNAc (sialylation independent)</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>RPBAl</td>
<td>Plant</td>
<td><em>Robinia pseudoacacia</em></td>
<td>Black locust lectin</td>
<td>Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>PA-I</td>
<td>Bacteria</td>
<td><em>Pseudomonas aeruginosa</em></td>
<td>Pseudomonas lectin</td>
<td>Gal, Gal derivatives</td>
<td>Gal</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>SNA-II</td>
<td>Plant</td>
<td><em>Sambucus nigra</em></td>
<td>Sambucus lectin-II</td>
<td>Gal/GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>SJA</td>
<td>Plant</td>
<td><em>Sophora japonica</em></td>
<td>Pagoda tree lectin</td>
<td>β-linked GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>DBA</td>
<td>Plant</td>
<td><em>Dolichos biflorus</em></td>
<td>Horse gram lectin</td>
<td>GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>GHA</td>
<td>Plant</td>
<td><em>Glechoma hederacea</em></td>
<td>Ground ivy lectin</td>
<td>GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>SBA</td>
<td>Plant</td>
<td><em>Glycine max</em></td>
<td>Soy bean lectin</td>
<td>GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>VVA-B4</td>
<td>Plant</td>
<td><em>Vicia villosa</em></td>
<td>Hairy vetch lectin</td>
<td>GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>BPA</td>
<td>Plant</td>
<td><em>Bauhinia purpurea</em></td>
<td>Camels foot tree lectin</td>
<td>GalNAc/Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>WFA</td>
<td>Plant</td>
<td><em>Wisteria floribunda</em></td>
<td>Japanese wisteria lectin</td>
<td>GalNAc/sulfated GalNAc</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>ACA</td>
<td>Plant</td>
<td><em>Amaranthus caudatus</em></td>
<td>Amaranthin</td>
<td>Sialylated/Gal-β-(1,3)-GalNAc</td>
<td>Lac</td>
<td>Vector Labs</td>
</tr>
<tr>
<td>ABL</td>
<td>Fungus</td>
<td><em>Agaricus bisporus</em></td>
<td>Edible mushroom lectin</td>
<td>Gal-β-(1,3)-GalNAc, GlcNAc</td>
<td>Lac</td>
<td>EY Labs</td>
</tr>
<tr>
<td>PNA</td>
<td>Plant</td>
<td><em>Arachis hypogaea</em></td>
<td>Peanut lectin</td>
<td>Gal-β-(1,3)-GalNAc (sialylation intolerant)</td>
<td>Lac</td>
<td>EY Labs</td>
</tr>
<tr>
<td>GSL-II</td>
<td>Plant</td>
<td><em>Griffonia simplicifolia</em></td>
<td>Griffonia lectin-II</td>
<td>GlcNAc</td>
<td>GlcNAc</td>
<td>EY Labs</td>
</tr>
<tr>
<td>sWGA</td>
<td>Plant</td>
<td><em>Triticum vulgaris</em></td>
<td>Succinyl WGA</td>
<td>GlcNAc</td>
<td>GlcNAc</td>
<td>EY Labs</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Plant</td>
<td>Nomenclature</td>
<td>Lectin Type</td>
<td>Heterogeneity</td>
<td>Source</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------</td>
<td>--------------</td>
<td>-------------</td>
<td>---------------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>DSA</td>
<td>Plant</td>
<td><em>Datura stramonium</em></td>
<td>Jimson weed lectin</td>
<td>GlcNAc</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>STA</td>
<td>Plant</td>
<td><em>Solanum tuberosum</em></td>
<td>Potato lectin</td>
<td>GlcNAc</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>LEL</td>
<td>Plant</td>
<td><em>Lycopersicum eyculentum</em></td>
<td>Tomato lectin</td>
<td>GlcNAc-β-(1,4)-GlcNAc</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>NPA</td>
<td>Plant</td>
<td><em>Narcissus pseudonarcissus</em></td>
<td>Daffodil lectin</td>
<td>α-(1,6)-Man</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>GNA</td>
<td>Plant</td>
<td><em>Galanthus nivalis</em></td>
<td>Snowdrop lectin</td>
<td>α-(1,3)-Man</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>HHA</td>
<td>Plant</td>
<td><em>Hippeastrum hybrid</em></td>
<td>Amaryllis agglutinin</td>
<td>Man-α-(1,3)-Man-α-(1,6)</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>Con A</td>
<td>Plant</td>
<td><em>Canavalia ensiformis</em></td>
<td>Jack bean lectin</td>
<td>Man, Glc, GlcNAc</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>Lch-B</td>
<td>Plant</td>
<td><em>Lens culinaris</em></td>
<td>Lentil isoelectin B</td>
<td>Man, core fucosylated, galactosylated biantennary N-linked oligosaccharides</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>PSA</td>
<td>Plant</td>
<td><em>Pisum sativum</em></td>
<td>Pea lectin</td>
<td>Man, core fucosylated trimannosyl N-linked oligosaccharides</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>WGA</td>
<td>Plant</td>
<td><em>Triticum vulgaris</em></td>
<td>Wheat germ agglutinin</td>
<td>NeuAc/GlcNAc</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>MAA</td>
<td>Plant</td>
<td><em>Maackia amurensis</em></td>
<td>Maackia agglutinin</td>
<td>Sialic acid-α-(2,3)-Neu5Ac</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>SNA-I</td>
<td>Plant</td>
<td><em>Sambucus nigra</em></td>
<td>Sambucus lectin-I</td>
<td>Sialic acid-α-(2,6)-Neu5Ac</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>PHA-L</td>
<td>Plant</td>
<td><em>Phaseolus vulgaris</em></td>
<td>Kidney bean leukoagglutinin</td>
<td>Tri- and tetra-antennary β-Gal/Gal-β-(1,4)-GlcNAc</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>PCA</td>
<td>Plant</td>
<td><em>Phaseolus coccineus</em></td>
<td>Scarlet runner bean lectin</td>
<td>GlcNAc in complex oligosaccharides</td>
<td>Sigma Aldrich</td>
<td></td>
</tr>
<tr>
<td>PHA-E</td>
<td>Plant</td>
<td><em>Phaseolus vulgaris</em></td>
<td>Kidney bean erythroagglutinin</td>
<td>Biantennary with bisecting GlcNAc,β-Gal/Gal-β-(1,4)-GlcNAc</td>
<td>EY Labs</td>
<td></td>
</tr>
<tr>
<td>RCA-I/120</td>
<td>Plant</td>
<td><em>Ricinus communis</em></td>
<td>Castor bean lectin I</td>
<td>Gal-β-(1,4)-GlcNAc</td>
<td>Vector Labs</td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Type</td>
<td>Plant</td>
<td>Lectin Name</td>
<td>Sugar Structure</td>
<td>Monosaccharide(s)</td>
<td>Source</td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>------------------------</td>
<td>-------------------------------------------------------</td>
<td>-------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>CPA</td>
<td>Plant</td>
<td><em>Cicer arietinum</em> Caragana arborescens</td>
<td>Chickpea lectin Pea tree lectin</td>
<td>Complex oligosaccharides</td>
<td>Lac</td>
<td>EY Labs</td>
</tr>
<tr>
<td>CAA</td>
<td>Plant</td>
<td><em>Erythrina cristagalli</em></td>
<td>Cocks comb/coral tree lectin</td>
<td>Gal-β-(1,4)-GlCNac Gal-β-(1,4)-GlCNac oligomers</td>
<td>Lac</td>
<td>EY Labs</td>
</tr>
<tr>
<td>ECA</td>
<td>Plant</td>
<td><em>Erythrina arborescens</em></td>
<td>Pea tree lectin</td>
<td>Gal-β-(1,4)-GlcNAc oligomers</td>
<td>Lac</td>
<td>EY Labs</td>
</tr>
<tr>
<td>AAL</td>
<td>Fungi</td>
<td><em>Aleuria aurantia</em></td>
<td>Orange peel fungus lectin</td>
<td>α-(1,6)-Fuc and α-(1,3)-Fuc</td>
<td>Fuc</td>
<td>Vector Labs</td>
</tr>
<tr>
<td>LTA</td>
<td>Plant</td>
<td><em>Lotus tetragonolobus</em> Ulex europaeus</td>
<td>Lotus lectin Gorse lectin-I</td>
<td>α-(1,3)-Fuc, α-(1,6)-Fuc and α-(1,2)-Fuc</td>
<td>Fuc</td>
<td>EY Labs</td>
</tr>
<tr>
<td>UEA-I</td>
<td>Plant</td>
<td><em>Ulex europaeus</em></td>
<td>Gorse lectin-I</td>
<td>α-(1,2)-Fuc</td>
<td>Fuc</td>
<td>EY Labs</td>
</tr>
<tr>
<td>GSL-I-B4</td>
<td>Plant</td>
<td><em>Griffonia simplicifolia</em> Maclura pomifera</td>
<td>Griffonia lectin-I</td>
<td>Terminal α-Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>MPA</td>
<td>Plant</td>
<td><em>Maclura pomifera</em></td>
<td>Osage orange lectin</td>
<td>Terminal α-Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>VRA</td>
<td>Plant</td>
<td><em>Vigna radiata</em></td>
<td>Mung bean lectin</td>
<td>Terminal α-Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>MOA</td>
<td>Fungus</td>
<td><em>Marasmius oreades</em></td>
<td>Fairy ring mushroom lectin</td>
<td>Terminal α-Gal</td>
<td>Gal</td>
<td>EY Labs</td>
</tr>
<tr>
<td>Empty</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PBS</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BSA</td>
<td>Animal</td>
<td>-</td>
<td>Bovine serum albumin</td>
<td>-</td>
<td>-</td>
<td>Sigma-Aldrich</td>
</tr>
</tbody>
</table>

* Reported recognition based on literature consensus or experimental evidence generated within our laboratory.
### Table S2. Neoglycoconjugates printed in order and their substitution ratio

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
<th>Substitution Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fetuin</td>
<td>Fetuin</td>
<td>Natural</td>
</tr>
<tr>
<td>ASF</td>
<td>Asialofetuin</td>
<td>Natural</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline, pH 7.5</td>
<td>N/A</td>
</tr>
<tr>
<td>Ov</td>
<td>Ovalbumin (Chicken)</td>
<td>Natural</td>
</tr>
<tr>
<td>RB</td>
<td>RNase B (Bovine)</td>
<td>Natural</td>
</tr>
<tr>
<td>Xferrin</td>
<td>Transferrin (Bovine)</td>
<td>Natural</td>
</tr>
<tr>
<td>4APBSA</td>
<td>4-aminophenyl-BSA (linker on BSA without sugar)</td>
<td>N/A</td>
</tr>
<tr>
<td>a-C</td>
<td>alpha crystallin from bovine eye lens</td>
<td>Natural</td>
</tr>
<tr>
<td>M3BSA</td>
<td>Man-α-(1,3)-(Man-α-(1,6)-Man-BSA</td>
<td>23</td>
</tr>
<tr>
<td>GlcNAcBSA</td>
<td>GlcNAc-β-BSA</td>
<td>39</td>
</tr>
<tr>
<td>LacNAcBSA</td>
<td>Gal-β-(1,4)-GlcNAc-β-Sp3-BSA</td>
<td>11</td>
</tr>
<tr>
<td>3SLNBSA</td>
<td>Neu5Ac-α-(2,3)-Gal-β-(1,4)-GlcNAc-BSA</td>
<td>13</td>
</tr>
<tr>
<td>3SLacHSA</td>
<td>Neu5Ac-α-(2,3)-Gal-β-(1,4)-Glc-APD-HSA</td>
<td>6</td>
</tr>
<tr>
<td>6SLacHSA</td>
<td>Neu5Ac-α-(2,6)-Gal-β-(1,4)-Glc-APD-HSA</td>
<td>15</td>
</tr>
<tr>
<td>2FLBSA</td>
<td>Fuc-α-(1,2)-Gal-β-(1,4)-Glc-Sp3-BSA</td>
<td>7</td>
</tr>
<tr>
<td>3SFLBSA</td>
<td>Neu5Ac-α-(2,3)-Gal-β-(1,4)-Fuc-α-(1,3)-Glc-Sp3-BSA</td>
<td>7</td>
</tr>
<tr>
<td>H2BSA</td>
<td>Fuc-α-(1,2)-Gal-β-(1,4)-GlcNAc-β-1-APE-BSA</td>
<td>26</td>
</tr>
<tr>
<td>BGABSA</td>
<td>GalNAc-α-(1,3)-[Fuc-α-(1,2)]-Gal-β-(1,4)-GlcNAc-BSA</td>
<td>19</td>
</tr>
<tr>
<td>BGBHSA</td>
<td>Gal-α-(1,3)-[Fuc-α-(1,2)]-Gal-β-(1,4)-GlcNAc-BSA</td>
<td>21</td>
</tr>
<tr>
<td>GGGNHSa</td>
<td>Gal-α-(1,3)-Gal-β-(1,4)-GlcNAc-HSA</td>
<td>9</td>
</tr>
<tr>
<td>Ga3GBSA</td>
<td>Gal-α-(1,3)-Gal-Sp3-BSA</td>
<td>20</td>
</tr>
<tr>
<td>Gb4GBSA</td>
<td>Gal-α-(1,4)-Gal-Sp3-BSA</td>
<td>N/A</td>
</tr>
<tr>
<td>Ga2GBSA</td>
<td>Gal-α-(1,2)-Gal-Sp3-BSA</td>
<td>12</td>
</tr>
<tr>
<td>ManaBSA</td>
<td>Man-α-BSA</td>
<td>N/A</td>
</tr>
<tr>
<td>LNFPIBSA</td>
<td>Fuc-α-(1,2)-Gal-β-(1,3)-GlcNAc-β-(1,3)-Gal-β-(1,4)-Glc-BSA</td>
<td>20</td>
</tr>
<tr>
<td>LNFPIIBSA</td>
<td>Fuc-α-(1,3)-Gal-β-(1,3)-GlcNAc-β-(1,3)-Gal-β-(1,4)-Glc-BSA</td>
<td>15</td>
</tr>
<tr>
<td>LNFPIIIBSA</td>
<td>Gal-β-(1,4)-[Fuc-α-(1,3)]-GlcNAc-β-(1,3)-Gal-β-(1,4)-Glc-BSA</td>
<td>20</td>
</tr>
<tr>
<td>Cerulo</td>
<td>Ceruloplasmin</td>
<td>Natural</td>
</tr>
<tr>
<td>LebBSA</td>
<td>Fuc-α-(1,2)-Gal-β-(1,3)-[Fuc-α-(1,4)]-GlcNAc-BSA</td>
<td>10</td>
</tr>
<tr>
<td>LexBSA</td>
<td>Gal-β-(1,4)-[Fuc-α-(1,3)]-GlcNAc-BSA</td>
<td>24</td>
</tr>
<tr>
<td>DiLexBSA</td>
<td>Gal-β-(1,4)-[Fuc-α-(1,3)]-GlcNAc-β-(1,3)-Gal-β-(1,4)-[Fuc-α-(1,3)]-GlcNAc-β-1-O-APE-BSA</td>
<td>28</td>
</tr>
<tr>
<td>A1AT</td>
<td>alpha-1-antitrypsin (human)</td>
<td>Natural</td>
</tr>
</tbody>
</table>
### Chapter IV

<table>
<thead>
<tr>
<th>Name</th>
<th>Structure</th>
<th>Chapter IV</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>3LexHSA</td>
<td>Gal-(\beta)-(1,4)-[Fuc-\alpha-(1,3)]-GlcNAc-\beta-(1,3)-Gal-\beta-(1,4)-[Fuc-\alpha-(1,3)]-GlcNAc-\beta-O-APD-HSA</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>3SLexBSA3</td>
<td>Neu5Ac-\alpha-(2,3)-Gal-\beta-(1,4)-[Fuc-\alpha-(1,3)]-GlcNAc-Sp3-BSA</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Fibrin</td>
<td>Fibrinogen (Human)</td>
<td>Natural</td>
<td></td>
</tr>
<tr>
<td>6SuLexBSA</td>
<td>(SO4)6-Gal-\beta-(1,3)-[Fuc-\alpha-(1,4)]-GlcNAc-SP3-BSA</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>6SuLeaBSA</td>
<td>(SO4)6-Gal-\beta-(1,4)-[Fuc-\alpha-(1,3)]-GlcNAc-SP3-BSA</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>3SuLeaBSA</td>
<td>(SO4)3-Gal-\beta-(1,4)-[Fuc-\alpha-(1,3)]-GlcNAc-BSA</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>3SuLexBSA</td>
<td>(SO4)3-Gal-\beta-(1,3)-[Fuc-\alpha-(1,4)]-GlcNAc-BSA</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>DFPLNHSHA</td>
<td>Gal-\beta-(1,3)-[Fuc-\alpha-(1,4)]-GlcNAc-\beta-(1,3)-Gal-\beta-(1,4)-Glc-APD-HSA</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>LeaBSA</td>
<td>Gal-\beta-(1,3)-[Fuc-\alpha-(1,4)]-GlcNAc-BSA</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>LeyHSA</td>
<td>Fuc-\alpha-(1,2)-Gal-\beta-(1,4)[Fuc-\alpha-(1,3)]-GlcNAc-\beta-1-O-APD-HSA</td>
<td>13.3</td>
<td></td>
</tr>
<tr>
<td>3FLeyHSA</td>
<td>Fuc-\alpha-(1,2)-Gal-\beta-(1,4)[Fuc-\alpha-(1,3)]-GlcNAc-\beta-(1,3)-Gal-\beta-(1,4)-[Fuc-\alpha-(1,3)]-GlcNAc-\beta-1-APD-HSA</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>LNNnTHSA</td>
<td>Gal-\beta-(1,4)-GlcNAc-\beta-(1,3)-Gal-\beta-(1,4)-Glc-APD-HSA</td>
<td>9.2</td>
<td></td>
</tr>
<tr>
<td>LNTTHSA</td>
<td>Gal-\beta-(1,3)-GlcNAc-\beta-(1,3)-Gal-\beta-(1,4)--APD-HSA</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>SLNFVSHA</td>
<td>Fuc-\alpha-(1,2)-Gal-\beta-(1,3)-[Neu5Ac-\alpha-(2,6)]GlcNAc-\beta-(1,3)-Gal-\beta-(1,4)-Glc-APD-HSA</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>MMLNnHSHA</td>
<td>Neu5Ac-\alpha-(2,3)-Gal-\beta-(1,4)-GlcNAc-\beta-(1,3)-[Gal-\beta-(1,4)-(Fuc-\alpha-(1,3))GlcNAc-\beta-(1,6)]Glc-\beta-(1,4)-Glc-APD-HSA</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>SLNnTHSA</td>
<td>Neu5Ac-\alpha-(2,3)-Gal-\beta-(1,4)-GlcNAc-\beta-(1,3)-Gal-\beta-(1,4)-Glc-APD-HSA</td>
<td>9.3</td>
<td></td>
</tr>
<tr>
<td>GM1HSA</td>
<td>Gal-\beta-(1,3)-GlcNAc-\beta-(1,4)-[Neu5Ac-\alpha-(2,3)]-Gal-\beta-(1,4)-Glc-APD-HSA</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>aGM1HSA</td>
<td>Gal-\beta-(1,3)-GlcNAc-\beta-(1,4)-Gal-\beta-(1,4)-Glc-APD-HSA</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>GlobNTHSA</td>
<td>GalNAc-\beta-(1,3)-Gal-\alpha-(1,4)-Gal-\beta-(1,4)-Glc-APD-HSA</td>
<td>8.4</td>
<td></td>
</tr>
<tr>
<td>GlobTHSA</td>
<td>Gal-\alpha-(1,4)-Gal-\beta-(1,4)-Glc-\beta-1-APD-HSA</td>
<td>21</td>
<td></td>
</tr>
</tbody>
</table>

N/A – Not assessed
Chapter V

Transcriptional response of human intestinal epithelial cells to bovine buttermilk

Sarah A. Ross, John O’Callaghan, Lokesh Joshi and Rita M. Hickey

Submitted to Journal of Dairy Science

Acknowledgements: IMGM, Germany performed microarray analysis. Sarah Ross performed all other experimental work for this chapter. The authors would like to graciously acknowledge the advice and contribution of Dr. Aoife Thompson.
Chapter V

Abstract

Human milk oligosaccharides (HMO) are associated with many health-promoting properties including anti-inflammatory activities. However, the large quantities of HMO required for studies and for incorporation into functional foods make HMO isolation difficult. Therefore, bovine milk is of interest as an alternative sustainable source of glycosylated components with structural similarity to HMO. Bovine buttermilk in particular is an excellent source of glycoconjugates and is a by-product of butter making. In this study, the transcriptional response of human colonic epithelial cells (HT-29) to bovine buttermilk was examined in order to assess the potential of this ingredient as an alternative to HMO. Since it is known that both the abundance of proteins and their glycosylation status changes over the course of lactation, we compared the transcriptional response of the human cells to buttermilk generated from colostrum and from mature milk. Analysis of the data identified a number of processes which were affected by both treatments including apoptosis, immune system processes, barrier function, metabolism and stimulus response. Immune system processes were the most significantly modulated with RT-PCR validation revealing that mature and colostrum buttermilk affected the expression of chemokines and cytokines including CXCL10, CXCL11, IL-8, IL-17D, IL-33 and anti-viral genes including OAS2, OAS3, IFIT1, IFIT3, DDX58 and genes involved in the JAK/STAT pathway including JAK2, STAT1 and STAT2 in a similar manner. Transcriptional responses to both samples also differed with colostrum buttermilk leading to downregulation of CCL5, CCL20, CXCL2, CXCL3 and IFNβ1 while mature buttermilk resulted in upregulation of the same genes. Additionally, IL-7 was upregulated by colostrum buttermilk but downregulated by mature buttermilk. Overall, the present study demonstrated the potential of colostrum and mature buttermilk to elicit an immunological response and in particular an anti-inflammatory phenotype. As such, buttermilk may have a potential role as a functional food in the prevention or treatment of inflammatory diseases.
Chapter V

Introduction

Human milk oligosaccharides (HMO) have been shown to possess many health benefits in humans including prevention of pathogen adhesion, prebiotic effects and modulation of gut microflora (Hickey 2009). HMO also possess immunomodulatory activities which play an important role in the infant gut. For instance, HMO and human milk glycoconjugates may function in immune homeostasis through interactions with C-type lectin receptors (García-Vallejo and Van Kooyk 2009, Koning, Kessen et al. 2015) and acidic HMO may promote a T-cell phenotype which contributes to allergy prevention (Eiwegger, Stahl et al. 2010). However, the limited availability of human milk has made it difficult to isolate and purify large quantities of HMO. In this respect, recent research has focused on domestic animal milks as a source of health promoting oligosaccharides. Indeed, bovine milk oligosaccharides (BMO) share some structural similarity with HMO (Gopal and Gill 2000, Barile, Tao et al. 2009). The transcriptional responses of human intestinal epithelial cells to both HMO and BMO have been compared. For example Lane et al. (2013) demonstrated that exposure to both HMO and BMO resulted in similar gene regulation in HT-29 cells, highlighting the potential of BMO as a viable alternative to HMO in functional foods. Moreover, the authors revealed the potential contribution of BMO to the development and maturation of the intestinal immune response in vitro. Unfortunately, the concentration of BMO in whole milk is low, particularly when compared to HMO in whole milk (0.09-1.2 g/L and 6-23 g/L, respectively) (Peterson, Cheah et al. 2013). This has resulted in researchers turning their attention to the many glycoconjugates present in bovine milk which possess several potential health-promoting bioactivities in humans (Ross, Lane et al. 2015) and some of which are present at substantially higher concentrations than free oligosaccharides. For example, the main protein of bovine milk fat globule membrane (MFGM) is the glycoprotein butyrophilin, which comprises approximately 40% of total MFGM proteins (Spitsberg 2005). Glycoproteins of bovine MFGM are highly sialylated consisting predominantly of O-linked mono- and di-sialylated core-1 glycans (Wilson, Robinson et al.
Chapter V

Moreover, bi-, tri-, and tetra-antennary complex-type and hybrid-type high-mannose N-linked oligosaccharide chains (Sato, Furukawa et al. 1993) and gangliosides (sialylated lipids) are also found associated with MFGM (Takamizawa, Iwamori et al. 1986), demonstrating some of the rich and diverse glycosylation present in bovine MFGM.

Bovine MFGM has been shown to possess anti-adhesive and anti-infective (Kvistgaard, Pallesen et al. 2004), anti-cholesterol (Ito, Kamata et al. 1992), and anti-viral properties (Yolken, Peterson et al. 1992). Additionally, bioactivities including improvement of cognitive function in infants has also been attributed to bovine MFGM (Timby, Domellöf et al. 2014). Many reviews have highlighted the potential use of bovine MFGM and its associated glycoconjugates in functional foods for the benefit of human health (Singh 2006, Ross, Lane et al. 2015). Another aspect of MFGM making it an attractive alternative to milk oligosaccharides is the ease at which it can be sourced. For instance, buttermilk is rich in MFGM and its associated glycoconjugates and represents an undervalued dairy fraction. It is produced as a by-product of butter-making (Jiménez-Flores and Brisson 2008) and is available in large quantities in the dairy industry, highlighting its commercial potential as a source of bioactive glycoconjugates. Buttermilk is rich in glycoconjugates such as mucins, butyrophilin and lactadherin (Ross, Lane et al. 2015). It may be feasible to isolate MFGM from buttermilk at a commercial scale through the use of processes such as micro- and ultrafiltration (Dewettinck, Rombaut et al. 2008). Interestingly, the type and abundance of glycan structure as well as protein concentration can change over the course of lactation for individual bovine glycoproteins (Ujita, Furukawa et al. 1993, Takimori, Shimaoka et al. 2011, O'Riordan, Gerlach et al. 2014), for whole MFGM (Ujita, Furukawa et al. 1993, Wilson, Robinson et al. 2008) and for buttermilk (Ross, Gerlach et al. 2016). These compositional temporal changes may affect the bioactivities of milk glycoproteins.
Chapter V

Non-human MFGM has previously been shown to have an immunomodulatory response and was shown to inhibit the proliferation of LPS- and ConA-activated spleenocytes through the counteraction of pro-inflammatory cytokine production such as interferon (IFN)-γ, TNF-α and IL-4 (Zanabria, Tellez et al. 2014). Individual proteins associated with MFGM have also been shown to have an effect on immunomodulation. For example, lactadherin may modulate immune function by contributing to T cell differentiation and enhancing IL-10 secretion by dendritic cells, which may be linked to modulation of intestinal allergic reactions (Zhou, Gao et al. 2010). Murine butyrophilin can modulate immune function through the inhibition of T cell activation (Smith, Knezevic et al. 2010) while the bovine whey protein lactoferrin may stimulate the mucosal immune system (Debbabi, Dubarry et al. 1998). However, very few studies to date have examined the transcriptomic response of human epithelial cells to bovine MFGM from different lactation stages. In this study, we compared the effect of bovine colostrum buttermilk (cBM) and mature buttermilk (mBM) on the transcriptional response of colonic adenocarcinoma epithelial HT-29 cells. Specifically, we investigated their contribution to the development of the intestinal immune response by concentrating on the expression profiles of immune system associated genes. The study indicated some differences exist in the transcriptional responses on HT-29 cells as a result of exposure to cBM and mBM. For instance, the genes CCL5, CCL20, CXCL2, CXCL3 and IFNβ1 were downregulated in response to cBM but upregulated in response to mBM. Furthermore, IL-7 was upregulated by cBM but downregulated by mBM. The study suggests that buttermilk may be an attractive source of glycoconjugates for food applications.
Chapter V

Materials and Methods

Generation of buttermilk
Milk was collected from 3 multiparous Holstein-Friesian cattle at the Teagasc Food Research Centre, Moorepark (Fermoy, Co. Cork). The colostrum sample was generated from the combination of milk taken from 3 animals on days 1, 2 and 3 post-parturition. The mature sample was generated from the combination of milk taken from the same 3 animals on days 30, 70 and 90 post-partum. Whole milk was separated into cream and skim milk using an FT15 disc bowl centrifuge (Armfield Ltd.) and the cream was stored at 4 °C for 12-24 h. Cream was separated into buttermilk and butter using a mixer. The buttermilk was then passed through glass wool (Sigma) to remove contaminating butter granules. Samples were lyophilised and stored in a desiccator at room temperature.

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis
Sample preparation and reduction was performed as per manufacturer’s instructions (NuPAGE system, Life Technologies, Thermo Fisher Scientific Inc.). Briefly, 2.5 µL of sample buffer and 1 µL of reducing agent were added to 6.5 µL of sample 10 mg/mL (by mass of powder). The sample was heated to 70 °C for 10 min and 10 µL of sample was added to each well of a 4-12% Bis-Tris gel (1.00 mm x 9 well, Life Technologies). 10 µL of a molecular weight standard solution (Thermo Scientific PageRuler Prestained Protein Ladder) was also loaded onto the gel. Electrophoresis was performed at 200 V for 50 min using MOPS buffer with MOPS buffer supplemented with 0.25% NuPAGE Antioxidant (Life Technologies) in the upper chamber. Protein bands were visualized on the gels using Coomassie blue stain (Invitrogen SimplyBlue SafeStain) following the manufacturer’s procedure. MFGM proteins and contaminating proteins were identified through using the bovine milk protein standards β-lactoglobulin A, β-lactoglobulin B, α-lactalbumin, κ-casein, α-
casein, β-casein, lactoferrin, and immunoglobulin G (IgG) and comparison with previously published SDS-PAGE analysis (Rombaut, Dejonckheere et al. 2006, Singh 2006, Rombaut, Dejonckheere et al. 2007).

**Cell culture and exposure**

The human colonic adenocarcinoma cell line, HT-29, was purchased from the American Type Culture Collection. HT-29 cells were routinely grown in McCoy’s 5A modified medium (Sigma-Aldrich®) supplemented with 10% FBS. All cells were routinely maintained in 75 cm² tissue culture flasks and incubated at 37 °C in 5% (v/v) CO₂ in a humidified atmosphere. Cells were passaged when the confluency of the flask was approximately 90%. For transcription analysis, HT-29 cells were seeded into 75 cm² tissue culture flasks and grown for 4-5 days with the medium being changed every second day until the confluency of the flask was 100%. The media was changed to McCoy’s 5A modified medium supplemented with 2% FBS at least 24 h prior to treatment. During treatment cells were exposed to filter sterilised McCoy’s 5A modified medium supplemented with mBM or cBM (5mg/mL). Unsupplemented McCoy’s 5A modified medium was used as a control. After 24 h cells were harvested in RNA later (Qiagen). RNA later was removed by centrifugation at 14,000 RPM for 1.5 min and removal of supernatant from RNA pellet. Cell pellets were frozen at -80 °C.

**RNA isolation**

Total RNA was extracted from treated and untreated HT-29 cells using Roche RNA Isolation Kit following manufacturer’s protocol. RNA quantity and quality was assessed on a NanoDrop 1000 (Wilmington, DE, USA). Only samples with \( A_{260}/A_{280} > 1.95 \) and \( A_{230}/A_{260} > 1.92 \) were used during the analysis. Sample quality was further assessed by agarose gel electrophoresis. RNA samples were subsequently shipped to IMGM laboratories (Martinsried, Germany) for labelling and microarray analysis. Prior to their use in the microarray experiments RNA concentration and purity was assessed using
NanoDrop 1000 and RNA integrity was assessed using 2100 Bioanalyzer and RNA 6000 Nano LabChip (both Agilent Technologies). Only RNA samples with an RNA integrity number (RIN) > 8.0 were used for microarray analysis.

**Microarray hybridisation**

Prior to labelling, the RNA was spiked with synthetic polyadenylated transcripts (RNA Spike-In Mix, Agilent Technologies) and labelled using an RT-IVT protocol. Briefly, spiked total RNA was reverse transcribed to cDNA and subsequently converted into Cyanine-3 labelled cRNA (Quick-Amp Labelling Kit One-Colour, Agilent Technologies). 600ng of labelled cRNA was fragmented and prepared for one-colour-based hybridization (Gene Expression Hybridization Kit, Agilent Technologies). Cyanine-3-labelled cRNA samples were subsequently hybridised to Agilent SurePrint G3 Human Gene Expression 8x60K v2 microarrays at 65 °C for 17 h. Arrays were washed (Gene Expression Wash Buffers, Agilent Technologies) and dried (acetonitrile, Sigma) prior to detection of fluorescent signal intensities. Scan Control A.8.4.1 software (Agilent Technologies) on the Agilent DNA Microarray Scanner was used to detect fluorescent signal intensities and images were extracted using Feature Extraction 10.7.3.1 software (Agilent Technologies) and the design file 039494_D_F_20140326.

**Quantitative PCR**

To validate the expression of target genes HT-29 cells were re-exposed to cBM or mBM samples. Total RNA was isolated from treated and untreated HT-29 cells using High Pure RNA Isolation Kit (Roche) and RNA quantity and quality was assessed as described above. Only samples with $A_{260}/A_{280} \geq 2$ and $A_{230}/A_{260} \geq 2.1$ were used for cDNA synthesis. The purity of the RNA and absence of genomic DNA contamination was confirmed by PCR analysis. cDNA was synthesised using a Transcriptor First Strand cDNA Synthesis kit (Roche) with
1 µg of total RNA incubated with 1 µL anchored-oligo(dT) primer and 2 µL random hexamer primer in a final volume of 20 µl as per manufacturer’s instructions. This was incubated at 25 °C for 10 min, followed by 50 °C for 60 min. Reverse Transcriptase was inactivated by heating to 85 °C for 5 min. cDNA was subsequently stored at -20 °C until PCR analysis. The cDNA product was amplified (RT-PCR) using RealTime Ready Custom Panels 96-24 (Roche) and a LightCycler® 480 (Roche). Reference genes used included ribosomal phosphoprotein P0, 18S RNA and β-2 microglobulin (Dydensborg, Herring et al. 2006). Each PCR contained LightCycler® 480 Probes Master, PCR-grade water and 50 ng/µL complementary DNA as per manufacturer’s instructions. The PCR running conditions were 10 min denaturation step at 95 °C, 45 amplification cycles consisting of 10 s at 95 °C, 30 s at 60 °C and 1 s at 72 °C, and a final 30 s cooling step at 40 °C.

**Data analysis**

*Gene expression analysis.*

The raw microarray results were analysed using Feature Extraction 10.7.3.1, GeneSpring GX 12.6.1 (both Agilent Technologies), Excel 2010 (Microsoft) and the tool marfin v1.9. Data was normalized by quantile normalization. The fold changes in gene expression were calculated as fold change relative to the control samples (cells treated with unsupplemented media). The cut-off for identifying genes of interest was a fold change >2 or < -2 and uncorrected p-value ≤0.05. Statistical significance was calculated using Welch’s approximate t-test. P-values were not adjusted for multiple testing errors as false detections would be detected by real time quantitative PCR analysis, which was used to confirm a panel of the differentially expressed genes identified from the microarray.
Chapter V

Quantitative PCR analysis.

All samples were run in triplicate on three separate occasions. Data was analysed by the ΔΔCT method (Livak and Schmittgen 2001) and presented as fold changes in gene expression, compared to control, after normalization to an internal reference gene expression level. Data is presented as a mean of triplicate data and error bars represent the range of possible relative expression values defined by the standard error of the ΔCT.
Chapter V

Result and Discussion

Sample characterization

The purpose of the study was to generate samples rich in glycoconjugates from bovine colostrum and mature milk and assess their effect on the transcriptional response of HT-29 cells. Buttermilk was chosen as a source given the rich glycoconjugate complement associated with the milk fat globule membrane. Compositional and comparative analyses of both samples were performed by SDS-PAGE analysis. SDS-PAGE gels were stained for total proteins (Figure 1). A molecular mass standard and previously published data (Rombaut, Dejonckheere et al. 2006, Singh 2006, Rombaut, Dejonckheere et al. 2007) were used to identify specific proteins. Approximately 15 protein bands were identified. High intensity bands were identified for the free milk proteins α-casein, β-lactoglobulin and proteose peptone 3 (approximately 34, 19 and 28 kDa, respectively). Low intensity bands were observed for α-lactalbumin, β-casein and κ-casein (approximately 14, 31 and 28 kDa, respectively) and IgG (approximately 26 kDa). High intensity bands were also observed for MFGM glycoproteins such as xanthine oxidoreductase and butyrophilin (approximately 130 and 58 kDa, respectively). Low intensity bands were seen for the MFGM glycoproteins CD36, Pas6 and Pas7 (approximately 67, 52 and 50 kDa, respectively). These findings were in agreement with previous publications (Rombaut, Dejonckheere et al. 2006, Singh 2006, Rombaut, Dejonckheere et al. 2007).

The glycosylation of the samples used in this study have been profiled through the use of lectin microarrays (Ross, Gerlach et al. 2016). The presence of structures including mannose, N-acetylglucosamine, N-acetylgalactosamine, fucosylated structures, sialylated structures, O-linked mucin core type 1 glycans and galactose were shown to be present in the buttermilk samples. Interestingly, the study highlighted the differences in glycosylation of buttermilk generated from colostrum compared to mature milk. For instance, relative proportions of
mannose, biantennary structures, N-acetylglucosamine and fucosylated structures were much greater in colostrum than in mature buttermilk samples.

**Overall transcriptional response**

The transcriptional response of the HT-29 cells to the above samples was assessed using total RNA and Agilent Human GE microarrays. Differentially expressed transcripts (DET) were identified by comparing the intensity data for untreated cells (no buttermilk) with those for the treated cells (buttermilk). Importantly, the samples did not affect HT-29 cell viability as confirmed by simple trypan blue staining. The number of DET induced by cBM and mBM was 841 and 527 respectively with 286 DET common to both treatments (Figure 2). These shared responses indicate the structural similarity of both samples. For example, both mature milk and colostrum are known to contain structures including Pas6 and Pas7, xanthine dehydrogenase, mucins and α-lactalbumin (Reinhardt and Lippolis 2008, Takimori, Shimaoka et al. 2011). The number of DET unique to the cBM and mBM was 555 and 241 respectively. For both treatments, the number of DET that were downregulated was substantially greater than the number upregulated. Furthermore, analysis of the DET for statistically over-represented genes indicated that the downregulated DET contained a number of gene ontology groups that occurred at a frequency of 3-5 fold higher than expected (Table 1). In contrast, upregulated genes did not contain any identifiable groups that were over-represented, possibly due to the smaller number of upregulated genes after each treatment. For cells treated with cBM the number of ontology groups over-represented was 8 while for mBM treated cells there were 7 groups over-represented (Table 1). The most significant of the over-represented groups common to both treatments was a group of genes associated with the response to IFN-γ (Table 1). IFN-γ is an important component of the cellular response to pathogenic microorganisms however dysregulation of the IFN-γ response can lead to inflammatory disorders and concomitant impairment of epithelial function (Ferrier, Mazelin et al. 2003). IFN-γ induced inflammation was found to be reduced in model systems by exposure to probiotic bacteria (Resta–Lenert
and Barrett 2006). It is interesting that in the current study, both mBM and cBM resulted in similar effects. The fact that mBM had this effect is of particular interest given the wide availability of buttermilk produced from mature bovine milk.

The higher number of DET unique to cBM could be as a result of different concentrations of proteins and oligosaccharide structures in cBM when compared to mBM. For instance, it is known that the concentrations of proteins such as IgG and lactoferrin, and glycans such as fucose and sialic acid are higher in bovine colostrum (day 1 post-partum) when compared to whole milk samples taken at later timepoints (weeks 1-4 post-partum) (Takimori, Shimaoka et al. 2011). Furthermore, the concentration of O-linked core 2 glycans are higher in colostral MFGM samples (3 and 7 days post-partum) when compared to mature MFGM samples (6-8 weeks post-partum) (Wilson, Robinson et al. 2008). These compositional differences may be linked to the importance of colostrum in protection of the mammalian gut from pathogenic infection and development of the immune response in the infant (Newburg and Walker 2007).

**Differential transcription of immune related genes**

Genes involved in the immune response were also identified as being some of the most highly differentially expressed in response to treatment with cBM or mBM (Table 1). The number of immune related genes over-represented following exposure to buttermilk was similar for both mature (73 genes) and colostrum (74 genes) and the breakdown of immune response and immune system process related genes was also similar for both types of buttermilk (Table 2). The resulting genes could be separated into 3 classes: chemokines, interleukins and intercellular signalling molecules however the significance of the immune-related categories differed for each of the treatments (Table 2). The classes identified were similar to those observed by Lane et al. (2013) where the effects of BMO and HMO generated from milk obtained one day post-
partum on HT-29 cells was studied. The genes were assigned to canonical pathways defined in the PANTHER database. The JAK/STAT signalling pathway was found to be significantly over-represented in the genes of the mature sample data set however this pathway was not statistically significantly influenced by cBM (Table 2).

**Improved Barrier function**

A number of cadherin/protocadherin genes were upregulated by both treatments (Table 3). These genes are essential for the correct functioning of the intestinal epithelium through their role in the maintenance of tight junctions. Modulation of cadherin gene expression has been cited as a possible mode of action of probiotic bacteria in the maintenance of intestinal function (Hummel, Veltman et al. 2012) and it is interesting that both buttermilk samples upregulated these genes. In this study, upregulation of genes involved in formation of tight junctions including PCDH11X, PCDHGB7, PCDHGC4, CDH16 and FAT2 were identified by microarray analysis (Table 3). Cadherin type 1, cadherin type 23, protocadherin 6, claudin 4 and claudin 25 are upregulated by exposure to *Bifidobacterium bifidum* (Turroni, Taverniti et al. 2014). The upregulation of such genes potentially increases cell-cell adhesion and strengthens intestinal epithelial barrier, promoting protection from pathogenic infection. Upregulation of the anti-microbial beta defensin gene (DEFB1) in response to mBM as observed in this study (Table 3), is also similar to the upregulation observed for *B. bifidum* (Turroni, Taverniti et al. 2014). Along with the cell-cell tight junctions, the mucin layer overlaying intestinal epithelial cells also has a vital role in epithelial function by shielding the epithelial cells from the intestinal contents (Moehle, Ackermann et al. 2006). Both treatments resulted in an upregulation of MUC5AC which may enhance the epithelial barrier (McGuckin, Lindén et al. 2011). Another mucin gene, MUC13, was downregulated by both treatments (Table 3). This may result in a beneficial effect as mucin 13 has oncogenic potential and over-expression of MUC13 is
often associated with cancer cells, such as epithelial carcinomas, rather than healthy epithelial cells (Maher, Gupta et al. 2011).

**Validation of microarray results by qPCR**

To validate and examine the differences in expression of a panel of selected immune related genes (Table 4), real time PCR analysis was performed using RealTime ready custom panels that included genes shown to be differentially expressed in the microarray analysis, compared to the control. Differential expression was confirmed for the chemokines chemokine (C-X-C motif) ligand 2 (CXCL2), chemokine (C-X-C motif) ligand 3 (CXCL3), chemokine (C-X-C motif) ligand 10 (CXCL10), chemokine (C-X-C motif) ligand 11 (CXCL11), chemokine (C-C motif) ligand 5 (CCL5), chemokine (C-C motif) ligand 20 (CCL20) and interleukins, IL-7, IL-8, IL-17D and IL-33. Furthermore, differences in the expression of members of the JAK/STAT pathway, including signal transducer and activator of transcription (STAT) 1, STAT2 and Janus kinase (JAK) 2 were verified. Additionally, genes involved in viral immunity including interferon regulatory factor (IRF) 7, IRF9, oligoadenylate synthetase (OAS) 2, OAS3, IFN-induced protein with tetratricopeptide repeats (IFIT) 1, IFIT3 and DEAD (Asp-Glu-Ala-Asp) Box Polypeptide 58 (DDX58), which encodes retinoic acid-inducible gene I (RIG I), were shown to be differentially expressed (Table 4) in response to the buttermilk samples.

Exposure of the HT-29 cells to cBM and mBM resulted in the differential expression of IFNβ1, a type I interferon. IFNβ expression in mice has previously been shown to be regulated by oral administration of bovine milk lactoferrin (Wakabayashi, Takakura et al. 2006). The importance of IFNβ is highlighted by a study conducted by Deonarain et al. (2000) whereby the high susceptibility of mice to viral infection following targeted destruction of the IFNβ gene was demonstrated. During bacterial or viral infection pathogen associated molecular patterns are detected by pattern recognition receptors. Downstream signalling through MyD88- or TRIF-dependent signalling
pathways, leads to the induction of IRF7 and IRF3 expression which subsequently induces secretion of IFNβ (Zhou, Michal et al. 2013). Subsequent binding of IFNβ to cell receptors leads to activation of IFN stimulated genes (ISGs) (Goubau, Deddouche et al. 2013) including IFIT, ribonuclease (RNase) L and OAS family members. Downregulation of many of these ISGs in response to the buttermilk treatments was observed, as well as a differential regulation of IRF7 and IRF9 (Table 4). This may indicate the potential of MFGM as a functional ingredient since exposure did not lead to unnecessary expression of anti-viral genes. This highlights the potential use of buttermilk as a functional food as an unwanted immune response may not be elicited.

PCR analysis highlighted the potential of bovine mBM and cBM to elicit a response in JAK/STAT signalling pathway components. In particular, JAK2, STAT1 and STAT2 were downregulated. During viral infection, secretory IFNβ binds its extracellular receptor leading to activation of the JAK/STAT signalling pathway. Phosphorylated Stat1 and Stat2 form a complex with IRF9 (which was also shown to be regulated in response to the buttermilk samples) known as the IFN-stimulated gene factor-3 (ISGF-3). This complex subsequently binds to promoters of ISGs, such as OAS and IFIT, which in turn leads to their expression (Diamond and Farzan 2013). OAS proteins catalyse ATP oligomerisation into 2’,5’-linked oligoadenylate in response to RNA. 2’,5’-linked oligoadenylate subsequently activates RNase L which degrades viral RNA (Hovanessian 2007). IFIT proteins also play a vital role in anti-viral immunity where they function in the prevention of viral replication through binding viral RNA and preventing initiation of transcription (Zhou, Michal et al. 2013). OAS2, OAS3, IFIT1 and IFIT3 were shown to be downregulated in response to both cBM and mBM in this study (Figure 3). The level of downregulation of the OAS and IFIT family members was much higher in response to cBM when compared to mBM. This could be as a result of the higher levels and diversity of milk components in colostrum when compared to mature milk. Milk components are known to have an effect on the transcription of anti-viral genes. For instance, IFIT1 and IFIT3 are upregulated in non-
transformed human intestinal epithelial cells by human and commercial bovine lactoferrin (Jiang and Lönnertal 2014). Although these authors demonstrated upregulation of IFIT transcription in response to lactoferrin, it is possible that the mixture of glycoconjugates in the buttermilk samples in the present study led to the downregulation observed. Our data also demonstrated the differential regulation of DDX58 gene by the buttermilk samples compared to the control (Figure 3). RIG-I activity is promoted by the binding of the protein encoded by DDX60 and activation of RIG-I by foreign RNA leads to downstream signalling events causing further activation of transcription factors, such as IRF3 and IRF7, leading to further increases in IFNβ expression (Goubau, Deddouche et al. 2013). A positive feedback loop exists between IFNβ1 and RIG-I in viral infection, where JAK/STAT pathways induced by IFNβ result in further expression of the DDX58 gene (Hu, Nudelman et al. 2011). In our study, both mBM and cBM led to downregulation of RIG-I which could further signify the potential use of buttermilk as a functional food, as expression of genes involved in the immune response are not activated in response to the buttermilk.

The interleukins IL-8, IL-17D, IL-33 and IL-7 were also differentially regulated in response to the buttermilk samples compared to the control (Figure 3). IL-33 is known to function in the chemoattraction of Th2 cells (Komai-Koma, Xu et al. 2007) and in inflammatory processes whereby it induces cytokine production in Th2-polarized naive T cells (Schmitz, Owyang et al. 2005). The downregulation observed in this study (Table 4) may indicate an anti-inflammatory response in the HT-29 cells. Similarly, IL-7 plays a role in the immune response whereby it functions in T cell production through involvement in thymopoiesis (Beq, Delfraissy et al. 2004). This interleukin was upregulated in response to the cBM and downregulated in response to the mBM. This may reflect the higher concentration and diversity of oligosaccharides and proteins present in the colostrum (Ross, Gerlach et al. 2016) and may indicate the requirement of increased protection in the neonatal gut against pathogenic infection in the days after birth. Studies suggest the
interleukin IL-17D plays an important role in inflammation as it has been shown to increase the production of cytokines including IL-6 and IL-8 in human endothelial cells (Starnes, Broxmeyer et al. 2002), human subepithelial myofibroblasts (Yagi, Andoh et al. 2007) and in chicken fibroblasts (Hong, Lillehoj et al. 2008). IL-17D also seems to play a role in the prevention of tumour progression (Saddawi-Konefka, O'Sullivan et al. 2013). Expression of this cytokine was reduced in HT-29 cells treated with both cBM and mBM, indicating a potential decrease in pro-inflammatory cytokine production which may be beneficial in the absence of pathogenic infection. Anti-microbial activity has been attributed to IL-8 with activity evident against *Escherichia coli* (Yung and Murphy 2012). Expression of this pro-inflammatory interleukin decreased in response to both cBM and mBM. Buttermilk components are known to play a role in regulation of interleukin expression. Recombinant mouse lactadherin has been shown to reduce the expression of IL-17 in vivo by modulating Stat3 (Cen, Aziz et al. 2015) and bovine and human milk lactoferrin reduce the levels of pro-inflammatory IL-6 in mouse plasma after LPS stimulation (Machnicki, Zimecki et al. 1993, Kruzel, Harari et al. 2002). In addition, the bovine whey proteins, α-lactalbumin, and β-lactoglobulin are known to decrease IL-8 and IL-6 release in rats (Yamaguchi and Uchida 2007).

The current study suggests that some chemokines may be regulated by cBM and mBM. The chemokine displaying the highest level of differential regulation, compared to the control, was CCL20 as validated by PCR (Figure 3) in response to both cBM and mBM samples. Interestingly, cBM led to a reduction in expression of all chemokines while mBM led to a reduction in CXCL10 and CXCL11 expression but an increase in CXCL2, CXCL3, CCL5 and CCL20 expression. The expression of pro-inflammatory genes may lead to an inflammatory response which in some cases can have a negative effect, such as in allergic reactions including psoriasis, asthma and rheumatoid arthritis (Charo and Ransohoff 2006). The downregulation of pro-inflammatory genes such as IL-8 as observed in this study may counteract the upregulation of other pro-inflammatory genes, such CCL20 in response to mBM, thus potentially
resulting in a more balanced inflammatory response while leaving the immune system primed to respond rapidly to pathogenic invasion and other challenges. Additionally, anti-microbial effects have been attributed to these chemokines (Yung and Murphy 2012). CCL20 is known to have anti-microbial properties against pathogens such as *E. coli*, *Pseudomonas aeruginosa*, *Streptococcus pyogenes* and *Staphylococcus aureus* (Yang, Chen et al. 2003). Additionally, CXCL2 and CXCL3 are known to possess anti-microbial activity against *E. coli* and *S. aureus* (Yang, Chen et al. 2003). Reports on anti-microbial activity associated with CCL5 vary and activity seems to be dependent on the assay conditions used (Yang, Chen et al. 2003). The upregulation of these chemokines in response to mBM may have a beneficial effect in protection from microbial infection.

Overall, a large number of pro-inflammatory genes were downregulated in response to both buttermilk treatments, particularly cBM (Figure 3). Eliciting an anti-inflammatory response may play an important role in the treatment of inflammatory bowel diseases (IBD) including ulcerative colitis and Crohn’s disease. For instance, goat milk oligosaccharides were able to reduce intestinal inflammation and associated symptoms in rat models with dextran sodium sulfate-induced colitis, and aid in the recovery of damaged colonic mucosa. These oligosaccharides decreased the production of pro-inflammatory cytokines such as IL-1β (Lara-Villoslada, Debras et al. 2006). Similarly, sialylated bovine glycomacropeptide, derived from κ-casein, was shown to have an anti-inflammatory effect in rats with trinitrobenzenesulfonic acid-induced colitis and resulted in a reduction in pro-inflammatory cytokine mRNA levels (Daddaoua, Puerta et al. 2005). Interestingly, the effects observed with glycomacropeptide were comparable to those observed with sulfasalazine, an established drug used for treatment of IBD, indicating the potential for milk-derived oligosaccharides in the management of IBD (Daddaoua, Puerta et al. 2005). Pro-inflammatory diseases, particularly those caused by Th1-type T cells, could potentially be prevented or treated by anti-inflammatory milk oligosaccharides. For example, the milk oligosaccharide lacto-N-neotetraose can modulate T-cell type by
suppressing Th1-type and driving towards a Th2-phenotype, inducing an anti-inflammatory response (Terrazas, Walsh et al. 2001).

The differential gene regulation observed in this study is similar to the transcriptional response of HT-29 cells to commensal bacteria. For instance, downregulation of genes including CXCL2 and CXCL3 is induced by *B. bifidum* (Turroni, Taverniti et al. 2014) while CXCL2 is downregulated by *Bifidobacterium longum* and *Lactobacillus plantarum* (Audy, Mathieu et al. 2012) *in vitro*, similar to the transcriptomic response to cBM. Upregulation of CXCL2 and CCL20 in Caco-2 intestinal cell lines has been indicated in response to *Lactobacillus salivarius* (O'Callaghan, Buttó et al. 2012), which is similar to the upregulation resulting from mBM in this study. Potentially, the combination of up and downregulated pro-inflammatory genes may indicate a balanced inflammatory effect whereby, the upregulation of some genes may allow them to be poised for a protective immune response upon infection while the downregulation of other genes may prevent an unwanted inflammatory response in the absence of infection. A reduction in IRF9 expression is also induced by *Bifidobacterium infantis* and *Lactobacillus helveticus* (Audy, Mathieu et al. 2012) which was also evident in this study in response to mBM and cBM.
Chapter V

Conclusions

In the present study, we demonstrated the potential of buttermilk to exert a wide ranging transcriptional response in colonic epithelial cells. The ability of cBM and mBM to modulate the transcription of barrier function related genes together with cytokines, chemokines and anti-viral genes highlighted a potentially very important effect of buttermilk consumption in the gastrointestinal tract. For instance, downregulation of the pro-inflammatory cytokine IL-8 may be of benefit to prevent or treat pro-inflammatory responses such as those found in IBD. Although there were many similarities between the transcriptional changes observed in response to both cBM and mBM in HT-29 cells, some differences were also observed. For instance the number of DET induced by each treatment and the level of induction as evidenced by the statistical significance varied for some gene categories. The direction of regulation also differed in some cases for each sample. In particular, the upregulation of IFNβ1, CCL20 and CCL5 in response to the mBM when compared to the downregulation observed for cBM may reflect differences in the composition of each buttermilk sample. The upregulation of such genes may be counteracted by the downregulation of other pro-inflammatory genes including CXCL10 and CXCL11. Considering these results, further in vivo studies are required to determine if these buttermilk fractions have an effect on health and if the transcriptional responses evidenced here would lead to a phenotypic change in humans. Additionally, bovine milk composition may change, for instance with season or breed, and thus it may be interesting to explore such variables to determine if the transcriptomic responses evidenced in this study may differ as a result. It is known that bovine buttermilk components are beneficial to human health and this study highlights the potential of bovine buttermilk fractions in the prevention or treatment of pro-inflammatory diseases. Additionally, buttermilk generated from widely available mature milk had similar effects to its colostrum counterpart, indicating mBM may be sufficient to elicit a beneficial transcriptomic response.
Chapter V

References


Chapter V


Chapter V


secrete anti-inflammatory cytokines and inhibit proliferation of naive CD4+ cells: a potential mechanism for immune polarization in helminth infections." The Journal of Immunology 167(9): 5294-5303.


Figure 1. Protein characterisation of buttermilk samples over the course of lactation. Colostrum and mature buttermilk samples were characterised by running on a NuPage 4-12% gradient sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions. Gels were Coomassie stained for total proteins. M – molecular weight ladder. Results are representative of duplicate experiments.
Figure 2. The number of DET affected by buttermilk samples. Venn diagram demonstrating the number of differentially expressed transcripts unique to the buttermilk treatments (colostrum buttermilk and mature buttermilk) and shared by both treatments. DET were chosen based on a p-value < 0.05 and a fold change of >2 (for upregulated genes) or <-2 (for downregulated genes).
Figure 3. Quantitative reverse transcription-polymerase chain reaction (qRT-PCR) analysis of immune-related genes in HT-29 cells treated with bovine colostrum buttermilk (cBM) or bovine mature buttermilk (mBM). Candidate genes were selected following microarray analysis and screened by qRT-PCR. The ΔΔCT method was used to calculate relative changes in gene expression compared to internal reference gene expression level. Data is presented as a mean of three experiments performed on separate occasions. Error bars represent the range of possible relative expression values defined by the standard error of the delta CT. CCL5, chemokine (C-C motif) ligand 5; CCL20, chemokine (C-C motif) ligand 20; CXCL11, chemokine (C-X-C motif) ligand 11; CXCL2, chemokine (C-X-C motif) ligand 2; CXCL3, chemokine (C-X-C motif) ligand 3, CXCL10, chemokine (C-X-C motif) ligand 10; IL-8, interleukin 8; IL-17D, interleukin 17D; IL-33, interleukin 33; IL-7, interleukin 7; Stat1, signal transducer and activator of transcription 1; Stat2, Signal transducer and activator of transcription 2; Jak2, Janus kinase 2; IRF7, interferon regulatory factor 7; IRF9, interferon regulatory factor 9; IFIT1, interferon-induced protein with tetraticopeptide repeats 1; IFIT3, interferon-induced protein with tetraticopeptide repeats 3; OAS2, 2'-5'-oligoadenylate
Chapter V

synthase 2; OAS3, 2'-5'-oligoadenylate synthase 3; DDX58, DEAD (Asp-Glu-Ala-Asp) box polypeptide 58; IFNβ1, interferon beta, 1.
Table 1. Gene ontology groups affected by buttermilk samples

<table>
<thead>
<tr>
<th>Gene ontology category</th>
<th>Fold enrichment</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Colostrum</td>
<td>Mature</td>
</tr>
<tr>
<td>Behavior (GO:0007610)</td>
<td>&gt; 5</td>
<td>&gt; 5</td>
</tr>
<tr>
<td>Response to interferon-gamma (GO:0034341)</td>
<td>&gt; 5</td>
<td>&gt; 5</td>
</tr>
<tr>
<td>Cell proliferation (GO:0008283)</td>
<td>&gt; 5</td>
<td>&gt; 5</td>
</tr>
<tr>
<td>DNA replication (GO:0006260)</td>
<td>4.4</td>
<td>-</td>
</tr>
<tr>
<td>Catabolic process (GO:0009056)</td>
<td>2.68</td>
<td>-</td>
</tr>
<tr>
<td>Immune response (GO:0006955)</td>
<td>2.42</td>
<td>4.43</td>
</tr>
<tr>
<td>Immune system process (GO:0002376)</td>
<td>2</td>
<td>2.98</td>
</tr>
<tr>
<td>Nucleobase-containing compound metabolic process (GO:0006139)</td>
<td>1.43</td>
<td>-</td>
</tr>
<tr>
<td>Hemopoiesis (GO:0030097)</td>
<td>-</td>
<td>&gt; 5</td>
</tr>
<tr>
<td>Response to stimulus (GO:0050896)</td>
<td>-</td>
<td>1.91</td>
</tr>
</tbody>
</table>

-, gene ontology category not identified as being affected by the sample
Table 2. Immune related gene ontology categories affected by buttermilk samples

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of differentially expressed transcripts</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biological process</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immune system process</td>
<td>70, 59</td>
<td>1.11E-03, 1.82E-06</td>
</tr>
<tr>
<td>Immune response</td>
<td>38, 29</td>
<td>8.88E-06, 5.04E-06</td>
</tr>
<tr>
<td>Response to interferon-gamma</td>
<td>8, 9</td>
<td>7.20E-04, 5.20E-06</td>
</tr>
<tr>
<td>Cellular defense response</td>
<td>25, 19</td>
<td>8.78E-05, 8.17E-05</td>
</tr>
<tr>
<td><strong>Molecular function</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytokine activity</td>
<td>14, 14</td>
<td>8.11E-04, 9.18E-06</td>
</tr>
<tr>
<td>Receptor binding</td>
<td>39, 38</td>
<td>2.65E-02, 1.99E-05</td>
</tr>
<tr>
<td>Chemokine activity</td>
<td>8, 6</td>
<td>5.35E-05, 2.14E-04</td>
</tr>
<tr>
<td><strong>Pathway</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inflammation mediated by</td>
<td>12, 14</td>
<td>3.32E-02, 1.19E-04</td>
</tr>
<tr>
<td>chemokine and cytokine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>signaling pathway</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JAK/STAT signaling pathway</td>
<td>4, 4</td>
<td>NS, 1.60E-04</td>
</tr>
</tbody>
</table>

NS, not statistically significant
Table 3. Differential expression of genes potentially involved in maintenance of epithelial barrier

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Colostrum buttermilk</th>
<th>Mature buttermilk</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDX60</td>
<td>DEAD (Asp-Glu-Ala-Asp) box polypeptide 60</td>
<td>-8.115127</td>
<td>-7.0339236</td>
</tr>
<tr>
<td>IL15RA</td>
<td>Interleukin 15 receptor, alpha</td>
<td>-3.8939435</td>
<td>-</td>
</tr>
<tr>
<td>PCDH11X</td>
<td>Protocadherin 11 X linked</td>
<td>7.8738313</td>
<td>-</td>
</tr>
<tr>
<td>PCDHGB7</td>
<td>Protocadherin gamma subfamily B7</td>
<td>2.7451339</td>
<td>-</td>
</tr>
<tr>
<td>PCDHGC4</td>
<td>Protocadherin gamma subfamily C4</td>
<td>-</td>
<td>3.1393077</td>
</tr>
<tr>
<td>PCDHA11</td>
<td>Protocadherin alpha 11</td>
<td>-2.0898988</td>
<td>-2.4319901</td>
</tr>
<tr>
<td>CDH16</td>
<td>Cadherin 16 (KSP-cadherin)</td>
<td>-</td>
<td>2.2631629</td>
</tr>
<tr>
<td>FAT2</td>
<td>FAT tumor suppressor homolog 2 (Drosophila)</td>
<td>-</td>
<td>3.5042636</td>
</tr>
<tr>
<td>DEFB1</td>
<td>Defensin, beta 1</td>
<td>-</td>
<td>2.1727304</td>
</tr>
<tr>
<td>MUC5AC</td>
<td>Mucin 5AC, oligomeric mucus/gel-forming</td>
<td>2.8463552</td>
<td>2.2013552</td>
</tr>
<tr>
<td>MUC13</td>
<td>Mucin 13, cell surface associated</td>
<td>-2.4739103</td>
<td>-2.409975</td>
</tr>
</tbody>
</table>

-, gene not identified as being differentially expressed in response to the sample
Table 4. Differential expression of immune associated genes identified by RT-PCR

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Colostrum buttermilk</th>
<th>Mature buttermilk</th>
</tr>
</thead>
<tbody>
<tr>
<td>CXCL2</td>
<td>Chemokine (C-X-C motif) ligand 2</td>
<td>0.477420802</td>
<td>1.03049202</td>
</tr>
<tr>
<td>CXCL3</td>
<td>Chemokine (C-X-C motif) ligand 3</td>
<td>0.571701243</td>
<td>1.172834949</td>
</tr>
<tr>
<td>CXCL10</td>
<td>Chemokine (C-X-C motif) ligand 10</td>
<td>0.20877198</td>
<td>0.897095409</td>
</tr>
<tr>
<td>CXCL11</td>
<td>Chemokine (C-X-C motif) ligand 11</td>
<td>0.171150016</td>
<td>0.543367431</td>
</tr>
<tr>
<td>CCL5</td>
<td>Chemokine (C-C motif) ligand 5</td>
<td>0.358488812</td>
<td>1.443929196</td>
</tr>
<tr>
<td>CCL20</td>
<td>Chemokine (C-C motif) ligand 20</td>
<td>0.679871186</td>
<td>2.422785474</td>
</tr>
<tr>
<td>IFNβ1</td>
<td>Interferon beta 1</td>
<td>0.667419927</td>
<td>4.277199994</td>
</tr>
<tr>
<td>IL-7</td>
<td>Interleukin 7</td>
<td>1.274560627</td>
<td>0.326842312</td>
</tr>
<tr>
<td>IL-8</td>
<td>Interleukin 8</td>
<td>0.281264621</td>
<td>0.559935802</td>
</tr>
<tr>
<td>IL-17D</td>
<td>Interleukin 17D</td>
<td>0.842841545</td>
<td>0.801069878</td>
</tr>
<tr>
<td>IL-33</td>
<td>Interleukin 33</td>
<td>0.2049485</td>
<td>0.070153878</td>
</tr>
<tr>
<td>STAT1</td>
<td>Signal transducer and activator of transcription 1</td>
<td>0.400534939</td>
<td>0.420448208</td>
</tr>
<tr>
<td>STAT2</td>
<td>Signal transducer and activator of transcription 2</td>
<td>0.655196702</td>
<td>0.718636109</td>
</tr>
<tr>
<td>JAK2</td>
<td>Janus kinase 2</td>
<td>1.156688184</td>
<td>1.094293701</td>
</tr>
</tbody>
</table>
### Chapter V

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>Value1</th>
<th>Value2</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRF7</td>
<td>Interferon regulatory factor 7</td>
<td>0.413702811</td>
<td>0.716977624</td>
</tr>
<tr>
<td>IRF9</td>
<td>Interferon regulatory factor 9</td>
<td>0.670511199</td>
<td>0.770215111</td>
</tr>
<tr>
<td>IFIT1</td>
<td>Interferon-induced protein with tetratricopeptide repeats 1</td>
<td>0.186856156</td>
<td>0.860551437</td>
</tr>
<tr>
<td>IFIT3</td>
<td>Interferon-induced protein with tetratricopeptide repeats 3</td>
<td>0.214146005</td>
<td>0.585064127</td>
</tr>
<tr>
<td>OAS2</td>
<td>2'-5' oligoadenylate synthase 2</td>
<td>0.069508505</td>
<td>0.109070536</td>
</tr>
<tr>
<td>OAS3</td>
<td>2'-5' oligoadenylate synthase 3</td>
<td>0.280615512</td>
<td>0.30495466</td>
</tr>
<tr>
<td>DDX58</td>
<td>DEAD (Asp-Glu-Ala-Asp) box polypeptide 58</td>
<td>0.266092546</td>
<td>0.51524601</td>
</tr>
</tbody>
</table>
General summary and discussion
The biological properties of human milk oligosaccharides and glycoconjugates have gained much interest in recent years (Morrow, Ruiz-Palacios et al. 2004, Bode 2006, Bode 2009, Sela and Mills 2010, Bode 2012, Manthey, Autran et al. 2013, Peterson, Cheah et al. 2013). However, since it is unlikely that glycans will be isolated from human milk for commercial use due to its limited availability, research began to focus on domestic animal milks as a source of bioactive glycans (Urashima, Saito et al. 2001, Zivkovik and Barile 2001, Sprong, Hulstein et al. 2012, Lane, O’Callaghan et al. 2013). In particular, many biological activities have been associated with various individual glycoconjugates from domestic animal milks including bovine MFGM (Urashima, Taufik et al. 2013). It is well established that glycosylation of milk components plays a role in the protection and development of the infant during the course of breastfeeding (Morrow, Ruiz-Palacios et al. 2004, Lane, O’Callaghan et al. 2013). It is possible that the use of buttermilk or whole MFGM may help narrow the gap between infant formula and breast milk in terms of composition and health promoting activities. Indeed, the use of whole MFGM in the prevention of infection and diarrhoea has been demonstrated (Timby, Hernell et al. 2015). Furthermore, Timby et al. (2014) highlighted that infants who were fed infant formula supplemented with bovine MFGM (Lacprodan MFGM-10, Arla Foods Ingredients) performed better in cognitive testing at 12 months of age than infants fed unsupplemented formula. This indicates the potential use of MFGM for the development of infant cognitive function. Additionally, the bioactivities of bovine buttermilk, including its immune-modulating (Zhou, Gao et al. 2010, Zanabria, Tellez et al. 2014) and anti-infective (Wang, Hirmo et al. 2001, Parker, Sando et al. 2010, Ross, Lane et al. 2016) activities, suggest a possible application of this dairy fraction in foods aimed at promoting health in individuals who are immuno-compromised such as infants and the elderly. The use of buttermilk as a functional ingredient to supplement other treatments may be a cheap and beneficial dairy-derived alternative to immune suppressing antibiotics. The use of such functional ingredients on a regular basis may reduce the need for antibiotics which in turn
ensures that beneficial commensal populations present in the gut will not be harmed. Indeed, human milk oligosaccharides are known to increase the adhesion of certain commensal bacterial strains to human intestinal epithelial cells \textit{in vitro} (Kavanaugh, O’Callaghan et al. 2013). These beneficial bacteria are also known to utilize milk oligosaccharides as a carbon source in order to increase their growth and promote their establishment in the gut (Asakuma, Hatakeyama et al. 2011). Glycosylated bovine MFGM may also provide similar beneficial effects and studies to investigate this are ongoing in our lab. Indeed, a patent exists to combine probiotics and MFGM to enhance the bioactivities of the probiotics (WO 2011069987 A1) (Benyacoub, Blum-Sperisen et al. 2012). Moreover, a study conducted by Favre et al. (2011) demonstrated that administration of a combination of MFGM and \textit{Bifidobacterium lactis} increased the number of mucosal IgA-secreting cells in a mouse model when compared to the administration of MFGM or \textit{B. lactis} alone. Interestingly, similar effects were still evident 12 weeks after the supplementation period had ended. This indicates the ability of probiotics in the presence of MFGM to promote the maturation of the immune system. Furthermore human MFGM components have been linked to preventing infection (Liu, Yu et al. 2012). Such components are thought to act as decoy receptors and reduce the threat of infection in infants during the course of breastfeeding (Morrow, Ruiz-Palacios et al. 2004). Interestingly, bovine MFGM-derived compounds have also been shown to play a role in the prevention of pathogenic infection \textit{in vitro} and \textit{in vivo} (Inagaki, Nagai et al. 2010, Parker, Sando et al. 2010). Bovine MFGM therefore represents an attractive source of food-derived decoy receptors which may have potential health benefits if used as prophylactics thereby reducing the requirement for antibiotic treatment in susceptible consumers. Bovine buttermilk, a by-product of butter-manufacture, is an ideal commercial source of MFGM as it is available in large quantities in the dairy industry. In 2011, 544,000 tonnes of drinking milk and buttermilk, 149,000 tonnes butter and 21,000 tonnes of cream were produced in Ireland (http://cso.ie/en/media/csoie/releasespublications/documents/agriculture/2011/
General summary and discussion

milksup_2011.pdf), highlighting the wide commercial availability of this dairy fraction. Due to the many bioactivities associated with buttermilk MFGM, buttermilk could be transformed from a low-value by-product of butter-manufacture to a novel high value product in the dairy industry through its use as a functional ingredient.

Many of the anti-infective studies performed to date have investigated purified MFGM components which can be both costly and time consuming to produce. Added to that, bacterial colonization seems to be a multifactorial event in most cases and it is probable that a cocktail of glycans with anti-infective capabilities may be the best solution. Such an approach might target several colonization factors while also targeting secreted bacterial products such as toxins. Therefore, from an economical point of view, using a mixture of glycoconjugates such as that found in MFGM is more cost effective. Considering this rationale, Chapter II focused on screening a protein/glycoprotein-enriched MFGM fraction in order to examine its anti-infective activities against *E. coli* O157:H7. *E. coli* O157:H7 was chosen due to the large numbers of outbreaks that occur worldwide. Diarrhoea is a major cause of malnutrition and mortality in developing countries and *E. coli* is one of the most common pathogenic etiological agents (WHO, 2009). In addition, *E. coli* outbreaks, often a result of food borne diseases, are highly prevalent in developed countries, leading to high public healthcare costs (Frenzen, Drake et al. 2005). In particular, EHEC serogroup O157:H7 can lead to serious illnesses (Sang, Saidi et al. 1996, Wong, Mooney et al. 2012) and even death. The severity of O157:H7 outbreaks in the USA in recent years have been summarised in Table 1.

Although a defatted MFGM fraction was used in Chapter II, to date, no delipidation methods are available that are scalable for food production and thus commercial scale up of a defatted MFGM glycoprotein/protein fraction is dependent on development of a new delipidation method. However, this may not be an issue as the lipids of the MFGM have also been shown to possess
potential health promoting activities including anti-infective properties (Iskarpatyoti, Morse et al. 2012, Salcedo, Barbera et al. 2013) and the ability to neutralize bacterial toxins (Wada, Hasegawa et al. 2010). A non-delipidated fraction may offer even further therapeutic effects although this requires investigation. It has been shown that the use of whole bovine MFGM as an infant formula supplement decreases the incidence of infection, including acute otitis media, in infants (Timby, Hernell et al. 2015), highlighting a potential use for MFGM in the promotion of health in formula-fed infants. Interestingly, the authors also noted that infants fed the MFGM-supplemented formula required significantly less use of antipyretics during the trial period than infants fed standard formula. Moreover, in their intact form as part of whole (non-defatted) MFGM, the glycoproteins are protected from digestion, potentially as a result of the steric shielding effect due to the presence of lipids (Le, Van de Wiele et al. 2012), thereby suggesting that their anti-infective activity is retained during gastric transit. In addition, a washed MFGM fraction was used in Chapter II to generate the dMFGM fraction. This wash step removed skim milk proteins including the caseins and whey which can be found in unwashed buttermilk (Dewettinck, Rombaut et al. 2008). Caseins also possess anti-infective activity, such as that seen for κ-casein against human rotavirus (Inagaki, Muranishi et al. 2014), thus an unwashed fraction, which would be even more commercially viable to produce, may also retain the anti-infective activity and possibly even broaden the pathogen range that can be targeted. However, it should be noted that the washing and delipidation steps used in Chapter II led to an enrichment of glycoproteins, and so an unwashed non-delipidated fraction may decrease the overall concentration of the glycoproteins. Further investigation would certainly be of interest to determine the impact whole MFGM may have on the bioactivities of the fraction.

Since it is widely accepted that glycans mainly act as decoy receptors in prevention of pathogenic infection, it would be interesting to determine if the glycans were in fact, the active portion of the dMFGM fraction. Treating dMFGM with various enzymes prior to incubation with the bacterial strains
could potentially confirm the role glycans play in the anti-infective activity identified for dMFGM. Enzymes including galactosidase (to remove terminal galactose residues), sialidase (for sialic acid removal) and peptide -N-Glycosidase F and A (to cleave N-linked glycans) may facilitate the identification of the glycans involved in imparting the bioactivity. Alternatively, treatment with pronase, denaturing the protein component of the dMFGM fraction and leaving all glycosylation intact, may identify if the carbohydrate portion alone was responsible for the anti-infective activity or if the linkage to the protein was also a determining factor.

Given the importance of glycosylation for MFGM bioactivities, it is interesting to note that the use of milk from a particular lactation stage may influence the glycosylation status of the buttermilk fraction. Mammalian colostrum is known to be a source of highly glycosylated ingredients, with the abundance of N-linked glycans decorated with sialic acid and fucose (Takimori, Shimaoka et al. 2011) and O-linked core 2 oligosaccharides (Wilson, Robinson et al. 2008) decreasing in the progression from colostrum to mature milk. This highlights the potential important role glycosylation plays in the infant gut. As a source of highly glycosylated ingredients, colostrum provides protection from pathogenic infection in the infant gut (Liu and Newburg 2013, Peterson, Cheah et al. 2013) and may also modulate the immature infant immune system (Lane, O'Callaghan et al. 2013). Knowledge of how bovine buttermilk glycosylation profiles change during lactation could help outline the role glycosylation plays during milk maturation (Chapter III). It may also be of importance on a commercial level, as isolating buttermilk from a particular lactational stage may be beneficial in obtaining the most bioactive pool of glycosylated ingredients. Thus, the glycosylation of bovine buttermilk was investigated through the use of lectin microarrays.

Factors which may influence glycosylation of buttermilk include season, feeding type (e.g. grass-fed versus ration-fed animals) and breed. It has been shown that Holstein-Friesian and Jersey cattle display different milk
General summary and discussion

oligosaccharide profiles (Sundekilde, Barile et al. 2012) while glycosylation of porcine milk glycoproteins has been demonstrated to be breed specific (Gustafsson, Hultberg et al. 2006). Interestingly, these varying glycoforms led to differing levels of anti-infective activity against *H. pylori*. It is likely that bovine MFGM glycosylation is affected by breed also. Further investigation into how breed affects milk glycosylation patterns and the subsequent effect on the associated bioactivities of the glycoconjugates warrants further investigation.

The bioactivities of bovine buttermilk are likely to change over the course of lactation due to the changes that occur in glycosylation. It would be of great interest to monitor the potential changes in bioactivities to fully understand how varying glycosylation may affect the potential health-promoting properties of bovine MFGM. This would be of further benefit at a commercial scale as isolation of buttermilk from a particular lactational stage could be targeted to impart particular bioactivities. For instance, the ability of an MFGM fraction to reduce the occurrence of *E. coli* binding to its glycan receptors could be monitored using lectin microarrays. Incubation of MFGM isolated from various stages of lactation with *E. coli*, prior to exposure to the lectin microarrays, could identify the potential glycan structures on bacterial cells which interact with the MFGM and prevent binding to the host cell receptor. Additionally, this would allow a direct comparison of MFGM isolated from colostrum, transitional and mature milks and may indicate the lactational stage from which MFGM should be sourced to most effectively act as a decoy receptor. Since lectin arrays allow for high throughput screening with minimal sample, it may also be possible to screen buttermilk from various lactation stages for the ability to reduce the adhesion of a large variety of pathogens to the lectins. This would help to identify the anti-adhesive capabilities of bovine buttermilk from various lactation stages against a wide range of pathogens.

The use of MFGM fractions as a source of glycoconjugates provides a highly glycosylated mixture of components, with a potential variety of anti-
infective decoy receptors. As stated, these have the potential to target the many
different cell surface receptors of *E. coli*, as well as other bacterial strains.
Proteins on the surface of pathogens, called adhesins, mediate the vast majority
of initial bacterial attachments to host cells and thus are the primary targets for
anti-adhesives. The specificity of the adhesins is one of the main determinants
for tissue tropism (the preferential colonization site) and host range (the species
the pathogen can colonize) (Zopf and Roth 1996). The initial adhesion of EPEC
is mediated by the adhesins bundle-forming pili and intimin α. Conversely,
initial adhesion of EHEC depends mainly on long polar fimbriae and intimin γ
(Nataro and Kaper 1998, Pereira and Giugliano 2013). However, it is important
to note that bacterial surface glycosylation is also an important factor for
pathogenesis. For instance, sialylation of *C. jejuni* lipooligosaccharide was
demonstrated to be an important contributor to invasion of human epithelial cell
lines (Louwen, Heikema et al. 2008). Furthermore, deglycosylation of the
surface glycoproteins of respiratory syncytial virus decreased viral infectivity
(Lambert 1988) while surface glycans of infectious hematopoietic necrosis
virus may be important for viral binding to fish epithelial cells (Nita-Lazar,
Mancini et al. 2016). Given the role of bacterial surface glycosylation in
pathogenesis, it would be of interest to investigate the glycosylation of *E. coli*
strains. However, most studies have focused on bacterial surface proteome in
host cell adhesion and the availability of literature investigating *E. coli* cell
surface glycosylation is limited. Furthermore, studies which have investigated
the glycosylation profiles of *E. coli* strains have not compared the profiles of
EHEC and EPEC strains, which display different tissue tropisms (Mellies,
Barron et al. 2007). Moreover, detailed comparisons of EHEC and EPEC host
receptors have not been investigated previously to the best of the authors’
knowledge. Such studies could further our understanding of bacterial adhesion
mechanisms and perhaps allow the tailoring of anti-infective compounds to
target bacterial strains of interest. In **Chapter IV**, the aim was to identify
differences in their surface glycosylation patterns and receptor binding
preferences which shed light on the localization of bacteria during infection.
Lectin microarrays and NGC microarrays were employed to investigate the potential surface glycome and receptor binding preferences of several *E. coli* strains, respectively.

Lectin microarrays have previously been used to profile bacterial glycosylation and are useful in distinguishing differences in glycosylation of similar bacterial strains (Hsu, Pilobello et al. 2006). Lectin microarrays also provide a platform in which secondary modifications, such as acetylation, can be detected (Pilobello and Mahal 2007). It was interesting to observe that the bacteria formed three distinct groups based on their surface glycosylation patterns considering that the anti-infective activity associated with the dMFGM fraction in Chapter II was active against *E. coli* NCTC 12900, DPC 6055 and DAF454, all of which are EHEC strains however it was not active against either of the EPEC groups. This could be as a result of the EPEC groups having distinct surface glycosylation and binding preferences when compared to the other EHEC strains screened, as identified by the microarrays in this study. Furthermore, the NGC microarrays highlighted the bacteria formed two distinct groups based on their binding preferences. Given that tissue tropisms differ for EHEC and EPEC, there is potential for the use of NGC microarrays to indicate the sites of bacterial colonization.

Phase variation, an adaptive process in which bacterial phenotype is altered by reversible and high frequency genetic changes, is a tool used by bacteria to avoid immune detection and survive stress conditions (Hallet 2001). Bacterial surface glycosylation can be affected by such processes. For instance, one of the two capsular glycan structures synthesized by *C. jejuni*, has been shown to be important for colonization and undergoes phase variation (Bacon, Szymanski et al. 2001). In addition, phase variation occurs for Le antigen expressed by *H. pylori* on its O-chain of LPS (Moran 2008). Changes in bacterial glycosylation due to phase variation or the stage of infection (Moran, Gupta et al. 2011) make it difficult to identify the bacterial glycome at a particular time point. It may be of value to investigate the binding of these
bacterial strains at different phase variations or infection time points to neoglycoconjugate microarrays in order to assess what changes in the bacterial cell surface glycome can occur under different conditions. This may provide invaluable information which could allow the compilation of a mixture of glycans that could increase the success of decoy receptor therapy irrespective of the stage of infection. The use of buttermilk, with its many glycosylated components, could be beneficial in such circumstances. The presence of different glycan structures in bovine buttermilk should increase the likelihood of a bacterial-MFGM interaction occurring and may also provide a variety of receptors for the ever-changing bacterial surface glycome. In addition, through profiling bacteria isolated from various sources of infection, a database could be composed based on their binding preferences. This could potentially contain information on bacteria with common host receptors which could be used as a guide to design cocktails of glycosylated ingredients to target multiple bacterial strains.

Although lectin microarrays provide a platform with which to rapidly screen for bacterial glycosylation, their use has limitations. For instance, some lectins can bind multiple glycan structures (Pilobello and Mahal 2007) making it difficult to determine the structure binding to the lectin. Furthermore, many bacterial cells contain glycans which are not found in eukaryotic cells, and so would not be identified by lectin microarrays, due to a lack of availability of lectins which recognize such sugars (Hsu, Pilobello et al. 2006, Pilobello and Mahal 2007). However, it is worth noting that there are other alternatives to lectins which could be used. For instance, peptides are functionally stable molecules which retain their activity under the majority of reaction conditions and can be used in microarray format to screen for glycans (Morales Betanzos, Gonzalez-Moa et al. 2009). Moreover, aptamers are single stranded nucleotides which have the ability to bind glycans and may provide another method to rapidly screen for the presence of pathogens (Arnaud, Audfrey and Imberty 2013).
General summary and discussion

The DNA microarray and qPCR validation experiments conducted in Chapter V provide novel insights into the role bovine buttermilk may play in immunomodulation in the gut at the transcriptional level. This may be beneficial since upregulating the expression of such genes can result in inflammatory disorders and concomitant impairment of epithelial function (Ferrier, Mazelin et al. 2003). Further studies are required to determine additional signaling pathways which can be potentially influenced by buttermilk treatments. Pathways such as the JAK/STAT pathway were found to be differentially regulated by the treatments compared to the control. However, it would be of interest to further characterize the molecular mechanisms by which buttermilk regulates these pathways in order to fully understand the potential of this fraction to modulate immune function. Additionally, since the transcriptomic response of HT-29 cells to buttermilk samples was investigated in Chapter V, it would be of interest to explore the expression of the corresponding proteins in the intestinal cells after stimulation with the buttermilk samples, such as through the use of enzyme-linked immunosorbent assays.

The benefits of bovine colostrum have been recognized by the dairy industry considering the availability of colostrum products from companies such as La Belle (Colostrum Capsules and Colostrum Chewables for dietary supplementation), APS Biogroup (bovine whole colostrum products including tablets and sachets), PerCoBa (colostrum supplements - Immune-Dynamics Inc) and Sterling Technology. ColoPlus, a bovine colostrum supplement, is known to alleviate the symptoms of HIV-associated diarrhoea in patients experiencing severe diarrhoea (Florén, Chineny et al. 2006). Although bovine colostrum has been recognized as an important health-promoting ingredient, the use of buttermilk as a source of ingredients for health supplements has not been fully realized. Since this dairy stream is already produced in such high quantities in the dairy industry, exploiting buttermilk’s beneficial properties in a similar manner seems like an obvious next step for dairy companies.
General summary and discussion

It is important to take into consideration the effect milk processing may have on the structure and bioactivities of MFGM components. Pasteurization, a process whereby milk is heated to 75 °C for 25 s to reduce microbial contamination (FSAI, 2009), is a routine procedure in the dairy processing industry. Pasteurization of milk prior to buttermilk production could alter the bioactivities of MFGM in buttermilk. Indeed, incorporation of whey proteins into MFGM can occur after heating to 80 °C and concentrations of existing MFGM glycoproteins such as Pas6/7 may be reduced (Lee and Sherbon 2002).

Homogenization, the process of mechanically breaking down fat globules in milk in order to form a homogenous solution, prior to buttermilk production may also lead to changes in the structure, and possibly the bioactivities, of the MFGM fraction. Adsorption of caseins to MFGM is observed after homogenization (Lee and Sherbon 2002) but interestingly, adsorption of skim milk proteins to the MFGM does not displace existing native MFGM components (Ye, Anema et al. 2008) and therefore it is likely that the bioactivities of the MFGM proteins and glycoproteins would be retained after the homogenization process. As discussed in Chapter II, the effect of digestion on MFGM bioactivities is an important factor to consider as the structure of MFGM during gastric and intestinal digestion is known to undergo alterations. However some glycoproteins such as mucins remain largely undigested. This is most likely due to the extensive glycosylation associated with these glycoproteins (Gallier, Ye et al. 2012, Gallier, Cui et al. 2013, Gallier, Zhu et al. 2013). Despite this, the health-promoting activities of MFGM as a whole may be affected by digestion. However, it may be possible to ensure that MFGM in its intact form (if required) reaches the site of action, by incorporating an encapsulation process, for example (Augustin and Hemar 2009, de Vos, Faas et al. 2010). Encapsulation is used in the food industry to protect, immobilize and control the release of sensitive compounds, by coating or entrapping in a matrix (Poncelet 2006). The ability to encapsulate large proteins has previously been shown. For instance, the use of polyethylene glycol hydrogels (McCall and Anseth 2012) and alginate micro gel particles
General summary and discussion

(Bokkhim, Bansal et al. 2016) have been demonstrated to be potential mechanisms to preserve protein bioactivity and stability in the food industry.

Overall, buttermilk offers great potential as a functional ingredient aimed at promoting human health and may offer an alternative approach to the more technically challenging isolation of pure MFGM and MFGM components. In addition, the presence of multiple glycosylated structures, free milk proteins and lipids may increase the bioactivity of the fraction when compared to purified MFGM. This may allow a wider array of pathogens to be targeted, as well as increasing other health benefits provided by its multiple components. Moreover, the use of buttermilk is commercially very attractive as it is produced regardless of its use as it is a by-product of butter-making. Therefore, the use of buttermilk in this manner may increase the value of milk in the dairy processing industry. This thesis supports the use of buttermilk derived from transitional or mature whole milk and negates the requirement to utilize buttermilk sourced from colostrum, which is less abundant. In addition, the transcriptional effects of mature and colostrum derived buttermilk are quite similar also, further validating the use of mature buttermilk as a widely available functional ingredient.
General summary and discussion

References


General summary and discussion


Food Safety Authority of Ireland (2009). “Mycobacterium avium subsp. paratuberculosis and the possible links to Crohn’s disease”.


General summary and discussion


General summary and discussion


General summary and discussion


General summary and discussion


Table 1. Cases of *Escherichia coli* O157:H7 infection in the USA

<table>
<thead>
<tr>
<th>Year</th>
<th>Source</th>
<th>Case count</th>
<th>Deaths</th>
<th>Hospitalisations</th>
<th>Cases of hemolytic uremic syndrome</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>Rotisserie chicken salad</td>
<td>19</td>
<td>0</td>
<td>26%</td>
<td>2</td>
</tr>
<tr>
<td>2014</td>
<td>Ground beef</td>
<td>12</td>
<td>0</td>
<td>58%</td>
<td>0</td>
</tr>
<tr>
<td>2013</td>
<td>Ready-to-eat salads</td>
<td>33</td>
<td>0</td>
<td>32%</td>
<td>2</td>
</tr>
<tr>
<td>2012</td>
<td>Pre-packaged leafy greens</td>
<td>33</td>
<td>0</td>
<td>46%</td>
<td>2</td>
</tr>
<tr>
<td>2011</td>
<td>Romaine lettuce</td>
<td>58</td>
<td>0</td>
<td>67%</td>
<td>3</td>
</tr>
<tr>
<td>2011</td>
<td>Beef product</td>
<td>14</td>
<td>0</td>
<td>23%</td>
<td>0</td>
</tr>
<tr>
<td>2011</td>
<td>Hazelnuts</td>
<td>8</td>
<td>0</td>
<td>50%</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>Gouda cheese</td>
<td>38</td>
<td>0</td>
<td>39%</td>
<td>1</td>
</tr>
<tr>
<td>2010</td>
<td>Beef product</td>
<td>21</td>
<td>0</td>
<td>43%</td>
<td>1</td>
</tr>
<tr>
<td>2009</td>
<td>Beef</td>
<td>26</td>
<td>2</td>
<td>73%</td>
<td>5</td>
</tr>
<tr>
<td>2009</td>
<td>Ground beef</td>
<td>23</td>
<td>0</td>
<td>70%</td>
<td>2</td>
</tr>
<tr>
<td>2009</td>
<td>Pre-packaged cookie dough</td>
<td>72</td>
<td>0</td>
<td>47%</td>
<td>10</td>
</tr>
<tr>
<td>2008</td>
<td>Ground beef</td>
<td>49</td>
<td>0</td>
<td>55%</td>
<td>1</td>
</tr>
<tr>
<td>2007</td>
<td>Pepperoni (pizza)</td>
<td>21</td>
<td>0</td>
<td>38%</td>
<td>4</td>
</tr>
<tr>
<td>2007</td>
<td>Ground beef</td>
<td>40</td>
<td>0</td>
<td>64%</td>
<td>2</td>
</tr>
<tr>
<td>2006</td>
<td>Taco bell</td>
<td>71</td>
<td>0</td>
<td>75%</td>
<td>8</td>
</tr>
<tr>
<td>2006</td>
<td>Spinach</td>
<td>199</td>
<td>3</td>
<td>51%</td>
<td>31</td>
</tr>
</tbody>
</table>

Source: Centers for Disease Control and Prevention