Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

_____ _____

Alexander J. Noll Date

Human Milk Glycan Interactions with Glycan-Binding Proteins of the Gastrointestinal Tract

By

Alexander J. Noll
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Microbiology and Molecular Genetics

Richard D. Cummings, Ph.D. Advisor	William M. Shafer Committee Member
Graeme L. Conn, Ph.D. Committee Member	David A. Steinhauer, Ph.D. Committee Member
Joanna B. Goldberg, Ph.D. Committee Member	David S. Weiss, Ph.D. Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Human Milk Glycan Interactions with Glycan-Binding Proteins of the Gastrointestinal Tract

By

Alexander J. Noll

B. S., Molecular Biology and Microbiology, University of Central Florida, 2011

Advisor: Richard D. Cummings, Ph.D.

An abstract of a dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

Graduate Division of Biological and Biomedical Sciences

Microbiology and Molecular Genetics

2016

Human Milk Glycan Interactions with Glycan-Binding Proteins of the Gastrointestinal Tract

By Alexander J. Noll

Human milk glycans (HMGs) represent one of the major biomolecules found in human milk. These non-nutritional carbohydrates are strongly associated with health benefits towards newborns and infants, including regulation of the gut microbiome and reducing the incidence of infection. More recently, HMGs have been shown to regulate immune responses and intestinal epithelial cell gene expression, but the underlying mechanisms are unclear. The human gastrointestinal (GI) tract, specifically epithelial cells and dendritic cells, express a number of glycan-binding proteins (GBPs) that may become exposed to HMGs during breast-feeding. GBPs, including the C-type lectin, Siglec, and galectin families of GBPs, are known to bind endogenous and exogenous glycan structures and subsequently initiate signaling pathways leading to changes in gene expression and immune responses. This thesis work was guided by the hypothesis that HMGs bind to some GBPs expressed by epithelial cells and dendritic cells (DCs) in the GI tract. To test this hypothesis, a number of galectins, C-type lectins, and Siglecs expressed by GI tract epithelial cells and DCs were screened for HMG binding. This study took advantage of glycan microarray technology to test binding to over 240 natural HMGs purified from human milk as well as chemically defined HMG and non-HMG glycans in a high-throughput format. The binding of GBPs to free HMGs in solution was also used to confirm GBP-HMG interactions in a more natural setting. The results of this study show that many galectins bind specific HMG structures, with each galectin exhibiting a unique binding specificity. Additionally, the DC-expressed C-type lectin DC-SIGN specifically binds numerous fucosylated HMGs. The HMG microarray technology also uncovered hitherto undefined glycan binding structures and determinants for galectins and DC-SIGN. The affinity of these GBPs for specific HMG structures occurs at or below HMG concentrations found in human milk, suggesting that these interactions may be physiologically relevant. The results of this study suggest that multiple GBPs in the GI tract may serve as HMG receptors. Future directions are aimed at understanding if these GBP-HMG interactions occur *in vivo* and if these interactions underlie the mechanism of regulation of gene expression and immune responses by HMGs.

Human Milk Glycan Interactions with Glycan-Binding Proteins of the Gastrointestinal Tract

By

Alexander J. Noll

B. S., Molecular Biology and Microbiology, University of Central Florida, 2011

Advisor: Richard D. Cummings, Ph.D.

A dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

Graduate Division of Biological and Biomedical Sciences

Microbiology and Molecular Genetics

2016

Acknowledgements

First of all, I would like to thank Richard Cummings for all his time and attention and letting me perform my graduate studies in his laboratory. Richard and his wife, Sandy Cummings, took me in as one of their own and helped to realize potential I didn't even know I possessed. For the training, development, and friendship they offered me, I am eternally grateful.

I would also like to thank all the members of my thesis committee. Graeme, Joanna, Dave, Bill, and Dave were all so eager to help me grow and develop as a scientist and person, even going as far as to drop whatever they were doing to focus on my development. They all taught me how to focus and set my goals and to never sell myself short. Thank you all for going above and beyond your ways as mentors, colleagues, and friends.

Many fellow co-workers, colleagues, and friends have also been there for me along the way. Jean-Phillipe Gourdine, Jamie Heimburg-Molinaro, Melinda Hanes, and Tanya McKitrick have been especially great friends, colleagues, and teachers. All of my friends and colleagues in the MMG program have also been very inspirational and helped challenge me to be beyond the best I can be. These are all people that I hope will be life-long friends and colleagues.

Lastly but most importantly, I would like to thank my family. They have been with me from the start and helped me develop into the person I am today. They taught me that education, science, and life in general is not just about brains but also devotion, self-sacrifice, and a strong work ethic. I would not have made it to this point in my life if it wasn't for their love and support.

Table of Contents

Figures and Tables

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

Chapter 1: Introduction

I. Human Milk Glycans

i. Background: Human Milk and Infant Health

Human milk is the ideal source of nutrition and health for neonates and infants as well as a vital component of the infant's diet up until about 2 years of age [1]. Compared to exclusive or predominant formula feeding, exclusive or predominant breast-feeding for 6 months is associated with multiple benefits including reduced risk of respiratory infections, gastrointestinal diseases, and mortality [2]. Breast-feeding even has long-term health benefits for infants including reduced risk of obesity and Type-2 Diabetes, reduced blood pressure, and improved school performance [3, 4]. A recent systematic study by the WHO also suggested a causal link between breast-feeding and a slight increase in intelligence test scores, including IQ scores [4]. For these and other reasons, human milk is regarded as the "gold standard" of infant health and nutrition [5], and the WHO recommends exclusive breast-feeding for the first 6 months of life [1].

Human milk not only provides nutrients (carbohydrates, proteins, fat, vitamins, and minerals) to newborns and infants but also other bioactive factors that promote proper immunity, physiology, and development. These bioactive factors include whole white blood cells, growth factors (ex-epidermal growth factor) and other hormones, cytokines (ex-TGF-β, IL-10), bile-salt stimulated lipase for more efficient fatty acid catabolism, innate immunity proteins (ex-lactoferrin, lysozyme), antibodies (mostly secretory IgA), essential fatty acids, glycoconjugates, and oligosaccharides (glycans) (reviewed in [1, 5, 6]). The quality and/or quantity of many of these bioactive factors in other nutrient sources such as fortified foodstuffs and infant formula do not match what

occurs in breast milk, which may at least partially explain the short- and long-term health benefits of breast-feeding. Of these bioactive factors, glycans of more complex structures than the nutritional disaccharide lactose are the major topic of this study and a topic of major interest due to their very high abundance in human milk and known roles in infant immunity, physiology, and development.

ii. Structural Features of HMGs

A non-lactose carbohydrate fraction in human milk has been known to exist for over 100 years, although it was not until the 1950's that the actual structures (and functions) of the carbohydrates, now termed human milk oligosaccharides (HMOs) or human milk glycans (HMGs), began to be elucidated (reviewed in [7]). Human milk is now known to contain over 200 different HMG structures [8], although ~15-20 structures make up the vast majority of HMGs by both mass and molar concentration. Excellent reviews on HMG structures and biosynthesis are available [9, 10], and a general overview of HMG structures and biosynthesis are shown in **Table 1** and **Figure 1**. The monosaccharide units of HMGs consist of D-glucose (Glc), D-galactose (Gal), D-N-acetylglucosamine (GlcNAc), L-fucose (Fuc), and 5-N-acetylneuraminic acid (Neu5Ac, a sialic acid), although glucose is only found at the reducing end of glycans. Other monosaccharides found in humans such as D-N-acetylgalactosamine (GalNAc), D-mannose (Man), D-xylose (Xyl), D-glucuronic acid (GlcA), and L-iduronic acid (IdoA) are not found within HMGs. HMGs are all biosynthesized from and contain lactose (Galβ1-4Glc), the most abundant carbohydrate structure in human milk (~40-70g/L), at the reducing end. The lactose can be modified with Fuc and/or Neu5Ac or extended with

Galβ1-3/4GlcNAc (N-acetyllactosamine, LacNAc) to create linear and/or branched structures. Three different isomers or linkages of LacNAc can be added to HMGs: Type I N-acetyllactosamine (Galβ1-3GlcNAcβ1-3; LacNAc I), Type II N-acetyllactosamine (Galβ1-4GlcNAcβ1-3; LacNAc II), or a branched Type II N-acetyllactosamine unit (Galβ1-4GlcNAcβ1-6). The following modifications at the non-reducing end lactose or LacNAc units prevent further extension from the non-reducing end Gal with LacNAc I or LacNAc II: α1-2-linked Fuc to Gal, α1-4-linked linked Fuc to -GlcNAc-, α1-3-linked Fuc to Glc on lactose, α2-3-linked Neu5Ac to Gal, or α2-6-linked Neu5Ac to Galβ1-3- or -β1-3GlcNAc-. The addition of LacNAc I to Galβ1-4- at the non-reducing end prohibits additions of other LacNAc units on that branch, but does not inhibit other "terminating" modifications from being added to it. For example, LacNAc I can be added to lactose to form lacto-N-tetraose (LNT; Galβ1-3GlcNAcβ1-3-Galβ1-4Glc), which can then be modified by the addition of α1-2-linked Fuc to form lacto-N-fucopentaose I (LNFPI; Fucα1-2 Galβ1-3GlcNAcβ1-3-Galβ1-4Glc) or both α1-2- and α1-4-linked Fuc to form lacto-N-difucohexaose I (LNDFH I; Fucα1-2 Galβ1-3(Fucα1-4)GlcNAcβ1-3-Galβ1-4Glc).

In contrast to α1-2- and α1-4-linked Fuc, structures have been identified containing α1-3-linked Fuc on GlcNAc residues that are not part of the non-reducing end LacNAc II unit. Thus, α1-3-linked Fuc does not always act as a terminating modification. However, it is unclear whether or not α1-3-fucosylation of GlcNAc occurs at the non-reducing end GlcNAc all the time (as is the case for LNFPIII) or if the enzyme can also modify internal GlcNAc residues. If the former is the case, this would suggest that α1-3-

fucosylation of the non-reducing LacNAc II unit is not a terminating modification whereas, if the latter is true, it may mean that α1-3-fucosylation of the non-reducing end LacNAc II but not internal LacNAc II units is a terminating modification. A major reason for this lack of knowledge is because the α1-3-fucosyltransferase responsible for this modification is unclear (described in more detail below), so there is currently no definitely way to study the substrate specificity and thus biosynthetic route leading to HMG structures carrying α1-3-fucose.

The addition of a LacNAc II unit to the non-reducing end allows further extension of the HMG with LacNAc, making this modification an "extendable" modification. A branching LacNAc unit can also be added to structures containing at least one LacNAc unit, such as LNT or lacto-N-neotetraose (LNnT; Galβ1-4GlcNAcβ1-3-Galβ1-4Glc) but not lactose. Like LacNAc II, the branching LacNAc unit is an extendable modification. In the case of HMGs containing a non-reducing LacNAc I unit, the addition allows further extension of the HMG structure, but only from the Galβ1-4GlcNAcβ1-6- branch. For example, the addition of branching LacNAc to LNT forms lacto-N-hexaose (LNH; Galβ1-3-GlcNAcβ1-3(Galβ1-4GlcNAcβ1-6)Galβ1-4Glc); only the Galβ1-4GlcNcβ1-6-branch can be further extended with LacNAc (**Figure 1**), although both branches can be modified with α-linked Fuc and/or α-linked Neu5Ac.

As a result, three different general biosynthetic routes can be described for HMGs (**Figure 1**). Starting from lactose, the lactose can be modified with:

1) α1-2-linked Fuc, α1-3-linked Fuc, α2-3-linked Neu5Ac, or α2-6-linked Neu5Ac, preventing further extension of the lactose unit with LacNAc. Additionally, both α1-2-linked and α1-3-linked Fuc can be added to lactose, forming lactodifucotetraose

(LDFT; Fucα1-2Galβ1-4(Fucα1-3(Glc); whose structure is not shown in **Figure 1**). In the case of colostrum (the concentrated "first milk" from the lactating mother), β-linked Gal can be added instead of Fuc or Neu5Ac to form 3'-, 4'-, or 6'-galactosyllactose. Essentially, this path leads to a modified lactose structure lacking LacNAc.

2) LacNAc I, forming LNT. This structure can be modified with α-linked Fuc and/or α-linked Neu5Ac, but only be further extended if branching LacNAc unit is added, and then only the Galβ1-4GlcNAcβ1-6- branch can be further extended with LacNAc. Thus, this path will only lead to structures carrying >1 LacNAc unit if the branching LacNAc unit is added.
3) LacNAc II, forming LNnT. This structure can be further extended with LacNAc and/or branched with the extendable branching LacNAc unit. Neu5Ac in α2-6-linkage at the non-reducing end LacNAc unit can also occur, thereby terminating branch extension. Fucose addition to GlcNAc in α1-3-linkage may also occur, although whether or not this modification to the non-reducing end LacNAc unit is a terminating modification is unclear. This path that can lead to both branched and linear (non-branched) structures containing >1 LacNAc unit.

It is remarkable that the specific glycosyltransferases responsible for the biosynthesis of all HMGs beyond lactose are unclear. It is believed that most of these glycosyltransferases are enzymes that are currently known, and genetic evidence discussed below is consistent with this, but some enzymes have not yet been defined either genetically or biochemically. For example, two enzymes are known, mainly because of the presence of a large number of individuals in the human population lacking

these enzymes and thus the modifications these enzymes produce. These enzymes are the FUT2 (Secretor) enzyme responsible for the addition of α1-2-linked Fuc to LacNAc I units or lactose to form 2'-fucosyllactose (2'-FL) and the FUT3 (Lewis) enzyme responsible for the addition of α1-4-linked Fuc to HMGs and α1-3-linked Fuc to 2'-FL to form lacto-difucotetraose (LDFT), as described in more detail below. One modification, the addition of α2-6-linked Neu5Ac to GlcNAc, is a relatively unique modification that rarely occurs on other human glycoconjugates and hence has received little study. As a result, the identity of this glycosyltransferase is unclear and represents an example of an unidentified glycosyltransferase specifically involved in HMG biosynthesis. For these reasons, the biosynthetic route of HMGs including those depicted in **Figure 1**, are primarily based on knowledge of known HMG structures.

Despite the relatively large number of HMGs (>200) in human milk, the structures contain only a handful of determinants, which are defined as the "functional" features of a glycan, especially based on their ability to interact with glycan-binding proteins (GBPs), antibodies, and other binding partners through the recognition of 2-5 monosaccharide residues [11]. Some of the major glycan determinants found on HMGs include LacNAc I, LacNAc II, the *Sambucus nigra* lectin (SNA) binding determinant (Neu5Acα2-6Galβ1-4GlcNAcβ-), Blood Group H Type I (Fucα1-2Galβ1-3GlcNAcβ-), Lewis a (Galβ1-3(Fucα1-4)GlcNAcβ-), Lewis b (Fucα1-2Galβ1-3(Fucα1-4)GlcNAcβ-), Lewis x (Galβ1-4(Fucα1-3)GlcNAcβ-), Lewis y (Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ-), sialyl Lewis a (Neu5Acα2-3Galβ1-3(Fucα1-4)GlcNAcβ-), and the so-called "sialyl Lewis c" determinant (Neu5AcGalβ1-3GlcNAc-) [8]. One of the most striking observations is the predominance of HMG structures containing LacNAc I and/or α-

linked fucose, including the Lewis a and Lewis b determinants [8, 12]. These features are highly unique to human milk glycans compared to milk glycans from other animals, suggesting an evolutionary adaptation specific for humans [12, 13]. Interestingly, HMGs do not express the A or B blood group antigens [14], although as discussed below, they can express the O(H) antigen.

The human milk glycome beyond lactose and derived only from the lactating mammary gland is remarkably unique compared to the milk glycomes of other mammals, even different primates, suggesting a strong evolutionary adaptation [15]. The term "glycome" is used to denote a constellation of glycans synthesized by a specific organism, organ, tissue or cell. The human milk glycome is especially and uniquely rich in both LacNAc I determinants and fucosylated determinants, contains significantly more structures, and has a much higher glycan concentration than in other mammalian milk samples (reviewed in [12, 16]). These features are especially true when comparing the human and bovine milk glycomes, which share a few similarities but even more differences. Human and bovine milk contain the sialylated milk glycans 3'-SL and 6'-SL, whereas bovine but not human milk contains N-glycolylneuramininic acid (Neu5Gc, a sialic acid) in place of Neu5Ac on some acidic milk glycans. Additionally, there is a nearly complete absence of fucosylated glycans and LacNAc I determinants in bovine milk along the presence of unique structures not seen in human milk such as structures containing LacNAc II instead of lactose at the reducing end. Finally, the glycan concentration and number of structures in bovine milk is substantially lower than in human milk [17]. Most importantly, even if adjusted for concentration differences, bovine milk glycans promote different functional effects on human cells than HMGs such

as differences in gene expression by intestinal epithelial cells [18]. For these reasons, bovine milk glycans are far from ideal substitutes for HMGs in infant formula.

Another interesting structural feature of HMGs is the structural diversity within and between mothers, especially diversity arising from genetic status. It has been shown that genotype affects the HMG structures and HMG concentrations in human milk. The major genetic variances seen in the human population that affect HMG structure and concentration are the *Secretor* and *Lewis* genes. The *Secretor* gene encodes the Secretor protein FUT2, an α1-2 fucosyltransferase responsible for the addition of α1-2 fucose preferentially onto Type I LacNAc structures (Galβ1-3GlcNAc-) to generate the Blood Group H Type 1 determinant (Fucα1-2Galβ1-3GlcNAc) [19]. In mothers lacking a functional copy of the *Secretor* gene (Secretor-negative), HMGs containing the Blood Group H Type 1 determinant, including lacto-N-fucopentaose I (LNFPI; Fucα1-2Galβ1-3GlcNAcβ1-3Galβ1-4Glc), are absent in their milk [20]. Additionally, Secretor-negative milk lacks 2'-fucosyllactose (2'-FL; Fucα1-2Galβ1-4Glc), the most abundant HMG in human milk [12, 20]. The prevalence of Secretor-negative mothers can range from 1-20% depending on the population [20-22].

The *Lewis* gene encodes another fucosyltransferase, the FUT3 fucosyltransferase, responsible for the addition of α1-3 and α1-4-linked fucose to GlcNAc, as in Galβ1-3GlcNAc- determinants to form the Lewis a glycan determinant (Galβ1-3(Fucα1-4)GlcNAc), and α1-3-linked fucose to 2'-FL to form lacto-di-fucotetraose (LDFT; Fucα1-2Galβ1-4(Fucα1-3)Glc) [23]. Of historical significance, this FUT3 enzyme responsible for the Lewis phenotype was actually first purified from human milk to biochemically define the specific substrates of this enzyme [23, 24]. In mothers lacking a

functional copy of the *Lewis* gene, LDFT and α1-4-fucosylated HMGs such as lacto-N-fucopentaose II (LNFPII; Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4Glc) are absent [20]. Similarly to the Secretor-negative phenotype, the Lewis-negative phenotype prevalence varies by population, with a range of about 5-25% [20]. [20]. It is noteworthy that the Lewis enzyme FUT3 is the only enzyme in humans capable of synthesizing two different glycan linkages, as it can generate both Fucα1-3GlcNAc and Fucα1-4GlcNAc modifications.

A small percentage of individuals lack functional copies of both the *Secretor* gene and *Lewis* gene. In these individuals, Blood Group H Type 1 and Lewis a determinants as well as 2'-FL are absent in milk. Fucosylated HMGs are still present though, specifically structures carrying α1-3-linked fucose such as 3-fucosyllactose (3-FL; Galβ1-4(Fucα1-3Glc) and lacto-N-fucopentaose III (LNFPIII; Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4Glc). This rare HMG genotype only makes up 1% or less of the population, at least within European populations [20].

To date, no milk samples devoid of α1-3-fucoslyated HMGs have been identified, but the fucosyltransferase responsible for this addition is unclear due to the lack of a genetic marker. Other known α1-3 fucosyltransferases (FUT4, FUT5, and FUT6) capable of generating such α1-3-fucoslyated glycans based on previous activity studies on these enzymes [25], although the expression of these enzymes in mammary tissue has not been thoroughly studied. Additionally, no studies to date have described variability in expression of other HMG glycan determinants besides α1-2- and α1-4-fucosylated HMGs.

Thus, Thurl et al. first proposed that the human population contains at least four different HMG "phenotypes" in milk: *Secretor*$^{+}$*Lewis*$^{+}$, *Secretor*$^{-}$*Lewis*$^{+}$, *Secretor*$^{+}$*Lewis*$^{-}$, and *Secretor*$^{-}$*Lewis*$^{-}$ [20]. These phenotypes can be determined based on genetic methods and/or the presence of 2'-FL and LNFPII in mothers' breast milk samples. The proportions of these four phenotype varies between different human populations; for example, Latin American populations have a much lower prevalence (~1%) of the Secretor-negative phenotype and a higher prevalence of the Lewis-negative phenotype vs. European populations [21].

iii. Analytical Studies of HMGs

Qualitative and quantitative studies on HMGs have shown that HMGs are a remarkable component of human milk. The HMG concentration of human milk is so high that HMGs are one of the major biomolecules in human milk along with lactose, lipids, and protein. Unlike the latter three, HMGs are mainly non-nutritional, which may explain why they have been much less studied than the nutritional components. Nonetheless, this high concentration of HMGs has made people wonder why HMGs are such a major component of human milk even though they do not represent a significant direct nutritional source for the infant (see the *Metabolism and Digestion of HMGs* section below for more information). An excellent review on the major analytical studies that determined the concentration of individual and total HMGs in human milk has been recently published [12].

The total HMG concentration in human milk ranges from about 8-25g/L [12, 26, 27]. The concentration ranges of the most common HMGs in human milk are displayed

in **Table 1**. It should be noted that the values of the individual and total HMG concentrations are based on the average concentration in samples as opposed to individual samples. The fact that these average concentrations also gave very high standard deviations (sometimes reaching a %CV as high as 150%) suggests that much individual variation is also present in total and individual HMG concentrations (refer to the data reviewed and presented in [12]). The high variability between studies is a result of both biological heterogeneity and the technical methodology used for HMG measurement [13].

The biological heterogeneity in HMG concentrations is due to a number of factors including genotype, number of days postpartum, and possibly other factors. In the case of genotype, lack of expression of the Secretor α1-2 fucosyltransferase results a reduction of total HMG concentration in human milk [28]. The most likely reason for this reduction is because 2'-FL, the major HMG in human milk, is no longer synthesized. As a result, less 2'-FL is produced from lactose, and hence less total lactose is converted to HMGs. Additionally, the large concentration range of certain fucosylated HMGs (**Table 1**) is likely primarily due to heterogeneity in the Secretor and Lewis phenotypes within the human population. Depending on the population studied, the proportion of Secretor- and/or Lewis-negative individuals may vary. If the population studied contains a lower prevalence of the Secretor- and/or Lewis-negative genotype, the lower the concentration of specific fucosylated HMGs, and vice-versa. The total HMG is also higher in colostrum than in mature milk, being around 20-25g/L in colostrum and 10-17g/L in mature milk, and the total HMG concentration as well as concentrations of most individual HMGs appears to decrease as the number of days postpartum increases [26, 28, 29]. Individual

variations in HMG concentrations are also seen that are independent of lactation stage or genotype [30], suggesting that other unknown factors may influence HMG concentration in human milk.

The technical methodology may also play a role in HMG concentration measurements. Most quantitative studies in the last 30 years have used quantitative HPLC analyses, including HPLC of fluorescently derivatized HMGs and HPAEC-PAD of underivatized HMGs, to measure total and individual HMG concentrations. These chromatographic methods, along with prior methods for HMG enrichment from milk, all introduce important analytical variables that may not be well controlled. For example, some methodologies require prior removal of lactose by an initial chromatography step [28, 31], which may unintentionally result in some HMG sample loss, and most studies did not include internal standards to measure and adjust for HMG loss during this step. Besides HMG loss during processing, analytical chromatography measurements rely on the selection and suitability of the reference standard(s) as well as a chromatography method that quantitatively separates HMGs from lactose, monosaccharides, and other contaminating biomolecules. Moreover, reporting the total HMG concentration in human milk using chromatography-based measurements presents an additional difficulty of being able to quantitate a highly structurally heterogeneous population. One method of total HMG quantitation is using a chromatography method where HMGs elute as a single peak separately from lactose and monosaccharides [26]. While very straightforward, this methodology may underestimate the total HMG concentration because not all HMG may elute together. A second methodology separates the HMGs by chromatography and measures the total HMG concentration as the sum of the concentrations of the individual

HMGs, which are mostly simple HMG structures including 2'-FL, 3-FL, LNT, LNnT, LNFPI, LNFPII, LNDFHI, LNDFHII, 3'-SL, 6'-SL, DSLNT, LSTa, LSTb, and LSTc [27-29, 32]. Some of these studies utilizing this methodology also separated neutral and acidic (i.e.-sialylated) HMGs by ion-exchange chromatography prior to the analytical chromatography to reduce chromatogram complexity [28, 31, 32], which may again lead to some uncontrolled HMG loss due to elution with contaminants, inadequate elution from column, etc. For these reasons, the actual individual HMG concentrations are likely an underestimate since the more complex HMG structures may not be quantitated and/or HMG sample loss may have occurred.

The major HMG structure in human milk is typically 2'-FL (1-5g/L), while the Type 1 LacNAc-containing HMG structures LNT and LNFPI (0.5-4g/L) are also highly abundant (**Table 1)**. LNnT, 3-FL, LDFT, LNFPII, LNFPIII, LNDFHI, and LNDFHII as well as the sialylated HMGs LSTa, LSTb, LSTc, DSLNT, 3'-SL, and 6'-SL (0.05-2g/L) are also relatively abundant (**Table 1**) [12]. Over 200 more HMG structures also exist, although these structures are much less abundant and not as well understood in terms of their actual concentration and concentration ranges in breast milk [8]. The biological significance of these complex HMGs is also unclear, especially because these structures are present at such low concentrations. One explanation may be that the more complex structures may be of low abundance simply because they have biological effects at very low concentrations in contrast to the simple HMG structures. This suggestion may be supported by the findings of this study, where some glycan-binding proteins appear to bind strongly to these more complex HMG structures, perhaps even stronger than to the simple HMG structures (described in more detail in **Chapter 2** and **Chapter 3**).

Therefore, more analytical studies and biological studies on these complex HMGs are needed.

These analytical studies have been essential for the understanding of the structural and quantitative aspects of human milk. They have also shown how complex human milk is since the concentrations of HMGs (and other bioactive components) vary over the course of lactation and even between individuals. Despite the complexity, these studies will prove useful from a clinical and pharmaceutical standpoint if research and clinical studies continue to show health benefits from HMGs and their supplementation in infant formula. Furthermore, the analytical studies have also raised some important questions about HMG diversity and concentration. For instance, does alteration in HMG composition and/or concentration affect infant health? The diversity of genetic status and its affect on HMG structures present in human milk also raises the question of whether or not HMG compositions from these different genetic backgrounds not only affect infant health, such as the gut microbiota and microbiome (as described below), but also why they exist in the first place. If one genetic background were found to be less efficient in its function, why would such a genetic background represent a significant fraction of the human population? In addition, would HMG supplementation into human milk from these less efficient genetic backgrounds (ex-2'-FL supplementation into Secretor-negative milk) help restore the beneficial effects lost? These questions are all worth future study to further unravel how HMGs affect and improve infant health, as well as what combination of structures, concentrations, and feeding duration leads to the best effect on infant health.

iv. Metabolism and Digestion of HMGs by Infants

Lactose, the predominant glycan found in human milk, functions as a significant carbon and energy source for the infant. On the other hand, HMGs generated from a lactose precursor may not serve as a significant nutritional source for the infant (at least not directly, since fermentation by the colonic microflora may generate some short-chain fatty acids that may be used as an energy source by humans). The major reason for this belief is because HMGs can be left out of infant formula without any consequence on infant *nutrition* although, as described in this study, the HMGs may still have *health-benefiting bioactive properties*. Some *in vitro* studies have been carried out to understand HMG digestion, role as a bioactive factor, and even localization in the infant, to further understand the role of HMGs in the infant.

In vitro studies have suggested that minimal hydrolysis of HMGs occurs in the upper GI tract. For example, studies showed little if any hydrolysis of HMGs by human intestinal brush border membranes [33, 34] and human duodenal aspirates [33]. Furthermore, acidic (pH = 2.5) conditions similar to the pH of stomach acid had little effect on HMGs, although very slight desialylation of sialylated HMGs was detected, after 2 hours of incubation [34]. On the other hand, whole porcine pancreas extract, which contains glycosidases that are normally localized to intracellular compartments, especially the lysosome, showed hydrolytic activity against HMGs [33]. While pancreatic secretions and small intestinal brush border membranes are thus incapable of hydrolyzing HMGs, intracellular enzymes (like lysosomal enzymes) are still capable of HMG digestion. Therefore, it cannot be ruled out that some HMG digestion may occur if HMGs are taken up (ex-by pinocytosis) and directed through the endosomal pathway.

Unfortunately, the contribution of pinocytosis and other methods (ex-specific transporters) for HMG uptake by enterocytes has not been well studied. These studies may also be important to understand another phenomenon of HMGs: their appearance in infant urine [35].

HMGs were determined to be present in the urine of both preterm [35] and 3-6 month postpartum infants [36], indirectly suggesting that HMGs enter the circulatory system. It should be noted, however, that lactose and sialic acid are also detected in the urine of both breast- and formula-fed infants [35], meaning that this phenomenon is not limited to glycans in human milk. The source of these HMGs was determined by feeding the mother ^{13}C-labeled galactose, the appearance of ^{13}C-labeled HMGs in the mother's milk, and subsequent appearance of ^{13}C-labeled HMGs in the infant urine [36]. Therefore, the HMGs were derived from the mother's breast milk. That study by Rudloff et al. also calculated an average of ~1% of the total amount of HMGs from mother's milk ends up in the infant's urine [36]; the limitations of this calculation were that only one infant was tested at five time points, and only the concentration of two HMGs (LNT and LNFPII) were measured, which does not take into account the potential of differential HMG absorption and/or excretion of different HMG structures. HMGs have also been detected in the urine of formula-fed infants [35]; the source of these glycans is likely the cow's milk used as the base for infant formulas. Recently, more direct evidence showed the presence of some but not all HMGs breast-fed infant plasma, although those HMGs that were present, such as 2'-FL, in the plasma were at more than a 10-fold lower relative concentration (~0.05% of total HMG concentration in milk) than seen in urine (~1-4%) [37]. Therefore, HMGs appear to enter the infant circulation.

The mechanism by which HMGs enter the circulation is still not fully understood. In the Caco-2 cell line, it was shown that neutral and acidic HMGs can be applied to the apical side and recovered from basolateral side [37, 38]. Indirect studies suggest receptor-mediated transcytosis may partially explain neutral HMG transport because neutral but not acidic HMGs appear in the intracellular fraction after exposure to the apical side only, and both brefeldin A and bafilomycin A partially inhibited LNFPI and LNT (neutral HMGs) but not 6'-SL (an acidic HMG) transcellular transport in Caco-2 cells [38]. Furthermore, the transport of at least LNFPI and LNT were saturable, showing specific receptors were involved [38]. On the other hand, 6'-SL transport occurred equally when applied to either the apical or basolateral side of the membrane, leading the authors to suggest that 6'-SL and other acidic HMGs are transported primarily by paracellular transport [38], although pinocytosis cannot be ruled out. Indeed, the role of pinocytosis in HMG transport has not been studied, which also play a role in both neutral and acidic HMG transport in a more non-specific manner. Studies of HMG transport within the primary intestinal epithelial cells as well as within the human intestinal tract are lacking and needed in order to help pinpoint where HMG transporters are expressed in the small and/or large intestine, the identity of the neutral HMG receptor(s), and if HMG transport is different for different HMG structures and in different parts of the infant intestine. Importantly, HMGs are more likely to be digested by microbes in the large intestine than the small intestine, since the small intestine (especially the upper parts) contains few microbes and low microbial diversity. In other words, HMG transport in the colon may be a "kinetic battle" between microbial HMG digestion vs. HMG uptake by the colonic

epthelium. Thus, HMGs pass largely intact through most of the GI tract but will be digested to some degree by the colonic gut microflora.

The significance of these studies is that it suggests HMGs may enter the infant's circulatory system and disseminate throughout the infant's body. In other words, the effects of HMGs may not merely be limited to the GI tract and may have more systemic effects, although the majority of known HMG functions occur in the GI tract. Five classes of HMG functions can be defined (**Figure 2**) and are described in more detail in the following five sections.

v. Functions of HMGs: Prebiotics and Microbiome Regulation

One of the oldest known functions of HMGs is to act as prebiotics to promote the growth of potentially symbiotic microbes in the infant intestine and, as a result, alter the infant gut microbiome in a potentially beneficial manner. Since around 1900, it has been known that human milk promoted the growth of a specific genus of bacteria, the *Bifidobacteria* that were, at the time, thought to be a single species termed *Lactobacillus bifidus* (reviewed in [7]). In 1953, György identified the unknown non-lactose carbohydrate fraction of human milk, termed "gynolactose" at the time, as the so-called "bifidogenic factor" of human milk [39]. This seminal study led to efforts to structurally and functionally characterize the glycans of this "gynolactose" fraction (reviewed in [7]). Due to the potential ability of these bacteria to inhibit pathogen colonization and infection by reducing the intestinal pH and/or competing for nutrients [7, 39], this prebiotic property of HMGs became a topic of major interest, especially in the 21st

century where a substantial role of the microbiota in controlling health and infection has been accepted.

HMGs are classified as prebiotics, which are biomolecules that are not digested by humans but rather support the growth and maintenance of beneficial microflora, particularly in the GI tract [40]. Some microbes use these prebiotic HMGs as their major source of carbon and/or energy, which allows HMGs to promote the colonization and growth of these microbes in the infant's distal GI tract, especially the colon. The experimental method of measuring prebiotic activity is by incubating microbial strains known to occur in the gut microbiota of human milk-fed infants in a medium containing total HMGs or specific HMG structures as the sole carbon and energy source [41]. Using this method, HMGs have been found to serve as a primary source of carbon and energy for symbiotic bacteria such as *Bifidobacterium longum* subsp. *infantis* and specific strains of *Bifidobacterium longum subsp. longum*, *Bifidobacterium breve*, and *Bifidobacterium bifidum* [41-45]. Interestingly, at least some strains of *Bifidobacterium longum* subsp. *infantis* preferentially utilize HMGs as the major carbon and energy source over glucose [41]. A few other species, including specific strains of *Bacteriodes thetaiotaomicron*, *Bacteriodes fragilis*, *Bacteriodes caccae*, *Bacteriodes vulgatus*, and *Lactobacillus delbruckii* may also utilize HMGs [45-47]. Very few other gut bacterial strains are known to metabolize HMGs. Since not all strains of a given species catabolize HMGs [42], the selection for gut colonization in human milk-fed infants is likely based on the strain-level rather than species-level. Hence, even species-level microbiome analyses must be interpreted with care and the knowledge that only specific strains of these bacterial

species are likely present, each of which may present its own unique set of functions of the species' pangenome.

How do HMGs enrich for specific *Bifidobacteria* species and other bacterial strains? There are at least two ways, the first being these bacteria express enzymes to catabolize HMGs and utilize specific monosaccharide or disaccharide structures as an energy and/or carbon source. Secondly, HMGs express the determinants that can be utilized by these bacteria. For example, the LacNAc I determinant (Galβ1-3GlcNAcβ-) is a major determinant found on HMGs such as LNT and a major target of digestion by specific *Bifidobacterium* strains [48, 49]. Interestingly, structures containing non-reducing LacNAc I (Galβ1-3GlcNAcβ1-3R) determinants predominate over Type II terminating LacNAc structures (Galβ1-4GlcNAcβ1-3R) in human milk, which is in stark contrast to nearly all other known animal milks where structures containing non-reducing LacNAc II determinants predominate [12]. This difference may possibly contribute to differences in microbiome regulation between infant humans and other primates. In fact, it has been hypothesized that these LacNAc I determinants may predominate in human milk because they are the major determinant metabolized by *Bifidobacterium* strains and thus represent the "true bifidus factor" of human milk and the HMG fraction [12, 48].

Other possible HMG determinants utilized by these bacteria are α-linked sialic acid and α-linked fucose. Sialylated HMGs such as 3'-SL and 6'-SL are utilized by strains of *Bifidobacteria* and *Bacteriodes* [45, 47]. Most of these sialylated HMG-utilizing strains produce neuraminidase enzymes and, in the case of strains of *Bifidobacterium longum* subsp. *infantis* and *Bacteriodes fragilis*, the liberated Neu5Ac can be directly catabolized as a source of carbon and energy [47, 50]. Fucosylated HMGs

such as 2'-FL, 3-FL, and LDFT are utilized by some *Bifidobacteria* and *Bacteriodes* strains [45]. It should also be α1-2 and α1-4 fucosylation are also variable within the human population (described in more detail in the *Analytical Studies of HMGs* section). These differences in HMG fucosylation may confer functional differences in infants, including differences in the microbiota composition. Indeed, Lewis et al. showed that Secretor-positive mother's milk is associated with a more beneficial microbiota, including higher *Bifidobacteria* levels, and faster acquisition of a beneficial microbiota in infants vs. Secretor-negative mother's milk [51]. Therefore, differences in HMG composition between mothers and a given mother over time (ex-colostrum vs. mature milk) has functional consequences on at least the microbiota composition and susceptibility to gastrointestinal infectious diseases (see *Inhibition of Infectious Diseases* section below). While α-L-fucosidase enzymes are expressed by most of these microbes [47, 52], whether or not the released L-fucose acts as a carbon and energy source is unclear, as is whether or not fucosylation itself or the increased HMG levels in Secretor-positive milk promoted the increased *Bifidobacteria* levels in the Lewis et al. study [51]. Therefore, a number of HMG determinants besides LacNAc I may serve as prebiotics for symbiotic bacterial species. For more intensive information on the mechanisms of HMG uptake and metabolism by *Bifidobacteria* and other bacteria, a number of useful reviews are available [52-54].

Human milk is likely one of the major factors controlling the gut microbiota composition of infants. In fact, of 18 different variables tested, only five were found to significantly affect the gut microbiota of infants, one of the two most dramatic being the method of feeding (exclusive breast-feeding vs. exclusive formula-feeding or

combination feeding) [55]. The difference in the gut microbiota between breast-fed and formula-fed infants is seen by both culture-based and genome-based methods and likely results from the presence of HMGs in human milk [56, 57]. As described earlier, breast-fed but not formula-fed infants also exhibit a *Bifidobacteria*-rich gut microbiota, as seen by an increased frequency of *Bifidobacteria* carriage and/or proportion of the total gut microbiota vs. formula-fed infants [58, 59]. The initial source of these microbes is most likely from the mother since monophylic strains have been identified in the infant gut that match those found in the mother's vagina and even mother's milk itself [60, 61], suggesting the vaginal births and breast-feeding are also important in delivering and establishing these probiotics in the infant intestine. HMGs are thought to be a major contributor to these differences in the microbiome between human milk-fed and formula-fed infants. To date though, no direct experimental or epidemiological evidence of the degree of HMG influence on these microbiota changes is available, such as studies on microbiota composition between infants fed formula and formula supplemented with HMGs.

The potential biological significance of these microbiome changes and symbiotic bacteria enrichment is manifold, and some of the known effects of these probiotics are described here. First, these microbes may inhibit colonization, levels, and/or virulence of potentially pathogenic organisms. For example, infant formula supplementation with *Bifidobacterium bifidum* along with *Streptococcus thermophilus* was shown to reduce the incidence of diarrhea and rotavirus carriage vs. unsupplemented formula [62]. Additionally, carriage rates of potentially pathogenic bacteria such as *Clostridium difficile* and *Enterococcus faecalis* is typically lower in breast-fed vs. formula-fed infants

[55, 56, 58]. The mechanisms of these antimicrobial and anti-virulence effects are likely multifactorial, including competition for nutrients and colonization sites due to high levels of carriage (up to 10^{11} *Bifidobacteria* cells/g stool) [55, 56], *Bifidobacteria*-mediated acidification of intestinal lumen contents [58, 63], production of antimicrobial compounds such as bacteriocins [64], and possibly other methods. Moreover, these bacterial species regulate a number of immunological responses, such as increased IL-10 and IFN-γ production, Th1:Th2 balance vs. the Th2-biased immature or germ-free intestine, reduced risk of inflammatory bowel diseases, and stimulation of T cell activation and immune system development mediated by *Bacteriodes fragilis* Polysaccharide A [65, 66]. *Bacteriodes thetaiotaomicron* also stimulates mucin production and goblet cell differentiation in the colon [67], which may be important for establishing the initial barrier between the gut microbiota and intestinal epithelium. The lack of or defects in this mucus barrier result in bacterial contact with the intestinal epithelium, which is thought initiate inflammation and ultimately result in inflammatory bowel diseases [68, 69]. Additionally, these bacteria supply a number of beneficial nutrients and bioactive components including short-chain fatty acids (SFCAs). These SFCAs not only serve as a major energy source for colonocytes and other cell types but also promote a number of health benefits including reduced weight gain and increased levels of anti-inflammatory regulatory T cell populations in the intestine (reviewed in [70]). Therefore, HMGs promote the establishment of these symbiotic bacteria that may assist with health, physiology, immunity, and development of newborns and infants. However, more studies are needed to confirm that the breast-fed infant gut microbiota

improves health, immunity, and/or development vs. that of formula-fed infants and if HMGs play a major role in these microbiota differences.

HMG regulation of the microbiota may also possibly contribute to disease. A recent study showed that mice fed milk deficient in α2-3-sialylated but not α2-6 sialylated HMGs were slightly more resistant to DSS-induced colitis [71]. The authors suggested that this effect was at least partially due to alterations in the gut microbiota since microbiota transfer from the α2-3 sialic acid-deficient milk mice to germ-free mice caused higher resistance to DSS-induced colitis vs. microbiota transfer from mice fed normal milk [71]. A follow-up study showed that this effect is at least partially due to stimulation of $CD11c^+$ dendritic cells in the gut in a TLR4-dependent manner [72]. The authors proposed that 3'-sialyllactose (3'-SL), a major α2-3 sialylated HMG in human and mouse milk, was the primary contributor [71, 72], although this was not experimentally verified. The use of cross-feeding of mice between mothers with or without the ST3Gal4 knockout also raises another variable that was not controlled for. Additionally, a genetic knockout of ST3Gal4 may have other effects such as loss of α2-3 sialylation of milk glycoproteins, alteration of milk composition, etc. The follow-up study [72] did use defined, LPS-free 3'-SL, which did induce colitis, but lactose was used as the control and not 6'-SL or some other non-digestible HMG; in addition, the results were only tested in an IL-10 knockout background. Therefore, additional studies are needed to confirm that these results are valid.

In spite of these prebiotic effects and possible relationships with health and disease, the prebiotic function of HMGs may not fully explain the total potential health benefits of HMGs. One major reason for this belief is because supplementation of infant

formula with plant-derived oligosaccharides, such as fructooligosaccharides (FOS) and/or galactooligosaccharides (GOS), causes the same prebiotic and intestinal lumen acidifying properties as human milk [73]. However, recent systematic studies have shown little to no health benefit of these FOS/GOS supplementations [74, 75]. It should be noted though that the microbial strains may be somewhat different between human milk- and prebiotic formula-fed infants since the strains metabolizing HMGs and FOS/GOS may not overlap [41, 43, 54]. Although far from a perfect model for drawing conclusions, this observation suggests the beneficial effects of HMGs may only be partially explained by their prebiotic effects.

vi. Functions of HMGs: Inhibition of Infectious Diseases

HMGs contain glycan determinants that are highly similar to those expressed on the cell surface, such as the Blood Group H determinant, Lewis determinants, LacNAc I, LacNAc II, and α2-6-sialylated LacNAc II (Neu5Acα2-6Galβ1-4GlcNAcβ1-3-). For this reason, the free, solution-based HMG structures are in a prime state to act as soluble inhibitors of cell surface binding by glycan-binding proteins (GBPs), including the GBP adhesins expressed by pathogens (reviewed in [76, 77]). Indeed, HMGs have been found to act as inhibitors of binding to and infecting cells by numerous pathogens; the HMG(s) responsible for this inhibition are typically mimics of the cell surface glycoconjugates receptor(s) for the given pathogen (reviewed in [78]). Hence, HMGs may function as "receptor decoys" against pathogens.

The first suggestion of an anti-adhesive function of HMGs was suggested by Kuhn and colleagues, who discovered that human milk inhibited influenza virus

infection; they suggested that sialic acid, a component of the HMG fraction and known receptor of influenza virus, mediated this effect via binding to the influenza and preventing the virus from binding the sialylated cell surface receptors (reviewed in [7]). It is currently well known that influenza viruses bind to 3'-SL and/or 6'-SL depending on the strain [79], although the binding strength to these sialyllactoses is probably non-physiological since the binding affinity is typically in the millimolar range. More recently, analysis of human influenza virus binding to natural human milk glycan microarrays showed a specificity of human influenza viruses for specific sialylated HMG structures, which may represent relatively high affinity structures [80]. Due to the paucity of studies though, more *in vitro* and *in vivo* studies on the effects of the protective effects of sialylated HMGs against human influenza virus infection are still needed to confirm these observations.

Many studies have shown anti-adhesive effects of HMGs towards a number of pathogens and toxins including *Campylobacter jejuni*, *Vibrio cholerae* and cholera toxin, enteropathogenic *Escherichia coli* strains and *E. coli* heat-stable toxin (ST-*E. coli*), specific strains of norovirus, *Pseudomonas aeruginosa*, *Steptococcus pneumoniae*, *Haemophilus influenzae*, and influenza viruses (reviewed in [78, 81]. While the interaction of HMGs with gastrointestinal pathogens is more likely to occur *in vivo* due to the very high local HMG concentration in the breast-fed infant GI tract, it is also plausible that HMGs interact with respiratory pathogens since a fraction of HMGs enters the circulation and thus could conceivably enter the respiratory tract [36, 37]. Support of these interactions between pathogens and HMGs is mainly derived from *in vitro* studies, such as inhibition of cell infection and/or pathogenesis, along with animal studies

supporting a role in protection against infection and/or disease severity (as an example, refer to [82]). Although the *in vitro* studies may support a mainly receptor decoy effect, other effects such as immune-modulation cannot be fully ruled out. Additionally, the results of the *in vivo* studies are complicated by other factors including regulation of the microbiota and gene expression, so the effects of HMGs *in vivo* may not be exclusively due to receptor decoy activity. However, strong evidence exists for HMGs inhibiting infection or virulence by three pathogens: *C. jejuni*, ST-*E. coli*, and noroviruses (described in more detail below).

Epidemiological studies also support the notion that HMGs protect against infection. For example, within Secretor-positive mother-infant pairs, higher levels of 2'-fucosylated HMGs are associated with slightly increased protection against gastrointestinal infectious diseases, including *C. jejuni*, heat stable-toxin-associated diarrhea, and norovirus [21, 83]. These epidemiological findings supported the *in vitro* and animal model studies that 2'-fucosylated HMGs inhibit infection and/or virulence of *C. jejuni*, ST-*E. coli*, and norovirus [14, 82, 84]. While these epidemiological findings are typically credited to the inhibition of pathogen adhesion (i.e.-"receptor decoy") property of HMGs, other HMG activities such as regulation of immune responses and the microbiota cannot be ruled out as contributors to this phenomenon. Nonetheless, HMGs may inhibit infection by a number of pathogens, especially the gastrointestinal pathogens *C, jejuni*, ST-*E. coli*, and norovirus strains, at least partially due to receptor decoy activity.

vii. Functions of HMGs: Regulation of the Immune System

Besides acting on microbes and the microbiota, HMGs can directly function as regulators of immunological and physiological processes. Most of these effects have been studied *ex vivo* by adding defined, purified HMGs to cell or tissue cultures and examining changes in the cell morphology, viability and proliferation, differentiation, factor (ex-cytokine) secretion, etc. The cells/tissues tested have been primarily intestinal epithelial cells and leukocytes, cell types that are likely to come in contact with ingested and/or absorbed HMGs. The effects on other cell types that likely come in contact with HMGs, especially epithelial cells in the upper GI tract (oral cavity, esophagus, stomach, and duodenum), remain to be tested.

The seminal study on the immune-regulatory function of HMGs was by Eiwegger et al. in 2004, who showed that total acidic HMGs slightly altered the percentage of T cells expressing specific cytokines, including IFN-γ and IL-13 [85]. Additionally, total sialylated HMGs increased the percentage of $CD4^+$ T cells that expressed CD25, a marker of T cell activation, suggesting that sialylated HMGs may also stimulate T cell activation and/or differentiation [85]. A follow-up study in 2010 by Eiwegger et al. showed that total sialylated HMGs also slightly stimulated IL-10 and IFN-γ but not IL-13 production by cord blood mononuclear cells [86]. In addition, total sialylated HMGs slightly reduced IL-4 production by Ara h1-specific T cells from individuals with peanut allergies but not T cells from healthy patients. Furthermore, there is controversial evidence that sialylated HMGs may interact with selectins on leukocytes to reduce leukocyte rolling and neutrophil activation [87, 88]; the reason for the controversy is described in more detail in the **HMG Interactions with Human Glycan-Binding**

Proteins section. Neutral HMGs can also regulate cytokine production; for example, physiological concentrations of 2'-FL reduced IL-8 expression by epithelial cells stimulated with LPS [89]. Surprisingly, even lactose may regulate immune responses; Cederlund et al. identified lactose as the major component of human milk that promoted increased expression of hCAP-18, the precursor of the cationic antimicrobial peptide LL-37, in colonic epithelial cells and monocytes [90]. Therefore, HMGs regulate a number of immune responses from cytokine and antimicrobial peptide expression to immune cell activation.

He et al. recently published a highly in-depth study on the immunoregulatory functions of colostrum. In that study, the authors showed that total colostrum HMGs but not total mature milk HMGs changed the gene expression profile of primary human fetal intestinal tissue, specifically cytokines, cytokine receptors, chemokines, and signaling factors [91]. Analysis of the cytokines and regulatory pathways altered by the colostrum HMGs suggested upregulation of Th1 cytokines, downregulation of Th2 cytokines, slight downregulation of Th17 cytokines including IL-17, downregulation of proinflammatory cytokines (TNF-α, IL-1β, and IL-6), and downregulation of chemokines (including the potent neutrophil-recruiting chemokine IL-8) in resting and/or PAMP-stimulated fetal intestinal tissue [91]. In contrast to the Eiwegger et al. 2010 study described above, the anti-inflammatory cytokine IL-10 showed increased mRNA expression but slightly reduced protein expression in the He et al. study [91]. This IL-10 downregulation may be a response to the concurrent downregulation of proinflammatory cytokines so that a balanced and moderate strength immune response is generated. The reason that colostrum but not mature milk HMGs stimulated these changes is still unclear, but it may involve

the physiology of preterm vs. term intestinal tissue in its capacity to respond to HMGs (ex-differential expression of GBPs, as suggested in this current study) and/or unidentified glycan structures that are present in colostrum but not mature milk. Therefore, HMGs appear to promote production of cytokines that promote anti-inflammatory effects and/or slightly redirect the immune system from Th2 to Th1 responses (ex-IFN-γ) and thereby lead to a more balanced Th1:Th2 response or a Th1:Th2 response most appropriate for newborns and infants. By reducing the pro-allergenic Th2 responses and lowering proinflammatory responses, including Th17 responses and reduced proinflammatory cytokine production, these cytokine expression changes are expected to reduce the likelihood of allergenic and highly damaging inflammatory responses.

3'-SL may also promote activation of specific intestinal dendritic cell populations in a Toll-like receptor 4 (TLR4)-dependent manner [72] (refer to *Functions of HMGs: Prebiotics and Microbiota Regulation* section for more information on and interpretation of this study). On the other hand, 2'-FL may suppress TLR4 signaling in epithelial cells by downregulating cell surface CD14 expression [89]. Therefore, a number of HMGs may regulate TLR signaling.

Two recent *in vivo* studies further support beneficial immunoregulatory functions of HMGs. The study by Jantscher-Krenn et al. showed that HMG supplementation into infant formula significantly increased protection from morbidity and mortality in a neonatal rat model of necrotizing enterocolitis (NE), an effect that was almost fully recapitulated with a single HMG, disialyllacto-N-tetraose (DSLNT; refer to **Table 1** for structure) [92]. NE is a fulminating, fatal inflammatory disease that is difficult to treat

and primarily occurs in infants born pre-term [93], and human milk feeding may significantly reduce the risk of NE vs. formula-feeding [94, 95]. Therefore, the Jantscher-Krenn et al. study not only provides a biological basis and support for NE protection by human milk but also suggests that HMGs, specifically DSLNT, are beneficial immune-regulatory molecules. Jantscher-Krenn et al. also suggested that the levels of DSLNT in human milk may be a biomarker for risk of developing NE and, if low, DSLNT supplementation may be beneficial [92]. It will be interesting to see if DSLNT also helps reduce the incidence of NE in humans and if DSLNT supplementation into infant formula reduces the risk of NE development when formula feeding is necessary. Another study showed that 2'-FL or 6'-SL slightly reduced symptoms in a mouse model of food allergy [96]. Therefore, HMGs appear to be true immunoregulatory molecules that may shape the neonatal and infant immune system to not only control infectious pathogens early in life but also prevent fulminating inflammatory and allergenic diseases during this critical development stage. The mechanisms by which these HMGs regulate immune responses are still unclear though and should be a major topic of future study.

It should be noted that all the studies described in this section used free, reducing, underivatized HMG structures. Ideally, experimental studies on the immunoregulatory and other regulatory properties of HMGs should only use free, reducing glycan structures and not multivalent conjugates of such molecules, such as proteins or beads derivatized with HMGs. The reason for this is because the multivalent presentation of glycans does not represent the natural biochemistry of HMGs and may lead to non-physiologically relevant results. For example, proteins such as human serum albumin (HSA) derivatized to carry multiple LNFPIII or LNnT molecules per protein (termed "neoglycoproteins")

were shown to induce spleenocyte proliferation. However, the free, underivatized LNFPIII and LNnT did not cause proliferation and, in fact, competed with and inhibited the mitogenic activity of the LNFPIII or LNnT neoglycoproteins [97]. For these reasons, any studies utilizing HMG glycoconjugates should be interpreted with much caution until results with the free, underivatized, reducing glycan structures are available.

viii. Functions of HMGs: Regulation of Gene Expression in and Differentiation by Intestinal Epithelial Cells

HMG-mediated regulation of gene expression in epithelial cells was a phenomenon first described in 2005, when Angeloni et al. showed that 3'-SL treatment of preconfluent Caco-2 cells reduced binding of multiple lectins to the cells [98]. The reduction of lectin binding in the 3'-SL-treated, preconfluent Caco-2 cells also coincided with a reduction of specific α2-3 sialyltransferase gene expression (although the data was not shown) and binding by an enteropathogenic *E. coli* (EPEC) strain. These effects were specific to preconfluent cultures since no effect was seen with confluent Caco-2 cells, which have a more highly differentiated phenotype than preconfluent Caco-2 cultures [98]. This confluence-dependent change in lectin-binding profile suggested that 3'-SL was specifically modulating gene expression in less differentiated intestinal epithelial cells such as preconfluent Caco-2 cells. Indeed, another study showed that 3'-SL and other neutral and sialylated HMGs promoted a decrease in cell proliferation and an increase of alkaline phosphatase activity, both of which are markers of differentiation, but only in less differentiated cell lines like HT-29 but not Caco-2 cells [99]. These results further suggest that HMGs directly alter gene expression in developing epithelial cells

that have not yet fully differentiated. Total neutral HMGs at high concentrations also appear to stimulate apoptosis of less differentiated intestinal cells such as HT-29 cells but not more differentiated cells like preconfluent Caco-2 cells [99]. The mechanism of this growth inhibition and differentiation was shown to be primarily due to cell cycle arrest during the G2/M phase that is correlated EGFR signaling as well as increased Cyclin-B and CDKI expression [100]. The HMG 2'-FL has also been shown to have anti-proliferative effects on the Caco-2 C2Bbe1 cell line, a model of absorptive intestinal epithelial cells (enterocytes) [101]. In that study, 2'-FL but not LNnT or 6'-SL at physiological concentrations slightly increased cellular alkaline phosphatase and sucrase but not lactase activity in postconfluent C2Bbe1 cells, which are markers of enterocyte differentiation and thus suggest increased differentiation to absorptive, villi-associated enterocytes [101]. 2'-FL can also reduce CD14 expression by the T84 and HCT8 intestinal epithelial cell line stimulated with LPS or pathogenic *E. coli* and thereby reduce infection and invasion by these bacteria as well as the inflammatory response [89]. In all, these studies suggest that HMGs promote changes in intestinal epithelial cell gene expression, which especially promotes differentiation into more highly absorptive enterocytes. This differentiation causes increased expression of catabolic enzymes for cellular uptake, more efficient nutrient absorption by enterocytes, and intestinal epithelium remodeling by apoptosis. Together, these changes may promote increased efficiency of nutrient absorption by the infant intestinal epithelium, which may be important for infant nutrition, health, and development.

The most striking evidence that HMGs regulate epithelial cell gene expression came in a 2013 study by Lane et al., who treated HT-29 cells with 4mg/ml total HMGs, a

4mg/ml of a single (3'-SL), or a mock treatment and measured changes in gene expression by gene microarray analysis. Treatment with total HMGs was found to change the expression of over 1,000 different genes, many of which are involved in signaling, immunity, and development [18]. Treatment of HT-29 cells with 4mg/ml 3'-SL also changed the expression of a little less than 1,000 genes, and about half of the genes expression changes by 3'-SL and total HMG overlapped (i.e.-total HMG and 3'-SL caused both similar and unique changes in gene expression) [18]. The fact that total HMGs but not 3'-SL caused changes in some genes is most likely explained by the fact that HMGs other than 3'-SL influence these changes. Surprisingly, 3'-SL caused changes in gene expression that total HMGs did not, which was unexpected but may possibly be explained by low relative abundance of 3'-SL in the total HMG pool, which is lower than 4mg/ml [12], and/or antagonistic effects among HMGs. It should be noted that the Lane et al. study also performed a treatment with 4mg/ml of total bovine milk glycans and saw changes in gene expression, and the changes in gene expression overlapped somewhat with total HMGs and 3'-SL. Therefore, bovine milk glycans cause a unique but somewhat overlapping transcriptional response as HMGs, likely because the bovine milk glycan profile is much different but still slightly overlapping with the human milk glycan profile (described in more detail in the *Structural Features of HMGs* section). A proteomic study that specifically focused on proteins involved in signaling pathways showed significant changes in expression of 28/512 proteins involved in signaling upon HMG treatment [89]. The changes in expression of these signaling proteins may at least partially explain how HMGs, including individual HMGs, can lead to changes in expression of hundreds or even thousands of genes.

Based on these studies, HMGs regulate gene expression, alter signaling pathways, and promote differentiation and/or apoptosis of human intestinal epithelial cells, activities that may be crucial for proper intestinal development and thus infant health and development. Moreover, HMGs may alter host-microbe interactions by altering the cell surface glycan receptors for microbes via changes in gene expression of glycosyltransferase activity and possibly other effects. In other words, HMGs may also prevent infection by a mechanism independent of HMG receptor decoy activity. Future studies should also be aimed at determining if HMGs also cause functional effects on cell biology and/or gene expression in epithelial cells of other parts of the GI tract including the oral cavity, esophagus, and stomach.

ix. Functions of HMGs: Neuro-regulatory Effects

Another function of HMGs is their effect on neuronal cells and tissues, including the brain and enteric nervous system. This phenomenon was first described as early as 1994, but basic research on these neurological and cognitive effects did not receive much attention until the last few years. The seminal study in 1994 by Krug et al. showed that intrahippocampal injection of 2'-FL into rats boosted long-term potentiation (LTP) responses of rats, a neurological response that underlies at least some forms of learning and memory [102]. A follow-up study confirmed and extended this observation by showing that this effect was specific for 2'-FL but not another HMG, 3-FL [103]. Most recently, Vázquez et al. extended these observations by showing that 2'-FL feeding (the most natural method of HMG exposure as opposed to intrahippocampal injection in the previous studies) not only increased long-term potentiation responses but also increased

learning and memory in mice and rats [104]. Therefore, 2'-FL and possibly other HMGs also appear to enhance cognitive functions, suggesting an effect of HMGs on brain development and/or function. This result may at least partially explain the link between breast-feeding and increased intelligence test scores [4].

Bienenstock et al. conducted the first study examining the regulatory effects of HMGs on the peripheral nervous system (PNS), specifically the enteric nervous system (ENS) in 2013. That study showed that physiological concentrations of 2'-FL and 3-FL, but not other HMGs tested, reduced the strength and frequency of murine colon smooth muscle contractions in an *ex vivo* organ culture model [105], suggesting that some HMGs may regulate ENS and possibly other PNS functions, including motor functions. Since this is currently the only study examining the effects of HMGs on ENS functions, more studies are needed to confirm this observation. Although the ENS is the most likely candidate for a nervous system branch in close contact with HMGs during breast-feeding, additional studies should be aimed at extending these studies to other PNS branches since a small proportion of HMGs enter the circulation and hence may come in proximity with these other branches [37].

These studies strongly suggest that HMGs have neuro-regulatory functions. Future studies should be aimed at examining the mechanism by which HMGs exert their neuro-regulatory effects, especially if the regulatory effects are mediated by direct or indirect interactions of HMGs with neurons.

Despite the knowledge of these neuro-, immune-, and epithelial cell-regulatory effects of HMGs on cells, the mechanisms by which these regulatory effects occur have not been investigated. However, one hypothesis is that HMGs interact with glycan-

binding proteins found on the surface of GI tract epithelial cells, leukocytes, neurons, and possibly other cell types, which may ultimately induce changes in signaling pathways, gene expression, and/or other cellular activities that ultimately manifest as changes in physiological and immunological responses. This hypothesis has been proposed by others as well but was limited to the immunoregulatory effects of HMGs. [10].

II. Galectins

i. Background: Galectins

Galectins are defined as a class of β-galactoside binding proteins with significant sequence similarity in the key amino acid residues involved in β-galactoside recognition [106, 107]. Other key properties of galectins are their solubility (specifically, lack of a transmembrane domain) and externalization from cells by an unknown mechanism [106, 107]. Proteins that have homology to galectins but lack β-galactoside binding activity or have not yet been proven to bind β-galactosides are instead classified as "galectin-like" proteins [108]. Although classically characterized as "β-galactoside binding proteins," this description is somewhat misleading because galectins do not bind all β-galactosides. Specifically, the glycan receptors for galectins are usually derived from galactose-containing precursors, including lactose (Galβ1-4Glc), Type 1 N-acetyllactosamine (LacNAc I; Galβ1-3GlcNAc), and/or Type 2 N-acetyllactosamine (LacNAc II; Galβ1-4GlcNAc). The amino acid residues involved in these interactions were found to be conserved among active galectins [109] and are thus currently used as a bioinformatics method to detect and define potential galectin genes [106, 107].

The first galectin was discovered in electric eels by Teichberg et al. and termed electrolectin [110]. This galectin was discovered by its ability to specifically hemagglutinate trypsinized rabbit erythrocytes, which are unique from other erythrocytes because they contain very high levels of α- and β-linked galactose at the non-reducing ends of glycans. Strong additional evidence for electrolectin binding to β-galactosides was provided by hemagglutination inhibition by the β-galactosides lactose and its synthetic analog thiodigalactoside (TDG; galactose that is α1-1 linked to galactose by a

thioacetal instead of acetal linkage) [110]. The study by Teichberg et al. also showed the presence of this β-galactoside-specific hemagglutinating activity in animal tissues and organs [110], although human tissues were not tested. Electolectin and mammalian β-galactoside-binding proteins were then purified by affinity purification using β-galactoside-linked affinity matrices, especially asialofetuin-Sepharose and lactosyl-Sepharose, for characterization of these proteins [111, 112].

Galectin-1 (Gal-1) was the first galectin identified in humans, specifically in the human lung due to a similar tissue expression of a similar lectin in rat lung tissue, by the use of the hemagglutination and affinity purification methods described above [111]. Other names for what is now termed galectin-1 include L-14, lactose-binding lectin 1, and 14kDa lectin. Additional lactose-binding lectins have been discovered in humans and other mammals since then. Due to their similarity in protein structure and lactose-binding specificity, these lactose-binding proteins were grouped into a single family of glycan-binding proteins termed the galectins in 1994, with each member numbered based on the order of discovery in mammals [106].

Humans are known to express galectin-1, -2, -3, -4, -7, -8, and -9, which have all been confirmed to be “true” galectins. Humans also express galectin-10, -12, -13, -14, -16, and -17. However, these latter galectins have some mutations of key amino acid residues involved in lactose binding [113-116]. Additionally, the evidence that these galectins bind lactose is unclear. For example, the evidence of lactose binding is controversial between studies with galectin-10 [116-118] and galectin-12 [114, 115, 119]. In the case of galectin-13, -14, -16, and -17, only one or a few studies are available that demonstrates lactose and/or LacNAc II binding [113, 120]. However, these studies

showed questionable biologically relevant affinity and/or specificity for lactose or LacNAc II. Therefore, galectin-10, -12, -13, -14, -16, and -17 should be regarded as "galectin-like" until unequivocal experimental proof of specific lactose/LacNAc II binding has been observed. A few other galectin-like proteins are also present in humans, including GRIFIN, but the focus of this study will be on the "true" human galectin family members described above.

Galectins are highly labile because many but not all galectins aggregate, precipitate, and lose activity in the absence of a reducing agent like β-mercaptoethanol due to oxidation of cysteine and/or tryptophan residues; this phenomenon was particularly observed with the earliest galectins discovered (electrolectin and galectin-1) [110-112]. Indeed, active mammalian galectins lack disulfide linkages [121], meaning that oxidation and formation of disulfide linkages leads to a loss of activity, likely due to subsequent protein aggregation and precipitation. This finding is in contrast to many other lectins, which require disulfide linkages for glycan binding. Glycan binding by galectins also stabilizes them against oxidative inactivation either by blocking and/or interacting with Cys/Trp residues at or near the carbohydrate-binding site [112] and/or by promoting or stabilizing oligomer formation of the galectin [112, 122]. Whether the oxidative susceptibility of galectins has biological significance is currently unclear, but may be a mechanism of reducing the overall activity and/or half-lives of galectins in the extracellular, oxidizing environment.

The fact that galectins are susceptible to oxidative degradation makes working with these proteins inherently difficult. The proteins are most stable when stored in a physiological buffer containing a reducing agent (typically β-mercaptoethanol or DTT)

and an excess of ligand (lactose). Removal of lactose is essential for glycan binding experiments, and quantitative removal is certainly necessary for highly specific analytical studies such as K_d measurements. However, once both the lactose and reducing agent are removed from the galectin, the life span of galectin activity becomes significantly decreased, which limits the time of most assays only a few hours. Removal of the reducing agent is needed when working with cells due to toxicity and/or introduction of artifacts by the reducing agent itself. One important example demonstrating the necessity of reducing agent removal is in experiments testing the effects of galectins on T cells and neutrophils. In the presence of the reducing agent DTT, galectin-1, -2 and -4 induce T cell phosphatidylserine exposure and apoptosis but have no effect when DTT is omitted from the solution [123], while galectin-1 induces phosphatidylserine exposure without apoptosis in neutrophils but induces neutrophil apoptosis in the presence of DTT [124]. Two ways to extend the time of galectin activity without the need for a reducing agent is to stably prevent disulfide linkages by reduction and subsequent iodoacetamide labeling of thiol groups [124, 125] or to the use oxidation-resistant mutants such as the galectin-1 C2S mutant [126]. However, galectin activity should be verified by methods such as lactosyl-Sepharose binding when these modified galectins are used. In conclusion, one must be extremely vigilant when performing experiment with galectin proteins, and the systems used should be optimized prior to beginning actual experimentation. This optimization is needed to understand the length of time of galectin activity as well as variables that may interfere with galectin activity such as the presence of redox-active components or free glycan ligands of galectins in the medium.

At the genetic level, galectins are conserved among Metazoans and thus can be identified in nearly all multicellular organisms from sponges to humans [106, 107]. Most studies to date have focused on mammalian galectins, particularly human galectins because of their role in health and disease as well as rodent galectins due to the use of rodents as animal models of diseases. However, whether or not other organisms (particularly invertebrates) produce active galectins is unclear in many cases because the lactose-binding ability of these proteins has not been elucidated; in these cases, the term "galectin-like" should be used for galectins that lack or have not yet been tested for lactose-binding ability [108]. The reduced identity of galectin-like proteins from plants, fungi, and invertebrates compared to mammalian galectins due to evolutionary divergence also makes it difficult to clearly identify these sequences by bioinformatics [108, 127]. For mammalian galectins, the galectins were named with a number corresponding to their order of discovery in the mammalian family (ex-galectin-1, galectin-2, etc.) [106], but no specific nomenclature for non-mammalian galectins currently exists. The galectins expressed in humans are the major focus of this review, with details on their specific or potential functions, tissue expression, glycan-binding specificity, and other details.

While the lactose/LacNAc core structure is a major component of the major target glycan receptors of galectins, these core structures alone are relatively low affinity ligands for galectins, with K_d's typically in the high micromolar to low millimolar range. Rather, the highest affinity ligands are derivatives of these core structures. High affinity ligands tend to contain multiple units of lactose and/or LacNAc, either as repeating units as in poly-N-acetyllactosamine $(\text{-3Gal}\beta\text{1-4GlcNAc}\beta\text{1-})_n$ or from multiple branches of

branched glycans such as bi-, tri-, and tetra-antennary N-glycans. Moreover, further modification of the LacNAc units can further increase binding strength depending on the galectin, such as the addition of α1-2-linked fucose (Fucα1-2Galβ1-4GlcNAcβ-; Blood Group H), α1-2 fucosylation with α1-3 linked galactose or N-acetylgalactosamine (Galα1-3(Fucα1-2)Galβ1-4GlcNAcβ- and GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ-, Blood Group B and Blood Group A, respectively), or α2-3 linked sialic acid (Neu5Acα2-3Galβ1-4GlcNAcβ-). The actual high affinity glycan structures differ between each galectin; in other words, each galectin has a more or less unique glycan-binding specificity, which can be most easily seen by glycan microarray analyses [128]. Indeed, unique binding specificities of galectins were first observed as early as 1986 between three different rat galectins [129]. It should be noted though that the actual physiological glycan receptors *in vivo* is complex and dependent on temporal and spatial expression of target glycans, the expression and concentration of the galectin, the local concentration of target glycan(s), and the presence of competing glycan structures for galectin binding.

ii. Structural Features of Galectins

Structurally, galectins are β-sheet-rich proteins that contain the same β-sandwich fold found among L-type lectin family members such as concanavalin A (ConA), and are generally lacking in α-helices. This β-sandwich fold is generated from two antiparallel β-sheets: one five-stranded (strands F1-F5) and one six-stranded (strands S1-S6) β-sheet. The conserved amino acids H45, N47, R49, N58, E68, and R/K70 (amino acid numbering relative to galectin-2) are involved in hydrogen bonding with the lactose moiety and the conserved amino acids while V56 and W65 promote van deer Waals

interactions with the lactose moiety [109]. The presence of these eight amino acids is one of the two criteria used for classifying a protein as a true galectin family member [106], as described in the *Background: Galectins* section. Other residues indirectly promote lactose binding by stabilizing these active site amino acids [109]. All of these amino acids involved in directly binding or indirectly promoting lactose binding are located on β-strands S3, S4, S5, and S6, which are all encoded by a single, highly conserved exon among galectin family members [109].

The galectin carbohydrate recognition domain (CRD) contains at least four subsites for glycan binding, which was first proposed by Knibbs et al. [130]. Further studies, which used known glycan specificity, thermodynamics, X-ray crystallography, and structural modeling data confirmed and further refined this observation [131-133]. The current model proposes that five subsites (A-E) are present, where subsites C and D are involved in binding the core lactose/LacNAc unit, subsites A and B promote binding to the modifications and/or extensions of this core unit, and subsite E may be involved in interaction with factors at reducing end side of the lactose/LacNAc core unit, which may include aglycone components [108]. The structural differences present in these subsites, particularly subsites A, B, and/or E, may ultimately allow each galectin to have unique glycan binding specificities.

Hirabayashi and Kasai proposed the first galectin classification method, which was based on structural features of the galectin proteins. This classification, which is still commonly used to this date, includes three galectin groups: prototypical, chimera, and tandem-repeat galectins [121]. The prototypical galectins, which includes human galectin-1, -2, -7, -10, -13, -14, -16, and -17 simply consist of the galectin CRD, although

all of these prototypical galectins are known to exist nearly exclusively as oligomers in nature. The chimeric galectins consist of a CRD in addition to other structural and/or functional domains; in humans, galectin-3 is the only member and consists of an N-terminal collagen-like domain followed by a C-terminal CRD. The tandem-repeat galectins consist of two different CRDs on the same protein molecule that are linked by an unstructured linker peptide, and this class includes human galectins -4, -8, -9, and -12. While this classification scheme is useful, there are some issues, which warrant using this classification system with the knowledge of its inherent limitations. These issues include differences in oligomeric structure, degree of oligomerization, and glycan specificities within a group [108, 109]. There is no metal ion dependence for glycan binding [121], in contrast to many members of the C-type lectin family. Furthermore, galectins are not themselves glycosylated [121], which is due to their translation in the cytosol as opposed to translation at the endoplasmic reticulum and subsequent entry into the secretory system.

iii. General Features and Functions of Glycan Binding by Galectins

Galectin functions are numerous such as in regulation of immune responses, host-microbe interactions, cell biology and gene expression, cancer and metastasis, and cell adhesion. Despite these numerous functions, these proteins have not yet been found to be essential in mammals, implicating them as “analog dials” that fine-tune these physiological functions. Each galectin appears to have unique functions, a feature that likely partially occurs due to their differential glycan specificities (described in more

detail below), although overlap in certain galectin functions exists. More specific details on these galectin functions are presented in the following sections.

The majority of galectin protein produced by cells remains intracellular. Multiple intracellular subcellular localizations of galectins have been defined, including the cytoplasm, nucleus, and mitochondria [134-136]. In these compartments, galectins primarily exert their functions independently of glycan binding. Many intracellular galectin functions occur via protein-protein interactions, where the galectins may play a role in fine-tuning processes such as cell cycle progression, apoptosis, and vesicular trafficking [135-138]. For these reasons, galectin-glycan interactions are not a general feature of galectin functions in intracellular compartments.

For glycan binding to occur, the galectins must instead be localized extracellularly and present at sufficient concentration for binding to occur. However, all mammalian galectins lack any known signal sequences for export from the cell; thus, galectins must be exported by a nonclassical secretory mechanism(s) (proposed mechanisms of secretion are reviewed in [139]). It should be noted though that more studies are needed to clearly identify the mechanism(s) of secretion, particularly in human cells. Nevertheless, a number of studies have been published that directly show extracellular localization of galectins from cells, particularly from cell/tissue culture studies [139-143]. Furthermore, the degree of galectin externalization appears to vary not only between different cell types and differentiation stages but also within the same cell type under different environmental conditions (reviewed in [139]). Studies with CHO cells have shown that galectin-1 secretion occurs at a constant rate of ~3% of total synthesized galectin-1 per hour under basal culture conditions. Additionally, secreted galectins in the culture

medium, but not those that are cell surface bound, are mostly if not completely inactive (i.e.-incapable of binding lactosyl-Sepharose) [144]. Very few studies have quantitated the amount of galectins in tissues or cells [142, 145], and those that have did not look for the presence of cell surface and extracellular galectins *in vivo*. Therefore, the concentration of active extracellular galectins *in vivo*, including under different conditions and in different tissues, as well as the half-life of these secreted galectins in the oxidizing extracellular environment are still unclear.

The major effect of glycan binding by extracellular galectins is the activation, inhibition, or modulation of cellular signaling pathways. Due to the oligomeric nature of galectins, the binding of galectins to glycans on glycoproteins and/or glycolipids leads to specific receptor cross-linking, which is believed to be the major mechanism by which galectins activate cellular signaling pathways (reviewed in [146, 147]). Depending on the number of CRDs on the galectin, the degree of galectin oligomerization, and the number of glycan determinants on the glycoconjugate receptor (termed the *valency* of the galectin and the glycoconjugates), these galectin-glycoconjugate interactions may form large complexes, sometimes even leading to formation of large lattice-like structures [146]. The evidence for formation of these complexes is based on cell-free, solution-based binding methods, where galectins have been shown to form large aggregates when incubated at specific ratios with target glycoconjugates such as asialofetuin [148]. Additionally, a similar phenomenon is believed to occur at the cell surface based on cellular fluorescence microscopy and fluorescence resonance energy transfer (FRET) studies that show dispersed glycoconjugates receptors specifically coalescing into a few large assemblies at the cell surface in the presence of specific galectins [149, 150]. These

complexes are also highly stable based on FRAP studies with neutrophils [150]. Further support of the functional role of these multivalent complexes comes from studies of galectin mutants defective in oligomerization, which are less capable of inducing cellular signaling and hence the functions induced by the galectins [122, 151]. However, X-ray crystallography, cryo-EM, and/or other direct structural studies on these galectin-glycoconjugate complexes are still lacking and should be a topic for future analysis. Nonetheless, specific receptor cross-linking most likely drives galectin-mediated cellular signaling effects.

These galectin-glycoconjugate complexes may occur under basal conditions but new complexes may form, and/or basal complexes may become modified, when environmental changes occur. These environmental changes may result in changes in galectin expression, the presence of galectins from exogenous cells (ex-galectins from infiltrating leukocytes), and/or changes in glycoconjugate receptor glycosylation. Therefore, the actual signaling pathway(s) affected depends on the galectin(s) present at the necessary concentration for signaling; the glycosylation status of potential target glycoproteins, which may be at least partially mediated by cell type and differentiation status; environment status such as basal conditions, inflammation, epithelial injury, etc.; and other environmental cues [147, 152-156]. The end result of these galectin-glycoconjugate interactions and receptor cross-linking is activation of cellular signaling pathways. These signaling pathways result in changes in gene expression and ultimately functional effects such as changes in immune responses, leukocyte turnover, cancer, cell migration and wound repair, and host-microbe interactions (examples of these galectin-regulated signaling pathways are reviewed in [157] and described in the following

sections). Another potential effect of glycan binding by galectins is stabilization of cell-cell and/or cell-ECM interactions, which can occur independently of cellular signaling (reviewed in [158]).

It should be noted though that galectins are not essential for physiological functions, but rather act as "analog dials" to fine-tune these functions. Nonetheless, this fine-tuning is needed for homeostasis and proper physiological response to environmental insults since aberrant galectin expression or activity results in loss of this homeostasis and may even lead to diseases such as inflammatory and allergenic diseases, atherosclerosis, and cancer. In other cases though, loss or inhibition of specific galectins actually helps ameliorate these diseases, which may make galectins a potential therapeutic target for specific immune and allergenic diseases. For more details on specific functions of specific galectin family members, as well as the effects of aberrant galectin expression or activity, excellent reviews are available [135, 157, 159-161]. A general overview of galectin functions in humans is presented in the following sections, which specifically focuses on functions believed to result from extracellular galectin activity, galectin binding to specific glycans, and subsequent cellular signaling.

iv. Galectin Functions: Regulation of Immune Responses

One of the most well studied functions of galectin-glycan interactions is their regulatory effect on immune responses, which primarily occur via interactions with cells of the immune system. In fact, galectins have been shown to regulate functions in nearly all types of leukocytes [157], although it is still unclear if all these activities are due to galectin-glycan interactions. The major regulatory effects of the galectin-glycan

interactions include leukocyte turnover, leukocyte recruitment, and leukocyte activation and effector functions.

One of the first immune-regulatory functions described for galectins was their role in leukocyte turnover. A seminal study in 1995 showed that galectin-1 induces apoptosis of activated and leukemic T cells [162]. Although many studies to date have shown an apoptotic role of multiple galectins towards leukocytes and other cell types (reviewed in [163]), the physiological significance of galectins in inducing cellular apoptosis is quite controversial. The major reason for this controversy is that apoptosis has only been shown to occur under certain experimental conditions, specifically when a reducing agent is used in the medium [123, 124], artificial cellular expression or overexpression is used (examples of such reviewed in [163]), or physical forces such as vortexing or other shearing methods are applied to the cells (unpublished data). Thus, most of these studies of leukocyte apoptosis should be interpreted with caution. An exception to this rule is galectin-3, which was shown to induce apoptosis of activated T cells but not leukemic T cells in the absence of reducing agents and other potential confounders [125], but whether or not this effect occurs *in vivo* is still uncertain.

Nonetheless, more recent studies have indicated that galectins still regulate leukocyte turnover in the absence of these potential confounders, albeit by a different mechanism than apoptosis in most cases. Namely, galectin-1, -2, -3, -4, and -8 have been shown to induce reversible phosphatidylserine (PS) exposure in the absence of apoptosis in a number of leukocytes, especially activated neutrophils [123-125, 151, 164, 165]. This process requires galectin binding to specific cell surface glycans, such as polyLacNAc for galectin-8 binding [164], although the identity of the glycoconjugate

receptor(s) is unclear. This cell surface binding induces a signaling pathway that leads to activation of the Src kinases Lyn and Hck, which promote tyrosine phosphorylation of phospholipase C γ2 (PLCγ2) and subsequent increases in cytoplasmic Ca^{2+} concentrations [165], although more details on this signaling pathway remain unclear. An exception to this rule is galectin-4, which does not promote Ca^{2+} flux in activated neutrophils and hence likely induces PS exposure through a different mechanism [123]. This signaling pathway ultimately results in cell surface PS exposure, which was shown to promote neutrophil uptake and destruction by macrophages [151]. This process of PS exposure in the absence of apoptosis for removal by macrophages has been termed *preaparesis* [125], and preaparesis only occurs when specific galectins interact with specific cell types and under specific conditions. On the other hand, galectin-9 has been shown to prolong the lifespan of eosinophils in a lactose-dependent manner [166], although the receptors and signaling pathways involved in sending these survival signals are unclear. Therefore, galectins specifically regulate leukocyte turnover by either enhancing cellular survival or stimulating leukocyte removal from the circulation by preaparesis or apoptosis.

Galectins are also important for leukocyte recruitment to sites of infection or injury. One of the first studied activities was the eosinophil chemotactic activity of galectin-9. Activated and transformed T cells were known to produce a potent eosinophil chemotactic protein termed eclectin, which was later biochemically characterized by Matsumoto et al. in 1998 to be galectin-9 [141]. Galectin-9 was shown to be secreted by activated T cells as well as the T cell line from which eclectin was purified, which confirmed that this galectin was secreted from T cells to act in a paracrine manner

towards eosinophils. The expression of galectin-9 in peripheral blood mononuclear cells (PBMCs) isolated from a patient with an allergy was also increased upon stimulation with the specific allergen, which suggested a mechanism for potent eosinophil recruitment during allergenic responses. Moreover, galectin-9 promoted potent eosinophil but not neutrophil, monocyte, or lymphocyte migration *in vitro* and eosinophil-rich exudates when injected into mice, demonstrating the specificity of galectin-9 chemotactic activity [141]. Galectin-9-mediated recruitment of eosinophils occurred in a lactose-dependent manner [166], suggesting that cell surface glycan binding by galectin-9 mediated this eosinophil chemotactic activity, although the actual eosinophil glycoconjugate receptor(s) remains unknown. Similarly, galectin-3 was found to promote monocyte and macrophage chemotaxis in a lactose-dependent manner as well as monocyte recruitment when injected into mice [167]. The mechanism underlying this galectin-3-mediated chemotactic activity includes calcium influx into the cytoplasm as well as a pertussis toxin-sensitive pathway [167], but the actual glycoconjugate receptor(s) and specific signaling pathway(s) involved remain unknown. On the other hand, galectins may also block recruitment to sites of infection. For example, galectin-1 administration was shown to inhibit neutrophil rolling and adhesion to inflamed endothelium in a partially lactose-dependent manner, which may inhibit neutrophil extravasation into inflamed tissue [168]. Galectin-1 knockout mice also had increased neutrophil infiltration into inflamed tissues, which further supported a role of galectin-1 in inhibiting neutrophil tissue homing [168]. Therefore, galectins regulate leukocyte recruitment, especially when the immune system is activated during infection, allergy, and/or injury.

Galectins have also been shown to regulate leukocyte activation and effector functions. Galectin-3 increases the threshold for T cell activation and hence inhibits T cell activation until strong, specific antigen binding to the TCR occurs. This phenomenon is associated with T cell receptor (TCR) and CD45 binding via highly branched N-glycans, TCR sequestration from activation microdomains, CD45 sequestration at microdomains to antagonize TCR signaling, and overall reduced activation of the TCR signaling pathway [147, 155, 169]. Galectin-3 also promotes differentiation of macrophages to the "alternative" (M2) phenotype via glycan-specific binding to macrophage cell surface CD98 and activation of the PI3K signaling pathway [170]. Additionally, galectin-3 may promote mast cell activation and effector release [171, 172] while galectin-9 may inhibit mast cell activation and effector release [173], although these studies were primarily carried out with rat and mouse mast cells and thus the effect of these galectins on human mast cells remains unclear. Besides leukocyte activation, galectins also regulate leukocyte effector functions. Galectins regulate direct antimicrobial functions such as reactive oxygen species (ROS) production in eosinophils [166], monocytes [142], and neutrophils [174, 175] in a lactose-dependent manner. The glycoconjugate receptors and signaling pathways that initiate ROS formation in these cells are currently unknown, although galectin-3 was shown to bind a number of unidentified high molecular weight glycoproteins on human neutrophil membranes [175]. Moreover, galectins modulate cytokine production [125, 176-178], which induces specific immune responses and effector functions. Therefore, galectins play a key role in fine-tuning immune responses towards specific effector functions via regulation of leukocyte activation, cytokine and chemokine secretion, and antimicrobial functions.

v. Galectin Functions: Host-Microbe Interactions and Antimicrobial Functions

Galectins have been shown to regulate immune responses as described above, but can also do this by directly interacting with microbes to influence the infectious process. Galectins can either promote or resist infection and disease, a topic that has been reviewed previously [179-182]. Besides the immunoregulatory effects of galectins described in the previous section, galectins may influence infectious processes via a number of mechanisms, which involve galectin binding to either or both microbial or host glycoconjugate receptors. Galectin binding to host glycoconjugates may block the glycoconjugate receptors of microbes and thereby prevent microbial colonization. On the other hand, the binding of galectins to microbial glycoconjugates may result in inhibition of microbial attachment to host cells, enhanced phagocytosis via microbial aggregation, direct antimicrobial (cidal or static) activity, or "bridging" microbes to target cells to promote microbe infection (of target cells) or destruction (by phagocytes) (reviewed in [179, 180, 182]). At this point in time though, these proposed galectin-microbe interaction effects remain speculative or still in the experimental phases since only a few studies are available and/or the available studies lack direct or well-controlled support.

One function of galectins with direct and well-controlled experimental support, and which was only elucidated over the last 10 years, is antimicrobial activity. Antimicrobial activity of lectins was first described for the L-type lectin wheat germ agglutinin (WGA) in 1975 towards the fungus *Trichoderma viridae*, likely via binding to cell wall chitin [183]. Kohatsu et al. described the first antimicrobial activity of galectins when this group showed that human galectin-3 showed fungicidal activity towards *C. albicans*, but only in specific strains that appeared to express different β1-2 mannan

structures [184]. However, the fungicidal activity of galectin-3 towards this yeast was very weak, although not enough time points were used, and thus more studies are needed to confirm if the antimicrobial effect is cidal and to determine if the β1-2 mannan structures are directly bound by hGal-3. In addition, an antimicrobial effect of galectin-3 towards *S. pneumoniae* has been seen [185], which was likely a bacteriostatic effect, although specific experiments are needed to confirm this.

More recent studies have shown that galectins also show antimicrobial activity towards a number of other bacterial strains, including those expressing "self-like" glycan antigens, such as structures resembling the Blood Group B and Galα1-3Gal antigens, on LPS O-antigen or capsular polysaccharide. For this reason, human galectins are hypothesized to have evolved to target microbes expressing glycans that would otherwise evade adaptive immune responses [186, 187]. Interestingly, galectin binding to bacteria does not always lead to antimicrobial effects. For example, human galectin -3, -4, and -8 all bind the Blood Group B-expressing *E. coli* O86:B7 strain but only galectin-4 and -8 but not galectin-3 induce an antimicrobial effect [187]. However, galectin-3, -4, and -8 all show antimicrobial activity towards bacterial strains expressing terminal Galα1-3Galβ1-3- determinants on LPS such as that of *Providencia alcalifaciens* serotype O5 and *Klebsiella pneumoniae* serotype O1 [186]. This phenomenon may explain why some galectins may actually bind directly to microbial glycoconjugates, such as *Helicobacter pylori* LPS, and *promote* infection by possibly "tethering" the bacteria to the host cell surface as opposed to having an antimicrobial effect [188]. Alternatively, the galectins may only have antimicrobial effects towards specific pathogens. Therefore, galectin binding is necessary but not sufficient for antimicrobial activity, and other factors

contributing to the antimicrobial activity should be a topic for future studies. Moreover, the mechanism of galectin antimicrobial activity remains to be elucidated, which may be useful to understand since this may help elucidate novel aspects of microbial physiology and growth and/or uncover novel mechanisms of antimicrobial activity to design a new generation of antimicrobial drugs.

In conclusion, galectins appear bind microbial glycoconjugates and, in some cases, exert an antimicrobial effect against these microbes. Additionally, galectin binding to host and/or microbial glycoconjugates may also influence host-microbe interactions by other mechanisms. For these reason, it is of interest to determine if these effect help prevent or sometimes promote colonization and infection by microbial pathogens and/or regulate the gut microbiome. Additionally, the antimicrobial effects of galectins may be important in controlling autoimmune responses since the galectins may inhibit immune cell sampling and B cell antibody generation towards microbes expressing self-like antigens, thereby preventing generation of potentially self-reactive antibodies. One may hypothesize that lectins containing the so-called "jelly-roll motif," which includes the L-type lectins and galectins, were initially antimicrobial factors that later evolved to have a more diverse range of functions.

vi. Galectin Functions: Cell Biology and Cancer

Galectins can mediate a wide variety of cell biological processes that ultimately result in changes in cellular signaling, gene expression, and cellular functions. This wide range of regulatory functions mainly results from the ability of galectins to bind and cross-link glycoconjugate receptors. Some of the downstream cellular functions mediated

by these galectin-glycoconjugate interactions include cellular differentiation, cell cycle regulation, cell migration, metastasis, and tumor progression.

Galectins have been shown to regulate cellular differentiation and proliferation, especially towards cells of the immune system (refer to the *Galectin functions: Regulation of Immune Responses* section). However, galectins may also regulate differentiation and proliferation of other cell types as well. In fact, myocytes (muscle cells) were one of the first cell types whose differentiation was suggested to be at least partially regulated by galectins. Myoblasts express very high levels of galectin-1 but fail to secrete this galectin until terminal differentiation [189]. Upon terminal differentiation, the secreted galectin-1 inhibits adhesion of differentiating myocytes to the substratum and thereby reduces the ability of the cells to fuse into myotubes [190]. This process might be important for fine-tuning the rate or degree of myocyte proliferation and myofiber development. Besides myocytes, galectin-3 has also been shown to have an inhibitory effect on osteoblast (bone cell) differentiation, which is associated with binding to the Notch receptor and activation of the Notch signaling pathway [191]. On the other hand, galectin-9 promotes osteoblast differentiation, which is associated with galectin-9 binding to CD44 on osteoblasts and activation of the Smad1/5/8 signaling pathway [192]. However, the dependence of galectin-9 binding to glycone or aglycone components on CD44 in this process was not tested and remains unclear. Interestingly, galectin-9 not only regulates osteoblast differentiation but also osteoblast proliferation via lipid raft clustering and activation of the c-Src/ERK signaling pathway in a lactose-dependent manner [193], although the receptor(s) involved in galectin-9-mediated lipid raft clustering is unclear. Galectin-1, -3, and -4 also slightly promoted osteoblast

proliferation but were not as potent as galectin-9 [193]. Additionally, galectin-3 stimulates proliferation of preadipocytes in a lactose-dependent manner [194], but the mechanism was not tested and is currently unknown. This galectin-3-mediated preadipocyte proliferation may be important in obesity since galectin-3 levels are upregulated in preadipocytes in an mouse model of obesity, which may result in increased accumulation of adipose (fat) tissue [194]. Galectins may thus serve as regulators of cellular differentiation and proliferation depending on the cell type and galectin(s) involved. This regulation of cellular differentiation and proliferation may not only be important for development but also cancer progression, as described in more detail below.

Galectins have also been shown to regulate epithelial cell migration. Immobilized galectin-8 was shown to bind epithelial cells and promote cellular attachment and migration across the substratum. The underlying mechanism involved galectin-8 binding to specific epithelial cell surface β_1 integrins, activation of integrin signaling pathways, and tyrosine phosphorylation of cytoskeletal proteins [195]. However, the physiological significance of this interaction is currently unclear. Another functional role of galectin-glycan interactions in epithelial cell migration is wound healing. Galectins promote epithelial wound repair in a glycan-dependent manner, which is mediated by cell migration and cytoskeletal reorganization with or without stimulation of cell cycle progression and cellular proliferation (reviewed in [196]). Besides epithelial cells, galectins can also regulate endothelial cell migration and thus the angiogenic process, which is commonly initiated by galectin binding to vascular endothelial growth factor (VEGF) receptors and activating VEGF signaling pathways (reviewed in [197-199]).

Therefore, galectin-glycan interactions may regulate cell migratory processes in epithelial and endothelial cells via receptor binding and cellular signaling, which is important in physiological processes such as wound repair and angiogenesis.

Due to their numerous fine-tuning functions in cell biology, it is no surprise that alterations in galectin expression and/or localization can lead to defects in physiological processes and result in diseases such as cancer. The expression of a galectin in a cell that normally does not express that galectin may also be detrimental to cell biology and lead to formation and/or increase the aggressiveness of cancerous cells. For example, galectin-7 was the most highly upregulated gene expressed in an aggressive T lymphoma cell clone vs. non-aggressive T lymphoma cell clone that were both derived from the same parental cell line [200]. Normally, galectin-7 is almost exclusively expressed in stratified epithelia such as the skin [201], so expression of galectin-7 in T cells is an aberration. Further support of a role of galectin-7 in this aggressive phenotype was established from *in vivo* studies where the non-aggressive T lymphoma cell line was injected into mice. Three tumor cell clones were isolated after just three *in vivo* passages, and all three clones now overexpressed galectin-7 and were more aggressive and lethal than the non-aggressive parental lymphoma cells [200]. This finding was also corroborated by ectopic expression of galectin-7 in a non-aggressive T lymphoma cell line, which led to a more aggressive phenotype *in vivo* [202]. Demers et al. suggested that the mechanism underlying this transformation might involve galectin-7-mediated upregulation of matrix metalloproteinase-9 (MMP9) expression and secretion in a glycan-dependent manner [202]. MMP9 is also associated with cancer cell invasion and metastasis [203], but the role of MMP9 in promoting this more aggressive phenotype was not directly tested in the

Demers et al. study. Nevertheless, these studies demonstrate how aberrant galectin expression may lead to cancer development and/or progression.

Galectin-glycan interactions are known to regulate tumor formation, the tumor microenvironment, and cancer cell aggressiveness and metastasis. One potential mechanism underlying galectin regulation of tumorigenesis is the effect of galectins on cell cycle progression, which was briefly described above as a mechanism for galectins promoting cell proliferation. One study showed that galectin-1 inhibits cell cycle progression and proliferation of carcinoma cell lines in a glycan-dependent manner [204], although relatively high concentrations of galectin-1 (at least 25μg/ml) are needed for just a slight effect. Many other regulatory effects of galectins on the cell cycle are known, but these occur via intracellular, lactose-independent mechanisms (reviewed in [205]). Thus, the importance of galectin-glycan interactions in cell cycle regulation is still unclear and should be a topic of future study. Besides dysregulation of cell cycle progression, galectins can also alter the tumor microenvironment. Some galectins have been shown to be pro-angiogenic factors for tumors (reviewed in [197-199]), and hence galectin-glycan interactions may promote capillary re-modeling to help "feed" tumors with additional oxygen and nutrients to support their growth and development.

Besides their role in tumor formation, galectin-glycan interactions regulate metastasis, the spread of cancer cells to secondary tissues. One of the first steps in metastasis is the conversion of tumor cells to invasive tumor cells. This transition can occur through multiple processes, one of the most well known being epithelial-mesenchymal transition (EMT), a process whereby differentiated cells begin to take on the properties of mesenchymal (less differentiated) cells, including reduced cellular

adhesion molecule expression and hence a higher potential of cell layer detachment and migration. Galectin-1 has been shown to be highly upregulated in squamous carcinoma cells that have undergoing SNAIL-induced EMT, and soluble galectin-1 caused increased cellular migration, Matrigel invasion, and frequency of cells that have undergone EMT [206]. The mechanism of these galectin-1-mediated effects is likely due to increased JNK signaling and subsequent upregulation of α_2, β_1, and/or β_5 integrins [206], although the cell surface receptor(s) bound and other details of the signaling pathway are currently unknown. Galectin-3 may also inhibit anoikis (cellular death induced by loss of anchorage, which is considered a major mechanism of preventing cancer cell metastasis) by promoting cancer cell aggregation [207]. The authors proposed that galectin-3-mediated cell aggregation may serve as a surrogate mechanism of cellular "anchorage," thereby allowing cancer cells to escape from tumors without undergoing anoikis and thus survive in the circulation and secondary tissues [207]. High concentrations (~62.5μg/ml) of galectin-3 may also inhibit anoikis independent of cellular aggregation [208], although the physiological relevance of such a high concentration of galectin-3 on anoikis is unclear. On the other hand, galectin-1 may promote anoikis, so the functional outcome of anoikis may be dependent on a "battle" between galectin-1 and galectin-3 expression and concentration at the site of the tumor [208]. Cancer cell-endothelial cell adhesion is important for cancer cell extravasation into secondary tissues, and this process may also be modulated by a number of galectins. Galectin-3 may promote increased cancer cell-endothelial cell adhesion, which is mediated by increased pro-inflammatory cytokine expression (IL-6, GM-CSF, G-CSF, and sICAM-1) and subsequently increased cell surface adhesion expression by endothelial cells [209]. Other studies suggest that

increased expression of galectin-3 in carcinoma cells is associated with increased rolling on and adherence to endothelial cells in a lactose-dependent manner [210], while exposure of carcinoma cells to galectin-2, -4, or -8 increases adherence to endothelial cells [211]. On the other hand, galectin-9 was shown to act as an anti-metastatic factor by reducing cancer cell-ECM adhesion and cancer cell binding to endothelial cell adhesion molecules in a lactose-dependent manner [212]. Support of the anti-metastatic role of galectin-9 was demonstrated in a mouse model of metastasis, where mice injected with recombinant galectin-9 or cancer cells recombinantly expressing galectin-9 showed substantially reduced numbers of pulmonary tumor colonies [212]. In order to escape the primary tissue as well as enter the secondary tissue, cancer cells need to make their way through connective tissue and the extracellular matrix. Degradative enzymes, especially matrix metalloproteinases (MMPs), play a key role in this process. Specific galectins have been shown to activate MMP gene expression [202], as well as accelerate pro-MMP processing into the active enzymes [213]. Thus, galectins may indirectly promote tumor escape and secondary tissue entry by regulating MMP expression or activity. Therefore, galectins may regulate the process of cancer cell metastasis at multiple stages from cell transformation to tissue invasion. For additional information on the roles of specific galectins in cancer and metastasis, intensive and extensive reviews are also available [205, 214].

vii. Galectin Functions: Cell-Cell and Cell-ECM Interactions

One of the first functions described for galectins was their role in adhesion, including cell-cell adhesion and cell-extracellular matrix (ECM) adhesion. A large

number of *in vitro* and cell culture studies have been conducted demonstrating glycan-dependent binding of galectins to cells and ECM glycoproteins including fibronectin, collagen, and laminin (reviewed in [158]). Despite this seemingly crucial role in physiology and morphogenesis, *in vivo* studies using galectin knockout mice revealed no gross abnormalities or defects in anatomy, physiology, development, or fertility [215-217]. For this reason, galectin-mediated adhesion processes are not essential but rather are more likely involved in fine-tuning cell-cell and cell-ECM interactions. Some key points about galectin-glycan interactions in adhesion are pointed out below.

Galectin-glycan interactions may promote or inhibit adhesion between cells or between cells and the ECM. The positive or negative effect on adhesion depends on the galectin(s) expressed, galectin conformation, cell type, cell surface receptor or ECM protein expression, glycosylation status, and environmental cues and conditions. For this reason, even the same galectin under different conditions may lead to different outcomes. For example, soluble galectin-8 was shown to inhibit cell adhesion to ECM components and tissue culture-treated plastic surfaces in a lactose-dependent manner, whereas immobilized galectin-8 actually promoted cell adhesion in a lactose-dependent manner [195]. Therefore, galectins presentation as soluble or surface-bound may dictate the adhesive effects of galectins. Notably, the inhibitory effect of soluble galectin-8 towards laminin binding by the 1299 pulmonary carcinoma cell line was specific to galectin-8 because galectin-1 and galectin-3 did not cause this inhibitory effect [143], demonstrating how specific galectins may mediate specific adhesion events. The cell type may also influence the effects of adhesion. For instance, galectin-1 promotes adhesion of human melanoma cells to the ECM glycoprotein laminin [218] but inhibits myocyte binding to

laminin [190]. Therefore, even the same galectin may have different adhesion effects on different cell types, even to the same substrate. A possible mechanism for this effect is that the oligomeric galectins may promote cell-ECM adhesion by binding both ECM components and cells, thereby "bridging" the two. However, this "bridging" effect may also prevent direct receptor-ligand contact and, in cases where such direct contact is needed for functional adhesion and/or cellular signaling for firm adhesion, this "bridging" effect may actually block or inhibit firm adhesion. Galectins may also inhibit cell-ECM adhesion by ECM binding without cell binding to block specific cell-ECM interactions. Other mechanisms are also possible (reviewed in [158]).

The mechanism of galectin-mediated adhesion may not only directly involve galectin-glycan interactions to help cells "stick" to other cells and ECM components but also regulation of signaling pathways induced by galectin-glycan interactions [158]. For example, galectin-8-mediated cell adhesion was due to galectin-8 binding to specific integrins [195]. This glycan-dependent galectin-8-integrin interaction led to integrin signaling pathway activation, resulting in phosphorylation of cytoskeletal proteins (including FAK and paxillin) that may explain the cytoskeletal rearrangements and cell spreading that occur upon cell attachment to galectin-8 [195]. However, whether integrin signaling or integrin binding and adhesion was the major mechanism underlying galectin-8-mediated cell adhesion was not tested and hence is still unclear. Nonetheless, these results suggest that galectin-glycan interactions may mediate adhesion by a number of mechanisms besides merely serving as "glue" or a "bridge" between cells and between cells and the ECM.

Galectin-glycan interactions may also be important in maintaining a tight barrier within epithelial cell layers. A recent study by Jiang et al. demonstrated that galectin-3 interacts with the desmosomal protein desmoglein isoform 2 (Dsg2) via Dsg2 N-glycans on human intestinal epithelial cell lines [219]. Lactose treatment, anti-galectin-3 antibody treatment, or siRNA knockdown of galectin-3 in human intestinal epithelial cells all caused reduced intercellular adhesion and dispersion the epithelial cell layer. Besides serving as a bridging mechanism, the galectin-3-Dsg2 interaction also promoted adhesion by maintaining Dsg2 cell surface expression. In the absence of galectin-3, Dsg2 was more readily degraded by the proteasome, most likely by internalization of cell surface Dsg2. The physiological relevance of this galectin-3-Dsg2 interaction in intestinal epithelium integrity was demonstrated by the use of isolated murine small intestines; the intestine was treated with anti-galectin-3 antibody, which led to a reduction in cell surface Dsg2 levels [219]. However, no studies were conducted on the integrity of this *ex vivo* intestine, such as by transepithelial resistance (TER) measurements, upon anti-galectin-3 or lactose treatment. Thus, the actual physiological contributions of galectin-3 to intestinal epithelium integrity and barrier function were not tested in *ex vivo* or *in vivo* experiments. Nonetheless, this study suggests a role of galectins in forming tight intercellular contacts within intestinal epithelial cells. It will be interesting to determine if other galectins play a role in intercellular adhesion in other simple epithelial layers such as the respiratory tract. Despite very early knowledge of galectin-glycan interactions in adhesion, this recent study was one of the first to strongly suggest an important physiological function of galectin-glycan interactions in epithelium adhesion and integrity. Moreover, this study leads to questions as to whether dysregulated galectin-3

expression or Dsg2 target glycan ligands contribute to diseases associated with reduced epithelial barrier function such as food allergies and inflammatory bowel diseases. Future studies should also be aimed at determining if galectin-3 knockout mice have reduced intestinal epithelial cell integrity and/or if other intestinal galectins such as galectin-4 regulate the integrity of the intestinal epithelial barrier.

III. Glycan-Binding Proteins of Dendritic Cells

i. *Background: Dendritic Cells and Dendritic Cell Glycan-Binding Proteins*

Dendritic cells (DCs) are a central player in the immune system, especially in their ability to bridge the innate and adaptive immune systems as well as act as phagocytes and pathogen sensors (reviewed in [220]). DCs are typically classified into two categories: classical (myeloid) DCs (cDCs) and plasmacytoid DCs (pDCs). pDCs are involved in pathogen sensing and production of Type I interferons, primarily through interactions with microbial nucleic acids via Toll-like receptor (TLR)-7 and -9 [221]. On the other hand, the cDCs are professional antigen-presenting cells that present antigens on MHC Class I or II molecules to activate $CD8^+$ cytotoxic T cells and $CD4^+$ T helper cells, respectively. Most cDCs exist as "immature" cDCs prior to stimulation and exhibit much phagocytic activity. Once stimulated by PAMPs such as LPS binding to TLR4, cDCs undergo differentiation into mature cDCs. Mature cDCs have reduced phagocytic activity but enhanced migratory activity and T cell stimulatory activity, so the mature cDCs cells are the DC subtype that activates T cells. These mature cDCs have increased cell surface expression of costimulatory molecules (especially CD80 and CD86) and cytokines that are critical for promoting activation and differentiation of T cells (reviewed in [220]). Depending on the type of microbe or microbial products encountered as well as other environmental and immunological stimuli, cDCs mature in different ways to promote T cell differentiation into different effector subtypes, such as $CD4^+$ T cell differentiation into Th1, Th2, Th17, T_{reg}, or other subtypes (reviewed in [222]). This maturation mechanism also depends on the pathogen recognition receptors (PRRs) expressed on the cDC surface, especially TLRs and C-type lectin receptors (CLRs),

which may vary between different cDC types. For example, the cDCs in the epidermis, known classically as Langerhans cells, express the CLR Langerin but not DC-SIGN [220]. For more information on DCs, interactions of DCs with pathogens and non-pathogen products, other functions of pDCs and cDCs, and functional variations between different DC subsets, excellent reviews are available [220-224].

DCs are very abundant in the intestine, where a variety of different DC subtypes are localized in the Peyer's Patches (PPs) and lamina propria (LP) (reviewed in [225]). These localizations and variety of DC subtypes may be important for sampling the numerous, diverse types of microbes that pass through the intestine to generate the appropriate immune responses to specific microbes, including differentiation of commensal vs. potentially parasitic organisms. Commensal and pathogenic microbes can pass through M cells in the intestinal epithelium and especially follicle-associated epithelium and then be sampled by DCs in the underlying LP and PPs, respectively [225]. Specific DC subsets have also been shown to directly migrate from the LP into the small intestinal lumen, although these cells do not appear capable of re-entering the LP and hence migrating to lymph nodes to activate T cells [226]. Thus, these intestinal lumen DCs may solely function in innate immune mechanisms such as phagocytosis to prevent pathogen dissemination from the intestinal lumen into the underlying tissue and circulation [226].

Interestingly, DCs have been shown to extend their dendrites through the small but not large intestinal epithelium from the LP to "sample" microbes in the intestinal lumen without perturbing epithelium integrity or completely entering the intestinal lumen [227, 228]. A more in-depth examination of this process in mice by Chieppa et al.

demonstrated that transepithelial dendrite extension is prominent in the apical (villus tip) but not basal (crypt) regions of the intestinal epithelium. Additionally, transepithelial dendrite extension occurs most frequently the jejunum and proximal ileum but not terminal ileum or cecum under basal conditions, although exposure to gastrointestinal pathogens like *Salmonella* increases the frequency of transepithelial dendrite extension into the lumen in the terminal ileum. This dendrite extension also appears to be dependent on the presence of an intact gut microbiota since antibiotic treatment strongly reduced the number of dendrite extensions in the proximal ileum of mice. The mechanism for microbe-induced transepithelial dendrite extension is specific Toll-like receptor 2 (TLR2), TLR4, and/or TLR9 recognition of target ligands and subsequent signaling via the MyD88 pathway by intestinal epithelial cells [229]. However, the mechanism by which TLR-stimulated epithelial cells send signals to dendritic cell to promote dendrite extension is unclear but is likely due to cytokine and/or chemokine production. Indeed, CCL20 has been suggested as a mechanism due to the ~20-fold upregulation of this gene in the terminal ileum in *Salmonella*-infected vs. uninfected mice, which was dependent on MyD88 expression [229]. However, this suggestion has yet to be confirmed through the use of CCL20 knockout mice. More recently, dendrite extension from specific DC populations in the Peyer's Patches through the follicle-associated epithelium has been also demonstrated. In contrast to transepithelial dendrite extension, dendrite extension occurred directly *through* M cells in the follicle-associated epithelium via so-called transcellular pores rather than *between* cells as is the case of dendrite extension between enterocytes from DCs in the LP [230]. Therefore, this mechanism of dendrite extension is not limited to LP DCs, although the mechanism of

dendrite extension varies depending on the cell type through which the dendrites will pass (paracellular for enterocytes vs. transcellular for M cells). Together, the studies suggest that DC sampling of microbes in the intestinal lumen can occur via transepithelial dendrite extension from DCs in the LP and PPs. A major question that remains to be addressed though is whether or not antigen sampling by dendrite extension is sufficient for DC migration into the lymphatic system to present these antigens to $CD4^+$ T cells, a question that has been hindered by technical limitations [229]. Since these studies were all performed using animal models, an additional question that remains unanswered is whether or not dendrite extension through the intestinal epithelium also occurs and is a common phenomenon in the human intestine. Additionally, the relative contribution of this dendrite extension mechanism to DC-mediated immune responses is still unclear and requires further study.

DCs utilize a number of cell surface receptors to recognize microbes and tailor DC-mediated immunity, including Toll-like receptors (TLRs) as well as the C-type lectin, Siglec, and galectin families of glycan-binding proteins. TLR binding to specific PAMPs, such as LPS to TLR4, is one of the major mechanism promoting DC maturation [231]. On the other hand, C-type lectins, Siglecs, and galectins do not typically promote DC maturation but rather play roles in endocytosis for antigen presentation and/or regulation of cellular signaling to “fine-tune” the ultimate immune response by the mature DCs. The latter function may be especially important for effector functions such as promoting proper T helper cell differentiation and inducing T cell activation or tolerance (reviewed in [223, 232]). In fact, it has been suggested that more efficient vaccines or therapeutics

may be produced via specific targeting of C-type lectins to generate specific T cell responses that would be most effective at pathogen clearance [232].

DCs are known to express galectin-1, -3, and -9 [233], which may play roles in DC-mediated immune responses (reviewed in [182, 223]). For more information on galectins, refer to the *Galectins* section of this chapter. The structures, ligands, and functions of specific C-type lectins and Siglecs expressed by DCs are discussed in the following sections.

ii. C-Type Lectins: DC-SIGN, Langerin, MGL, and Dectin-2

C-type lectins are classically defined as a family of transmembrane and secreted glycan-binding proteins that bind glycans in a calcium-dependent manner via a highly conserved domain (the C-type lectin domain, CTLD) [234, 235]. However, examples of calcium-independent C-type lectins (ex-Dectin-1) are now known [236], so calcium-dependence is not an absolute requirement for glycan binding by C-type lectins. Dendritic cells express a large number of membrane-bound C-type lectins that play a role in endocytosis of bound glycoconjugate ligands or regulation of downstream signaling pathways upon glycoconjugate ligand binding. These membrane-bound C-type lectins are thus referred to C-type lectin receptors (CLRs). Within all subsets and differentiation stages of DCs, at least nine different CLRs are known expressed [237], many of which play a key role in DC-mediated recognition of specific microbes and tailoring of specific immune responses towards these specific microbes (reviewed in [238]). The major focus here will be on the CLRs DC-SIGN, Langerin, MGL, and Dectin-2 due to their use in this

study. For information on the other DC-expressed CLRs, excellent reviews are available [232, 238-240].

One of the most well studied CLRs of DCs is Dendritic Cell-Specific Intercellular Adhesion Molecule-3 Grabbing Nonintegrin (DC-SIGN). DC-SIGN was first discovered in 1992 by Curtis et al. during a study on CD4-independent mechanisms of HIV-1 binding to human cells [241]. Using expression cloning of a human placental cDNA library into COS-7 cells, the authors identified a clone that promoted binding of cells to a recombinant version of the HIV-1 gp120 cell adhesion molecule. DNA sequencing of this clone revealed an ORF that encoded a protein with significant homology to transmembrane C-type lectins. This protein was functionally confirmed to be a C-type lectin by showing that gp120 binding was calcium-dependent and that mannose, fucose, and especially mannan specifically inhibited binding of gp120 to cells expressing this C-type lectin [241]. These results suggested that this C-type lectin was a mannose-binding and possibly fucose-binding lectin. At the time, the human cells or tissues expressing this C-type lectin other than placental tissue were unknown.

In 2000, Geijtenbeek et al. demonstrated expression of this C-type lectin in a specific cell type, DCs, in two complementary studies. The first study identified a protein that promoted DC-T cell complexes via interaction with Intercellular Adhesion Molecule-3 (ICAM-3) on T cells and subsequent T cell proliferation. Using DC-T cell adhesion-blocking monoclonal antibodies, the authors identified a 44kDa protein that, by peptide sequencing and gene cloning, was identical to the C-type lectin identified in the 1992 Curtis et al. study. Since this C-type lectin was also specifically expression on DCs but not other leukocytes, the lectin was aptly named DC-specific ICAM-3-grabbing

nonintegrin (DC-SIGN). The function of DC-SIGN as an ICAM-3-binding molecule and as a C-type lectin was confirmed by recombinantly expressing DC-SIGN in COS-7 cells and demonstrating ICAM-3 binding, which required calcium and was inhibited by mannan and mannose [242]. Therefore, this first study suggested that DC-SIGN was a cell adhesion molecule that bound ICAM-3 via DC-SIGN's CRD. The second study, which gained DC-SIGN much notoriety, demonstrated that DC-SIGN on DCs or recombinantly expressed in other cells could bind to gp120 and HIV-1 virions via the DC-SIGN CRD. More importantly, DC-SIGN binding did not lead to viral entry into the DCs or other cells recombinantly expressing DC-SIGN but rather allowed DCs to present these viruses to $CD4^+$ T cells. Moreover, DC-SIGN-mediated HIV-1 virion shuttling increased the length of time (from ~1 day to ~4 days) that HIV-1 virions efficiently infected $CD4^+$ T cells relative to the absence of cells expressing DC-SIGN [243]. Therefore, DC-SIGN-mediated *trans*-infection gained much attention as a potential mechanism underlying the early stages of the HIV-1 infectious process due to the presence of DC-SIGN-positive DCs in the lamina propria of mucosal tissue initially exposed to HIV-1 and the ability of DC-SIGN to prolong the lifespan of infectious virions [243]. DC-SIGN is also expressed by cells in the human small intestine, particularly the Peyer's Patches [244], so DC-SIGN may come into contact with intestinal lumen contents such as HMGs in breast-fed infants.

Studies on glycan specificity of DC-SIGN have confirmed the initial observation by Curtis et al. that DC-SIGN binds mannosylated and fucosylated glycans. Besides mannan, DC-SIGN was shown to bind high mannose N-glycans via a tetrameric extracellular CRD, which was suggested to be the major mechanism by which DC-SIGN

binds to the high mannose-rich HIV-1 gp120 glycoprotein [245]. Other studies revealed that DC-SIGN binds to specific fucosylated glycans, namely those containing Lewis determinants, which include the Lewis a (Galβ1-3(Fucα1-4)GlcNAcβ-), Lewis b (Fucα1-2 Galβ1-3(Fucα1-4)GlcNAcβ-), Lewis x (Galβ1-4(Fucα1-3)GlcNAcβ-), and Lewis y (Fucα1-2 Galβ1-4(Fucα1-3)GlcNAcβ-) glycan determinants [246]. In fact, DC-SIGN was shown to bind the heavily fucosylated *Schistosoma mansoni* egg antigen (SEA) specifically via Lewis x determinants and structurally-related LDNF (GalNAcβ1-4(Fucα1-3)GlcNAcβ-) determinants on SEA [247]. Moreover, DC-SIGN was found to be a major CLR involved in DC binding to SEA (along with MGL and the mannose receptor), suggesting that DC-SIGN is a major receptor for SEA and thus may at least partially underlie the pathological immune responses triggered by SEA [248]. X-ray crystallography studies with Lewis x determinants and high mannose N-glycans have revealed key contacts with specific amino acids and a coordinated calcium ion as well as similarities and differences in the amino acid residues involved in DC-SIGN binding to the Lewis x trisaccharide vs. the $Man_9GlcNAc_2$ high mannose N-glycan. Therefore, DC-SIGN is a calcium-dependent glycan-binding protein that binds to mannan, high mannose N-glycans, and Lewis glycan determinants, which may explain the variety of endogenous and exogenous ligands and microbes recognized by DC-SIGN (reviewed in [232, 249]).

The major functional consequences of DC-SIGN binding by ligands are endocytosis and/or regulation of signaling pathways in DCs. DC-SIGN binding to multivalent ligands such as glycoproteins or microbial cells carrying DC-SIGN-specific mannosylated or fucosylated determinants leads to endocytosis of the cargo [247, 248, 250-252]. DC-SIGN-mediated endocytosis has also been shown to result in cargo

trafficking through the endocytic pathway and subsequent co-localization with MHC Class II molecules [248]. Therefore, DC-SIGN-mediated endocytosis may be important for presentation of specific glycan or glycoconjugate antigens on MHC Class II for activation of $CD4^+$ T helper cells. However, not all multivalent ligands bound by DC-SIGN are endocytosed, such as HIV-1 virions. For this reason, more studies are needed on the physical and biochemical factors that promote DC-SIGN-mediated endocytosis.

DC-SIGN also functions as a modulator of DC signaling, particularly upon TLR activation. In DCs stimulated with LPS, mannosylated ligands and microbes containing mannosylated ligands such as the tuberculosis-causing pathogen *Mycobacterium tuberculosis* inhibited DC costimulatory molecule upregulation and caused increased IL-10 production via binding to DC-SIGN [251]. Paradoxically though, mannosylated ligands also promote increased cytokine expression, including increased expression or production of pro-inflammatory cytokines such as IL-6 and IL-8 as well as the Th1-biasing cytokine IL-12 [253]. This mechanism of increased cytokine production by mannosylated ligands involved Raf-1 phosphorylation in a Src kinase, PAK kinase, and Ras-dependent manner. Interestingly, Raf-1 phosphorylation led to activation of an alternative Raf-1 signaling pathway, which resulted in acetylation of the p65 subunit of NF-κB [254]. This DC-SIGN-dependent interaction prolonged and increased expression of genes encoding cytokines including IL8, IL-10, and IFN-β relative to LPS stimulation alone [254]. Moreover, a KSR1–CNK–Raf-1 complex as well as Lsp1, RhoA, and LARG were required for activation of this alternative Raf-1 signaling pathway [253], although more details of the pathway upstream and downstream are unclear. Additionally, the mechanism by which mannosylated ligands inhibit DC maturation, including CD80 and

CD86 expression, is still unclear. Thus, mannosylated ligands may regulate TLR4 (and possibly other TLR) signaling pathways in DCs by promoting a more pro-inflammatory and Th1-biased response but reduced costimulatory molecule expression and DC maturation. These regulatory effects suggest that DCs may take on a more of an innate immunity leukocyte role rather than a role as a professional antigen-presenting cell upon DC-SIGN binding to mannosylated ligands in the presence of TLR ligands.

In contrast to mannosylated ligands that increased cytokine expression, fucosylated ligands (specifically multivalent Lewis x-polyacrylamide beads, Le^x-PAA) actually reduced cytokine production by promoting dissociation of KSR1–CNK–Raf-1 complexes from DC-SIGN-Lsp1 [253], suggesting an anti-inflammatory role of fucosylated ligand binding to DC-SIGN. Moreover, fucosylated ligands uniquely bias DCs to promoting Th2 differentiation at least partially via DC-SIGN binding [248]. Recently, the mechanism underlying Th2-biasing by fucosylated ligands (including Le^x-PAA, SEA, and *Helicobacter pylori* strains expressing Lewis x antigen on LPS) was uncovered and determined to occur via Ser252 phosphorylation of Lsp1 by the MK2 kinase that is likely induced by the absence of a bound KSR1–CNK–Raf-1 complex. This Lsp1 phosphorylation event caused activation of an IKKε- and CYLD-dependent signaling pathway that led to the formation of an alternative Bcl-3-p50 NF-κB complex. This signaling pathway ultimately resulted in a Th2-biased cytokine profile by DCs, including upregulation of IL-10 and downregulation of IL-6 and IL-12 [255]. Additionally, DC-SIGN binding to fucosylated ligands and subsequent Lsp1 phosphorylation by fucosylated ligands also promoted activation of the ISGF3 signaling pathway, resulting in increased IL-27 secretion by DCs [256]. This increased IL-27

production promoted $CD4^+$ T helper cell differentiation into the follicular T cell phenotype (T_{FH}), which are T cells that promote and maintain germinal centers and strongly promote B cell activation, maturation, and isotype switching from IgM to IgG [256]. Therefore, fucosylated ligand binding to DC-SIGN on TLR-activated DCs induces a more anti-inflammatory response as well as T helper cell differentiation into both Th2 and T_{FH} phenotypes, which are the two key T cell phenotypes involved in humoral immunity.

Langerin is another CLR that is expressed by some subsets of DCs, especially DCs in the epidermis called Langerhans cells (LCs), but also in DCs localized in other stratified epithelia and other tissues such as the lung (reviewed in [257]). Interestingly, Langerin expression may replace DC-SIGN in LCs since this DC subset does not appear to express DC-SIGN [258]. This phenomenon suggests that different DC subtypes may have evolved to express different CLRs depending on anatomic localization, which dictates what endogenous and exogenous ligands such as microbes the DCs will encounter and respond to. Langerin was first characterized as CD207, an LC-specific cell surface marker, and hence was given the name Langerin [259]. Langerin was determined to associate with intracellular Birbeck granules characteristic of LC morphology, and transfection of Langerin into fibroblasts led to formation of Birbeck granules, suggesting that Langerin plays a major role in the presence and formation of Birbeck granules [260]. Additionally, Langerin was shown to encode a protein with homology to a C-type lectin, and was indirectly determined to require calcium for lectin activity and bind to mannan [260]. Finally, Langerin was shown to be endocytosed upon ligation with anti-Langerin antibodies or upon binding to glycoproteins, which was followed by Langerin

localization to Birbeck granules and degradation of the endocytosed cargo [260, 261]. These findings suggest that Langerin functions as an endocytic receptor that targets cargo for lysosomal degradation, although whether or not Langerin promotes antigen presentation onto MHC Class II molecules like DC-SIGN does not yet been demonstrated. However, Langerin has been shown to be essential for antigen presentation on CD1a, a receptor involved in presentation of lipid antigens [262]. The antigen loading is believed to occur in the Birbeck granules where both CD1a and endocytosed Langerin can localize [262], although no direct proof of this was presented. Interestingly, Langerin also promotes LC phagocytosis and degradation of HIV-1 virions [263], in stark contrast to DC-SIGN binding of HIV-1. Although Langerin has been shown to promote microbial phagocytosis and some forms of antigen presentation, to date no studies of a role of Langerin in the regulation of cellular signaling have been performed.

Structurally, Langerin was shown to form trimers via the interactions of α-helices in the extracellular neck domain to form coiled-coils. Of a panel of monosaccharides, Langerin preferentially binds to Man, Fuc, and GlcNAc with low millimolar K_i values. However, Langerin bound more strongly to high mannose N-glycan structures ($Man_{5-9}GlcNAc_2$), but the K_i for the $Man_9GlcNAc_2$ interaction was still relatively high for the monomeric CRD (0.23mM). This low affinity binding may be compensated for by higher avidity via trimerization; this is supported by the fact that glycoproteins containing high-mannose N-glycans are only bound by trimeric but not monomeric recombinant Langerin constructs [261]. CFG glycan microarray analysis of Langerin revealed specific binding to not only high mannose N-glycans but also structures terminating in β-linked GlcNAc such as the complex N-glycan structure GlcNAcβ1-2Manα1-3(GlcNAcβ1-2Manα1-

3)Manβ1-4GlcNAcβ1-4GlcNAcβ-, structures containing 6'-O-sulfated galactose such as Gal6Sβ1-4GlcNAcβ- and Gal6Sβ1-4Glc, and weak binding to a few structures containing histo-Blood Group determinants [264]. In addition, Langerin was shown to bind to β1-3 glucans that, along with mannan binding, likely explains the observation that Langerin serves as the major LC receptor for many fungal strains [265]. The binding of Langerin to high mannose N-glycans, 6'-O-sulfated LacNAc, histo-Blood Group determinants, and β1-3 glucans was confirmed by co-crystallization studies, and the X-ray diffraction studies also revealed similarities and differences in the mechanisms of how these different determinants bind to Langerin, including a unique mechanism for Langerin binding to the 6'-O-sulfated LacNAc [266]. Finally, the fine specificity of Langerin may vary within the human population due to known polymorphisms occurring at key amino acid residues involved in glycan binding. In an interesting example, a mutation destroying a key salt bridge (L313I) necessary for binding to the sulfate group on 6'-O-sulfated LacNAc not only abolishes binding to structures containing terminal 6'-O-sulfated LacNAc but also actually *increases* binding to structures containing terminal β-linked GlcNAc [267]. Therefore, Langerin may bind to multiple glycan determinants, which may be important for LCs and other Langerin-expressing DCs to recognize a wide variety of microbial invaders. Moreover, polymorphisms in Langerin that alter glycan recognition are found in the human population, which may have been evolutionarily selected over time so that the human population exhibits diversity in Langerin-mediated recognition of pathogens.

The macrophage galactose/N-acetylgalactosamine-binding lectin (MGL, also referred to as CLEC10A) was first discovered while screening of an activated monocyte

cDNA library for C-type lectins homologous to the prototype human C-type lectin called the hepatic asialoglycoprotein receptor. This screening was performed because a Gal/GalNAc-specific C-type lectin non-identical to the Gal/GalNAc-specific hepatic asialoglycoprotein receptor was known to be expressed by rat and mouse macrophages [268]. This led to identification of a calcium-dependent C-type lectin that bound to galactose-Sepharose and eluted with 0.1M Gal or GalNAc monosaccharides, and also weakly eluted with 0.1M fucose. Interestingly, the MGL also bound to glycopeptides containing the tumor-associated cancer antigen Tn antigen (GalNAcαSer/Thr) but not to the galactosylated T antigen (Galβ1-3GalNAcαSer/Thr) found on normal tissues [268], which not only suggested a preference for terminal α-linked GalNAc but also that MGL may interact with tumor cells and products. By glycan microarray analysis of monosaccharides, MGL was shown to interact with α-linked GalNAc and, to a lesser extent, β-linked GalNAc, but poorly bound to α- or β-linked Gal. Moreover, MGL specifically bound to glycan structures containing a terminal α- or β-linked GalNAc including LDN (GalNAcβ1-4GlcNAcβ-) and the glycolipids GM2 (GalNAcα1-4(Neu5Acα2-3)Galβ1-4Glc and GD2 (Neu5Acα2-8Neu5Acα2-3(GalNAcα1-4)Galβ1-4Glc), though an exception to this rule was binding to the Core 6 O-glycan structure GlcNAcβ1-6GalNAcα- containing a terminal β1-6-linked GlcNAc rather than GalNAc and an internal GalNAc residue [269]. MGL has also been shown to bind a number of microbes, specifically to glycoconjugates containing terminal GalNAc (reviewed in [270]). Therefore, human MGL's binding motif is simply GalNAcα- or GalNAcβ- and, due to poor binding of terminal Gal determinants, MGL is truly a "GalNAc-binding proteins" and not the original name "Gal/GalNAc-binding protein."

MGL was later shown to be expressed by immature but not mature monocyte-derived DCs and Langerhans cells, which acted as a cell surface endocytic receptor upon binding to antibodies or multivalent GalNAcα-PAA particles [237, 271]. Moreover, MGL partially contributed to the binding and endocytosis of SEA by DCs, along with DC-SIGN and the mannose receptor [248], likely via binding to LDN and the fucosylated analog LDNF (GalNAcβ1-4(Fucα1-3)GlcNAcβ-) [269]. Besides endocytosis, MGL has been shown to modulate TLR signaling and T helper cell differentiation, akin to DC-SIGN. These functions of MGL in DC signaling and T cell differentiation have been excellently reviewed [270], but will be briefly described here. In general, MGL signaling appears to promote a tolerogenic DC phenotype, which causes these MGL-stimulated DCs to dampen effector T cell responses and promote T helper cell differentiation to the tolerogenic and anti-inflammatory Tr1 phenotype. However, more studies on MGL-mediated signaling and regulation of immune responses are needed since only a few studies are available for the role of MGL on human DCs, and some of the evidence is conflicting or controversial, which may depend on the MGL ligand and/or mechanism of DC stimulation or differentiation. It should also be noted that mice contain two homologs of human MGL, MGL1 and MGL2, which have different glycan specificities and functions than human MGL [270]. Therefore, studies performed with mice on MGL function should be interpreted with caution when applied to humans.

Dectin-2 is another CLR expressed by some human DCs. While Dectin-2 was first cloned, characterized, and extensively studied in mice (reviewed in [272-275]), very few studies have focused on human dectin-2. Human dectin-2 was first cloned in 2004 [276], about 4 years after cloning and characterization of the mouse dectin-2 gene [277].

In addition to a few other leukocytes and lung tissue, human dectin-2 mRNA was detectable in both immature and mature monocyte-derived DCs and Langerhans cells [276, 278]. Moreover, higher mRNA levels were detected in pDCs vs. cDCs [278], although protein and/or cell surface expression levels were not compared. These results suggest that dectin-2 is expressed by at least some subsets of DCs.

Dectin-2 serves as a signaling receptor in DCs. Dectin-2 crosslinking by antibodies or binding to specific strains of *Candida albicans* alters cytokine expression by monocyte-derived cDCs, especially by promoting increased expression of the Th17-biasing cytokines IL-1β and IL-23 as well as decreased expression of the Th1-biasing cytokine IL-12 [279]. Importantly, no TLR activation was needed for dectin-2 to cause these changes in cytokine expression [279], which is in contrast to DC-SIGN, Langerin, and MGL. This is likely because dectin-2 and other C-type lectins in the dectin-2 family associate with the Fcγ receptor via cytoplasmic ITAM motifs to transmit signals that directly lead to changes in gene expression (reviewed in [239, 253]), and this Fcγ singling pathway was recently demonstrated for human dectin-2 as well [280]. In pDCs, dectin-2 binding to hyphae of the opportunistic fungal pathogen *Aspergillus fumigatus* caused increased cytokine secretion (of at least TNF-α and IFN-α) as well as increased direct pDC killing of the hyphae. Moreover, the binding of *A. fumigatus* hyphae to dectin-2 on pDCs was inhibited by mannan and weakly inhibited by laminarin (an extract rich in β1-3 glucan) [280], suggesting that human dectin-2 is a mannan receptor and possibly a weak β1-3 glucan receptor. In other cell types, dectin-2 was shown to function as a potential receptor for the Ebola virus glycoprotein to promote viral attachment to and invasion of cells, which suggests that dectin-2 may also serve as a receptor for Ebola virus infection

of DCs [281]. To date though, no direct proof of an endocytic function of human dectin-2 has been shown. These few studies suggest that human dectin-2 is expressed by DCs and functions as a regulator of fungal immune responses, including promoting antifungal immune responses in pDCs and Th17-biased cytokine responses by cDCs, via binding to fungal mannan and possibly β1-3 glucan structures.

iii. Siglecs

Siglecs (sialic acid-binding immunoglobulin-like lectins) are a group of glycan-binding proteins belonging to of the I-type lectin family of glycan-binding proteins. I-type lectins are characterized by the presence of an immunoglobulin (Ig)-like fold in the protein three-dimensional structure and glycan-binding activity via this domain [282]. The term "Siglec" was first proposed in 1998 and used to classify structurally and functionally related glycan-binding protein that met two criteria: binding to sialylated glycans and significant sequence similarity, especially the presence of and similarity within the single variable (V)-Ig-like domain and one or more C2 constant-Ig-like domains [283]. Siglecs are primarily expressed by leukocytes and, not surprisingly, function in numerous immunoregulatory processes (reviewed in [284, 285]). Some of the Siglec and their functions are described in more detail below, with the major focus being on Siglecs expressed by dendritic cells and the DC-specific functions of these Siglecs.

Most Siglecs bind to specific sialylated glycoconjugates and require sialic acid for this binding, although the fine glycan specificity varies between Siglecs. Although earlier studies using a smaller number of glycan structures identified specific sialylated glycan determinants for human Siglecs [286, 287], other and especially more recent studies

using much larger numbers of glycan structures have not confirmed these earlier observations. In particular, Siglec screenings on the glycan microarray available from the Consortium for Functional Glycomics (CFG) have generally shown Siglec binding sialylated glycans, but the actual glycan determinant(s) were indefinable except for the presence of sialic acid (refer to CFG glycan microarray data publically available at http://www.functionalglycomics.org/glycomics/publicdata/home.jsp). The exceptions to this rule are Siglec-2 and Siglec-8. Siglec-2 is known to specifically bind to the Neu5Acα2-6Galβ1-4GlcNAcβ- determinant [286, 288], which was also confirmed by CFG glycan microarray analysis (http://www.functionalglycomics.org/glycomics/publicdata/home.jsp). Siglec-8 specifically binds to a sialyl Lewis x determinant (Neu5AcGalβ1-4(Fucα1-3)GlcNAcβ-) that is sulfated on the 6-OH group of galactose (termed 6'-O-sulfo-sialyl Lewis x) [287, 289]. However, more recent CFG glycan microarray analyses suggest that the fucose moiety is not important in this interaction and thus the true determinant is likely Neu5Acα2-3Gal6Sβ1-4GlcNAcβ- as opposed Neu5Acα2-3Gal6Sβ1-4(Fucα1-3) GlcNAcβ- (http://www.functionalglycomics.org/glycomics/publicdata/home.jsp).

For other Siglecs, the glycan specificity results vary between studies. For example, Siglec-5 binding was specific for Neu5Ac/Neu5Gcα2-3Galβ1-4Glc in one study [286], but Siglec-5 gave much weaker and more broad binding on the CFG glycan microarray (http://www.functionalglycomics.org/glycomics/publicdata/home.jsp) and in other studies on Siglec-5 sialylated glycan specificity [290]. Additionally, Siglec-9 was specific for the 6'-O-sulfated sialyl Lewis x determinant [287]. Despite the strong binding of Siglec-9 to 6'-O-sulfated sialyl Lewis x in the latter study [287], the binding

strength for this determinant was quite low in CFG glycan microarray studies (http://www.functionalglycomics.org/glycomics/publicdata/home.jsp). Therefore, the specificity of many Siglecs for specific sialylated glycan determinants is equivocal at this point in time.

It is possible that sialic acid is truly the major determinant of most Siglecs and that the other components of sialylated glycans make only small contributions to binding and/or presentation of the sialic acid. For example, Siglec-7 on dendritic cells binds to sialylated glycans containing different types of sialic acid linkages, but the different sialic acid linkages result in different downstream effector functions [291]. However, some specificity in sialylated glycan binding is still observable in some studies [286, 287], suggesting that some elements beyond sialic acid are still important for Siglec binding. Moreover, the binding affinity of most Siglecs to known sialylated ligands is quite low (with IC_{50} values of ~0.5mM for most ligands) in solution [286]. The exceptions to this rule of relatively weak binding are again Siglec-2 and Siglec-8; Siglec-2 binds 6'-sialyllactose with a K_d of 32μM by equilibrium dialysis [292], and Siglec-8 binds 6'-O-sulfo-sialyl Lewis x with a K_d of 2-2.5μM by SPR [289]. However, the binding strength of other Siglecs is much stronger (low μM K_d's) using a solid-phase method of measurement (surface plasmon resonance, SPR) with glycans conjugated to polyacrylamide beads [293], although it is unclear if the solid-phase presentation and/or multivalent glycan presentation on polyacrylamide beads increased the strength of these binding interactions. This raises the question of how Siglecs physiologically interact with sialylated glycans. On one hand, avidity and/or the very high density of sialylated glycans at the target cell surface or in serum may still make these low affinity Siglec-glycan

interactions physiologically relevant [294]. Alternatively, the glycan binding specificity and strength by Siglecs may have been severely underestimated by some studies because of the methodology used, as was described above for Siglec-9. Hence, Siglec-glycan interactions may critically dependent on the glycan presentation factors such as solution vs. solid-phase presentation of the glycan, the substratum (glass, plastic, nitrocellulose, microspheres, cells, etc.) in solid-phase binding assays, the specific glycan backbone, and/or the presence or identity of an aglycone component (lipid, peptide, or protein). For all these reasons, more glycan specificity, thermodynamic, and structural studies on Siglec-glycan interactions are needed to identify the mechanisms of binding. Ideally, these studies should eventually focus on the binding of Siglecs to physiologically relevant endogenous and/or exogenous ligands, which can be identified by pull-down studies with Siglec-coated beads and target cell lysates and subsequent characterization of the glycopeptides promoting Siglec binding..

Human dendritic cells express a number of Siglec proteins, including Siglec-1 (sialoadhesin, CD169), Siglec-5, Siglec-7, Siglec-9, and Siglec-10. Some dendritic cells may also express other Siglecs, including Siglec-3 (CD33) [295]. However, the major focus of the following paragraphs will be on Siglec-1, -5, -7, -9, and -10 due to their use in the current study.

Siglec-1 (originally called sheep erythrocyte receptor, or SER) was the first Siglec identified and was discovered by Crocker et al. based on specific expression of a neuraminidase-sensitive hemagglutinin of sheep erythrocytes in mouse bone marrow macrophages but not mouse peritoneal macrophages that was specifically inhibited by addition of sialyllactose (containing either predominantly α2-3- or predominantly α2-6-

linked sialic acid) [296]. Using a monoclonal antibody that was found to block the neuraminidase-sensitive hemagglutination activity, Crocker et al. subsequently purified and characterized SER as a sialic acid-dependent lectin that preferentially bound α2-3-linked sialylated glycoconjugates; the protein was also renamed as sialoadhesin [297]. Siglec-1 expression on dendritic cells was only recently discovered because this Siglec is only expressed under certain conditions and in certain DC subtypes. The first evidence of Siglec-1 expression and function in human DCs was in a 2005 study, where Kirchberger et al. were studying the mechanism of DC anergy that results when DCs contact specific human rhinovirus strains. It was found that PD-L1 and an unknown receptor, later identified as Siglec-1, were responsible for this T cell anergic effect because antibodies against both these receptors almost completely removed the anergic phenotype [298]. Therefore, Siglec-1 may function as an inhibitory molecule towards T cell activation under conditions of expression. However, the mechanism of action and ligand(s) on T cells are currently unknown, although CD43 and PSGL-1 are known T cell ligands for Siglec-1 and hence may represent potential ligands for Siglec-1 on DCs [298, 299]. Siglec-1 is only weakly expressed by immature DCs but the expression dramatically increases in LPS-stimulated mature DCs [300]. In fact, Siglec-1 may serve as the major HIV receptor on mature DCs to promote $CD4^+$ T cell *trans*-infection by DCs in a similar manner as DC-SIGN [300]. Therefore, Siglec-1 is expressed on DCs under specific conditions or subtypes such as mature DCs and rhinovirus-stimulated DCs, where it acts as an inhibitory receptor towards T cell activation and as an HIV receptor. Due to the paucity of studies on Siglec-1 on DCs, future studies should be aimed at further understanding the functional roles of Siglec-1 in dendritic cell biology and immunity as

well as the mechanism of Siglec-1 gene and cell surface expression since this Siglec is not constitutively expressed.

Siglec-5, Siglec-7, Siglec-9, and Siglec-10 were all discovered within a span of three years based on their homology to other known Siglecs and demonstration of their sialic acid-dependent glycan binding activities [290, 301-303]. Siglec-10 was actually initially cloned from a DC cDNA library and subsequently shown to express Siglec-10 at the cell surface [303]. Siglec-5, Siglec-7, and Siglec-9 were shown to be expressed by dendritic cells, specifically immature monocyte-derived dendritic cells, in a 2004 study by Lock et al. [304]. LPS-induced mature DCs also expressed siglec-5 at similar levels as immature DCs, whereas the cell surface expression of Sigle-7 and -9 is significantly reduced upon LPS stimulation of DCs [304]. Interestingly, Siglec-5 is the only Siglec expressed by circulating plasmacytoid dendritic cells (pDCs) under basal conditions [304], although the functional role of Siglec-5 in pDCs is currently unknown. Indeed, very few structural and functional studies of Siglec-5, -7-, and -9 have been performed, let alone with DCs. No studies to date have been performed on Siglec-5 in dendritic cells. However, studies in other cell types have shown that siglec-5 may serve as an endocytic receptor [304], including for microbes carrying sialylated glycoconjugates such as the meningitis-causing pathogen *Neisseria meningitidis* [305]. Additionally, Siglec-5 may serve as a negative regulator of immune cell activation, and this function may be exploited by pathogens to evade immune responses. *Streptococcus agalactiae* produces a protein, β protein, which was shown to bind Siglec-5 in a sialic acid-independent manner and promote Siglec-5 -mediated inhibition of innate immune mechanisms such as ROS and NET production [306]. In contrast to *N. meningitidis*, the β protein of *S. agalactiae*

also inhibits phagocytosis by neutrophils, but likely via negative regulation of immune signaling pathways instead of directly acting as an endocytic receptor [306]. Therefore, Siglec-5 may serve as an endocytic receptor and negative regulator of immune functions in leukocytes, which may include DCs, although studies with DCs are needed to confirm this.

A few studies have been performed on the role of Siglec-7, Siglec-9, and Siglec-10 in human DCs. Siglec-7 was shown to bind multiple lipooligosaccharide (LOS) structural mimics of gangliosides from the gastrointestinal pathogen *Campylobacter jejuni* [291], but different structures containing different sialic acid linkages resulted in different functions. Specifically, LOS containing the α2-3-sialylated GD1a determinant (Neu5Acα2-3Galβ1-3GalNAcβ1-4(Neu5Acα2-3)Galβ-) caused DCs to secrete Th2 cytokines and promote Th2 differentiation of $CD4^+$ T helper cells while LOS containing the terminal α2-8-sialylated GD1c determinant (Neu5Acα2-8Neu5Acα2-3Galβ1-3GalNAcβ1-4Galβ-) promoted Th1 cytokine secretion by DCs and a Th1 phenotype towards T cells [291]. Therefore, Siglec-7 may recognize multiple sialic acid linkages in glycans, although the direct contribution of Siglec-7 to these effects remains associative and not causative since no antibody inhibition or Siglec-7 knockout studies were performed to confirm that the effects were due to Siglec-7. Siglec-7 may also promote endocytosis of multivalent ligands, such as bacterial cells or liposomes, and stimulate antigen presentation for activation of T cells [307]. For this reason, targeting Siglec-7 on DCs has been proposed as a mechanism for generating more specific, efficacious vaccines against pathogens such as *Mycobacterium tuberculosis* [307]. Siglec-9 on immature DCs may bind to MUC2 mucin from human colon carcinoma cells, which may

result in reduced IL-12 production by the DCs [308]. However, the effects of this interaction on production of other cytokines or immune responses was not tested, and it is unclear whether or not MUC2 from healthy human cells or samples has a similar effect and if this function has biological relevance. Nonetheless, the results of the latter study suggest that Siglec-9 may function as an inhibitor of DC Th1 immune responses and possibly other immune responses. In DCs, Siglec-10 (or the mouse homolog Siglec-G) binding to sialylated glycans on CD24 inhibited NF-κB signaling and inflammatory responses to the danger-associated molecular pattern (DAMP) protein high-mobility group box 1 (HMGB1) in a liver injury model [309]. Siglec-10 on DCs in the lamina propria may also serve as a receptor for the major flagellin protein (FlaA) of the gastrointestinal pathogen *Campylobacter jejuni*, which results in increased IL-10 production by the DCs via p38 activation [310]. Interestingly, Siglec-10 binding to the FlaA may be mediated by the presence of pseudaminic acid and derivatives [310]. Pseudaminic acid is nine-carbon monosaccharide (nonaose) similarly to sialic acids but is a different isomer with different functional groups at some of the carbons (5,7-diamino-3,5,7,9-tetradeoxy-L-*glycero*-L- *manno*-nonulosonic acid for pseudaminic acid vs. the 5-amino-3,5-dideoxy-D-*glycero*-D-*galacto*-nonulosonic acid core structure of neuraminic acids like Neu5Ac in humans) [311]. Therefore, Siglec-10 and possibly other Siglecs may even serve as a receptor for exogenous α-nonulosonic acid monosaccharides like pseudaminic acid and/or derivatives. Future studies should be aimed at understanding the mechanism of Siglec binding, if relevant, to exogenous sialic acids and other ulosonic acids such as 2-keto-3-deoxy-D-*glycero*-D-*galacto*-nonulosonic acid (KDN), 2-keto-3-

deoxy-D-*manno*-octulosonic acid (KDO), 5,7-diamino-3,5,7,9-tetradeoxy-D-*glycero*-D-*galacto*-nonulosonic acid (legionaminic acid), and derivatives of these structures [311].

In all, these studies suggest that Siglecs on DCs participate in a variety of immune functions in immune cells from direct endocytosis and subsequent antigen presentation to fine-tuning immune responses, especially influencing specific T helper cell responses such as Th1 or Th2 responses, by regulating DC signaling pathways. Future studies should be aimed at understanding more of the functional roles of Siglec-1, -5, -7, -9, -10 in DC biology and immune responses, especially the endogenous and exogenous ligands of these dendritic cell Siglecs.

IV. HMG Interactions with Human Glycan-Binding Proteins

Humans contain a number of different glycan-binding proteins (GBPs), including members of the C-type lectin, Siglec, and galectin families as well as other families. Interestingly, members of the former three GBP families have been shown to interact with specific HMG structures. It should be noted though that these studies primarily focused on understanding the glycan specificity of these GBPs and not the binding of these GBPs to HMGs *per se*.

One study showed that total acidic HMGs and 3-sialyl-3-fucosyllactose (Neu5AcGalβ1-4(Fucα1-3)Glc, a sialyl Lewis x tetrasaccharide analog with Glc instead of GlcNAc at the reducing end), could slightly reduce leukocyte rolling on and attachment to inflamed endothelial cells [87]. While these studies suggested that acidic HMGs or 3'-sialyl-3-fucosyllactose interacted with human selectins, a C-type lectin that mediates leukocyte rolling on endothelial cells for subsequent leukocyte adhesion and extravasation [312], this conclusion is questionable because no direct examination of selectin binding to these HMGs was performed, the inhibitory effect on leukocyte rolling (the actual function mediated by selectins) was only minimal (~25% decrease relative to untreated controls), very little if any change in leukocyte rolling with increased glycan concentrations was seen, and no correction for multiple comparisons was performed in the statistical analysis [87]. Another study also suggested sialylated HMGs inhibit platelet-neutrophil complex formation. While the authors suggested that the interaction was via P-selectin on the platelets [88], this study suffered from the same issues as the previously noted study (namely, little change in magnitude and lack of direct evidence of P-selectin binding). Future studies should be aimed at directly assessing selectin-HMG

interactions by glycan microarrays or functional assays with selectin gene knockouts and/or neutralizing anti-selectin antibody treatments.

Another study showed that the milk glycan LNFPIII (Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4Glc) co-crystallized with the human C-type lectin DC-SIGN [252]. However, that study used a single LNFPIII concentration of 10mM for co-crystallization [252], so the actual affinity of DC-SIGN for LNFPIII was unknown and still has not been directly measured. In addition, Hong et al. showed that human milk glycans could weakly inhibit HIV-1 gp120 binding to DC-SIGN, and this inhibition was proposed as a mechanism by which HMGs may protect infants from HIV-1 infection during breast-feeding [313]. The authors of that study suggested this inhibition was due to the presence of HMGs carrying terminal Lewis antigens based on the known glycan determinants of HMGs (refer to the *Structural Features of HMGs* section) and the glycan specificity of DC-SIGN [246]. However, the study by Hong et al. used *K. fragilis* β-galactosidase treatment of the milk to remove lactose [313], yet the substrate specificity of this enzyme has not been fully explored and thus may have led to structural alterations of the HMG population. Moreover, the weak inhibition of DC-SIGN binding suggests that HIV-1 gp120 may still bind DC-SIGN efficiently enough to allow DC-SIGN-mediated *trans*-infection of $CD4^+$ T cells. Nonetheless, the studies suggest that specific HMGs such as LNFPIII may bind to DC-SIGN, although the physiological relevance of these interactions is still uncertain.

Siglecs are another class of lectins found in humans, which bind sialylated glycans and are most well known for their role in modulating immune responses (reviewed in [284]). While not directly explored for the ability to bind HMGs, studies

have reported the interaction of Siglecs with simple sialylated HMGs. Human recombinant Siglec-2 (CD22) was shown to bind 6'-SL in solution with a K_d of 32μM by equilibrium dialysis [292]. Siglec-2 was also shown to bind multivalent 6'-SL in an ELISA assay [314] and to free 6'-SL in a competition assay [315]. Furthermore, human recombinant Siglec 1, 3, 5, 7, and 9 were shown to directly bind multivalent 3'-SL and 6'-SL conjugates by ELISA [314] and/or surface plasmon resonance [293]. Human Siglec-1 has also been co-crystallized with 3'-SL [316], although very high concentrations (25mM) of 3'-SL were used during the co-crystallization. Therefore, sialylated HMGs may interact with Siglecs.

Galectins have also been shown to bind HMG structures during studies on galectin specificity. LNnT, LNT, LNFPI, and 2'-FL have been shown to be ligands for most galectins [317-319], although the binding affinities for these HMGs are relatively low compared to the major binding determinants of these galectins. Some galectins can also bind LNFPII, LNFPIII, and lacto-N-difucohexaose I (LNDFH I; Fucα1-2Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4Glc), although these interactions are typically weaker strength than to LNnT, LNT, LNFPI, and 2'-FL due to the relatively unfavorable binding of galectins to Lewis glycan determinants [317, 319]. The sialylated HMG 3'-sialyllactose (3'-SL) has also been shown to be s strong ligand for galectin-8, specifically the N-terminal CRD, with a K_d of about 1μM for the free 3'-SL structure [174].

3'-SL may also interact with TLR4 [72], although this interaction has not yet been directly tested by biochemical and cell biology methods. Nevertheless, this finding at least suggests that TLRs may be another class of direct HMG receptors.

The results of these studies all point to the ability of some GBPs to interact with specific HMG structures. Specifically, these studies suggest that HMGs may bind to human GBPs, including C-type lectins, Siglecs, galectins, and possibly even TLRs. Since these GBPs are known to play important roles in physiology, immunity, and host-microbe interactions, HMG-GBP interactions may thus modulate infant physiology and immunity. Although others have hypothesized a biological role of HMG-GBP interactions [320], no studies to date have experimentally tested the functional effects of these interactions, such as regulation of gene expression and immune responses. Thus, these facts raise questions though about the significance of these interactions from both a biochemical and physiological standpoint.

V. Goals of Study

In contrast to the known mechanisms underlying the prebiotic and antimicrobial effects of HMGs, the mechanisms by which HMGs regulate immune responses, epithelial cell gene expression, and neurological effects are currently unclear. One interesting feature of HMGs and GBPs is their overlap in certain functions, especially that both are known to regulate gene expression and immune responses. GBPs expressed by epithelial cells and dendritic cells in the GI tract are also in a prime position for exposure to HMGs during breast-feeding, and the HMGs are not significantly digested in the GI tract until the large intestine is reached [33, 34]. Additionally, the fact that some GBPs can bind HMG structures raises questions about the biological significance of these interactions, including whether or not such interactions occur in breast-fed infants and lead to downstream effects. For these reasons, we hypothesize that some of these GBPs expressed in the GI tract bind HMGs. Although not a goal of this study, we further hypothesize that these HMG-GBP interactions in the GI tract are the mechanism underlying regulation of gene expression and immune responses by HMGs.

Although previous studies have shown that some GBPs bind a few simple HMG structures (refer to the **HMG Interactions with Human Glycan-Binding Proteins** section above), no comprehensive study on GBP-HMG interactions has been carried out to date. The physiological relevance of these interactions is also unclear. The major goal of this study was thus to test the first hypothesis that GBPs interact with HMGs, particularly by performing a comprehensive study utilizing screening platforms that encompass most of the HMG metaglycome such as shotgun HMG microarrays [321]. Additionally, experiments were also conducted to measure or estimate the K_d of specific

GBP-HMG interactions to determine the potential physiological relevance of these interactions. The importance of this study is that it will determine if GBP-HMG binding interactions can occur and if the binding occurs at physiologically relevant concentrations, especially with regard to the concentration of HMGs in human milk. In other words, do the *in vitro* interactions of GBPs with HMGs suggest that such interactions could theoretically occur *in vivo*? These finding will thus be a critical first step in determining if GBP-HMG interactions underlie some of the biological effects of HMGs, especially regulation of immune responses and gene expression. Moreover, the findings may also assist in understanding the glycan specificity of or even unravel novel glycan determinants of GBPs, especially GBPs that have not yet been screened on glycan microarrays such as human dectin-2. This knowledge may be highly useful for developing specific GBP agonists or antagonists as therapeutics for specific diseases, such as Siglec-5-specific antagonists to help treat *Streptococcus agalactiae* carriage or infection [306].

The GBPs tested in this study were those expressed by GI tract epithelial cells and dendritic cells since these two cell types express a number of GBPs that may come in direct contact with HMGs passing through the GI tract in breast-fed infants. These GBPs include members of the galectin, C-type lectin, and I-type lectin (Siglec) families of GBPs. This study also took advantage of a high-throughput technology, glycan microarrays, which are glass slides containing tens to hundreds of different glycan or glycoconjugate structures that are screened with GBPs in order to elucidate the glycan structure(s) bound (reviewed in [321-323]). The major goal of glycan microarray screenings is to define the glycan specificity of a GBP, including any common feature(s)

of the glycan structures bound. This common feature(s) may represent the actual portion of the glycan structure(s) responsible for GBP binding, which is termed the GBP's *motif* [324]. Hence, the goal of the glycan microarray screenings in this study were to not only identify if and which HMG structures bind GBPs but to also utilize these screenings to define the glycan specificity and motifs of these GBPs. To this end, the GBPs were screened on multiple glycan microarray platforms, including a shotgun HMG microarray consisting of over 200 natural, purified glycans encompassing the HMG metaglycome. A glycan microarray consisting of a small panel of chemically defined HMGs, including infant formula additives such as galactooligosaccharides (GOS), was also utilized to confirm and extend the results of the HMG shotgun glycan microarray analysis. Finally, the GBPs were screened on the CFG glycan microarray since this microarray consists of a very large number of glycan structures (610 total structures) encompassing many different glycan determinants, including many not present on HMGs. Besides elucidating if and which HMGs are bound by GBPs, the glycan microarray screening results were useful for understanding the HMG determinants bound by these GBPs as well as other potential physiologically relevant endogenous and exogenous glycan structures of these GBPs.

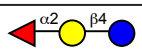
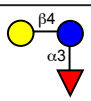
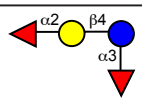
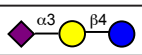
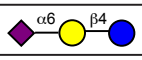
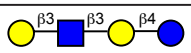
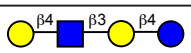
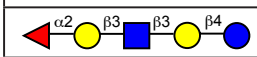
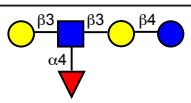
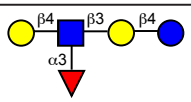
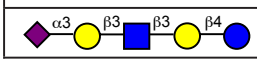
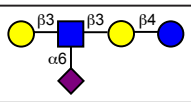
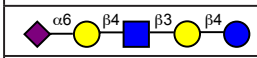
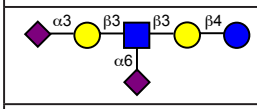
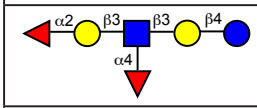
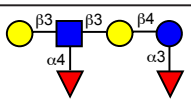
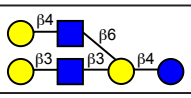
In addition to glycan microarray experiments, isothermal titration microcalorimetry (ITC) and/or hapten inhibition experiments were performed in order to measure or estimate the affinity of specific GBP-HMG interactions. These ITC and hapten inhibition experiments were also used to determine if binding occurred in solution (the natural mode of HMG presentation, as opposed to the solid-phase presentation of glycans on microarrays) and in a glycan- and CRD-specific manner. The results of these

experiments helped elucidate whether or not specific GBP-HMG interactions were strong enough to potentially occur *in vivo*.

The results of this study suggest that some GBPs expressed in the GI tract, particularly many galectins and the C-type lectin DC-SIGN, specifically bind HMGs. Additionally, the study actually helped further refine the glycan-binding specificity of these GBPs, including the identification of novel binding motifs for some GBPs, especially galectin-7. The affinity of some of the GBP-HMG interactions were also measured and determined to occur within physiologically relevant HMG concentrations. These results suggest that DC-SIGN-HMG and galectin-HMG interactions may occur breast-fed infants' GI tract, which warrants future studies on the biological effects of such interactions in proper infant development and immunity.

VI. Tables and Figures

Table 1. Structures, Names, and Concentrations of Common Human Milk Glycans

Structure[a]	Name	Concentration Range (g/L)[b]	Concentration Range (mM)[b]
	2'-fucosyllactose (2'-FL)	0.22 - 4.57[c]	0.45 - 9.36[c]
	3-fucosyllactose (3-FL)	0.07 - 2.35	0.14 - 4.81
	Lactodifucotetraose (LDFT)	0.07 - 0.43[c]	0.11 - 0.68[c]
	3'-sialyllactose (3'-SL)	0.076 - 0.35	0.12 - 0.55
	6'-sialyllactose (6'-SL)	0.19 - 1.77	0.30 - 2.80
	Lacto-N-tetraose (LNT)	0.55 - 3.90	0.77 - 5.51
	Lacto-N-neotetraose (LNnT)	0.08 - 2.04	0.11 - 2.88
	Lacto-N-fucopentaose I (LNFPI)	0.28 - 3.18[c]	0.33 - 3.72[c]
	Lacto-N-fucopentaose II (LNFPII)	0.14 - 1.25[d]	0.16 - 1.46[d]
	Lacto-N-fucopentaose III (LNFPIII)	0.31 - 0.44	0.36 - 0.52
	Lactosialyl-N-tetraose a (LSTa)	0.01 - 0.141	0.01 - 0.141
	Lactosialyl-N-tetraose b (LSTb)	0.05 - 0.131	0.05 - 0.131
	Lactosialyl-N-tetraose c (LSTc)	0.09 - 0.686	0.09 - 0.687
	Disialyllacto-N-tetraose (DSLNT)	0.29 - 1.274	0.22 - 0.9874
	Lacto-N-difucohexaose I (LNDFH I)	0.50 - 1.87[c, d]	0.50 - 1.87[c, d]
	Lacto-N-difucohexaose II (LNDFH II)	0.02 - 0.86[d]	0.02 - 0.86[d]
	Lacto-N-hexaose (LNH)	0.07 - 0.13	0.065 - 0.12

[a]Monosaccharide key: ● = Glc, ○ = Gal, ■ = GlcNAc, ▲ = Fuc, ◆ = Neu5Ac

[b]The concentration range includes the measured HMG concentrations in individual milk samples and the average HMG concentrations of all milk samples in a given study. The measurements were pooled from Urashima et al. [12]

[c]Absent in Secretor-negative milk

[d]Absent in Lewis-negative milk

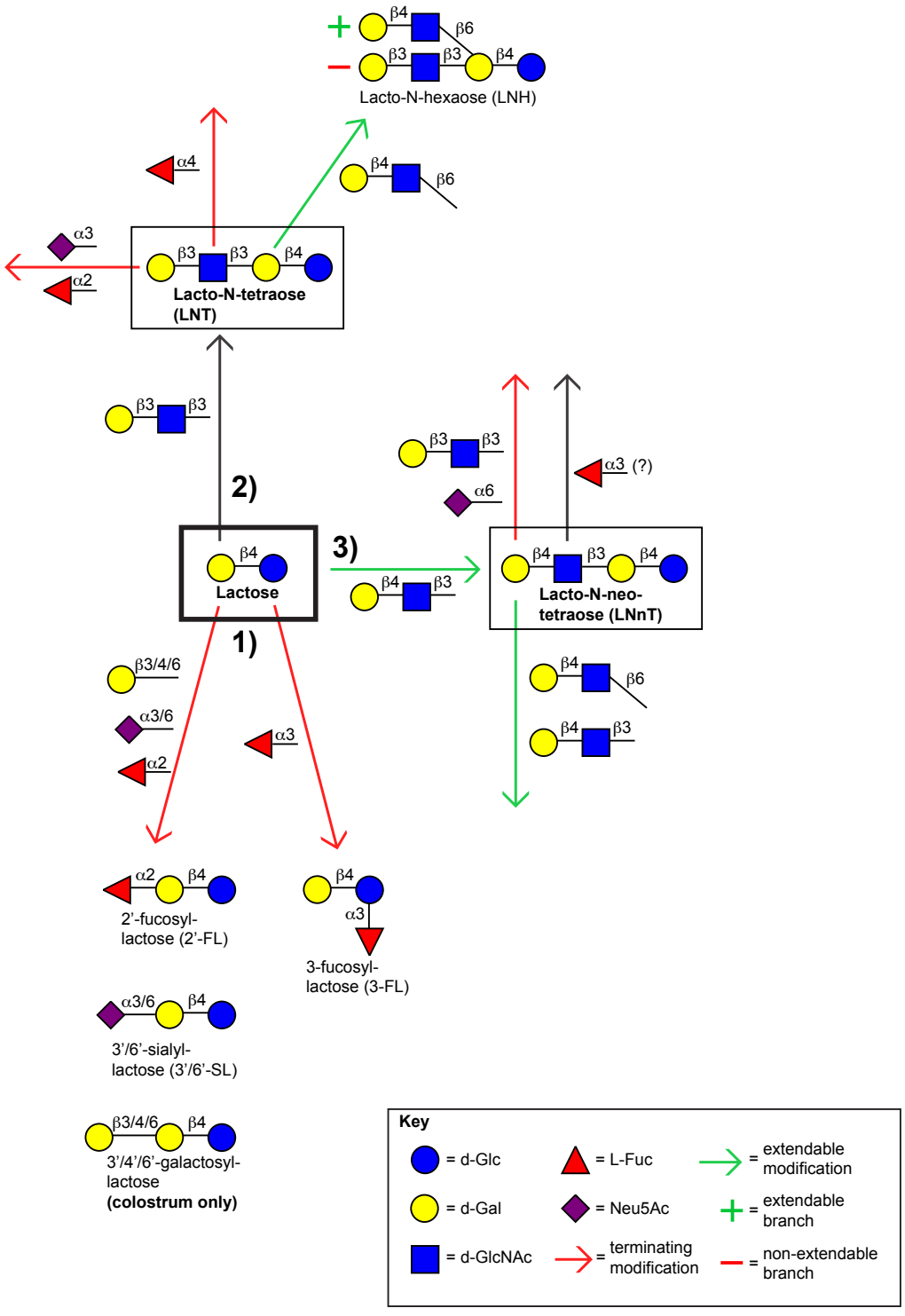

Lacto-N-hexaose (LNH)
Lacto-N-tetraose (LNT)
Lactose
Lacto-N-neo-tetraose (LNnT)
1)
2)
3)
2'-fucosyl-lactose (2'-FL)
3-fucosyl-lactose (3-FL)
3'/6'-sialyl-lactose (3'/6'-SL)
3'/4'/6'-galactosyl-lactose (colostrum only)
Key
= d-Glc
= d-Gal
= d-GlcNAc
= L-Fuc
= Neu5Ac
= terminating modification
= extendable modification
= extendable branch
= non-extendable branch

Figure 1. General Features of Human Milk Glycans (HMGs) and HMG Biosynthesis.

All HMGs contain a reducing end lactose (Galβ1-4Glc) unit that is further extended and modified with Gal, GlcNAc, Fuc, and/or Neu5Ac. Three pathways from lactose exist. The first pathway involves addition of α-linked Fuc, α-linked Neu5Ac, or (in colostrum only) β-linked Gal directly to the lactose unit, which prevents further extension with N-acetyllactosamine (LacNAc). Another HMG, lacto-difucotetraose (LDFT, Fucα1-2Galβ1-4(Fucα1-3)Glc, not depicted) can also be formed by this pathway. The second pathway involves the addition of a Type 1 N-acetyllactosamine unit (LacNAc I; Galβ1-3GlcNAcβ1-3-) to the reducing end lactose, forming lacto-N-tetraose (LNT). This LacNAc I unit itself cannot be further extended with LacNAc but can be modified with α-linked Fuc or α-linked Neu5Ac. Additionally, a branching LacNAc unit, Galβ1-4GlcNAcβ1-6-, can also be added to the Gal residue nearest to the reducing end forming lacto-N-hexaose (LNH). The LNH can be further extended with LacNAc, but only on the branch containing the β1-4-linked Gal since β1-3-linked Gal (LacNAc I) units cannot be further extended. The third pathway involves the addition of Type 2 N-acetyllactosamine (LacNAc II) to lactose, forming lacto-N-neotetraose (LNnT). Unlike LNT, LNnT can be further extended with LacNAc and further modified from there. Additionally, LNnT can be modified with α1-3-linked Fuc, α2-6-linked Neu5Ac, or branched by the addition of Galβ1-4GlcNAcβ1-6-.

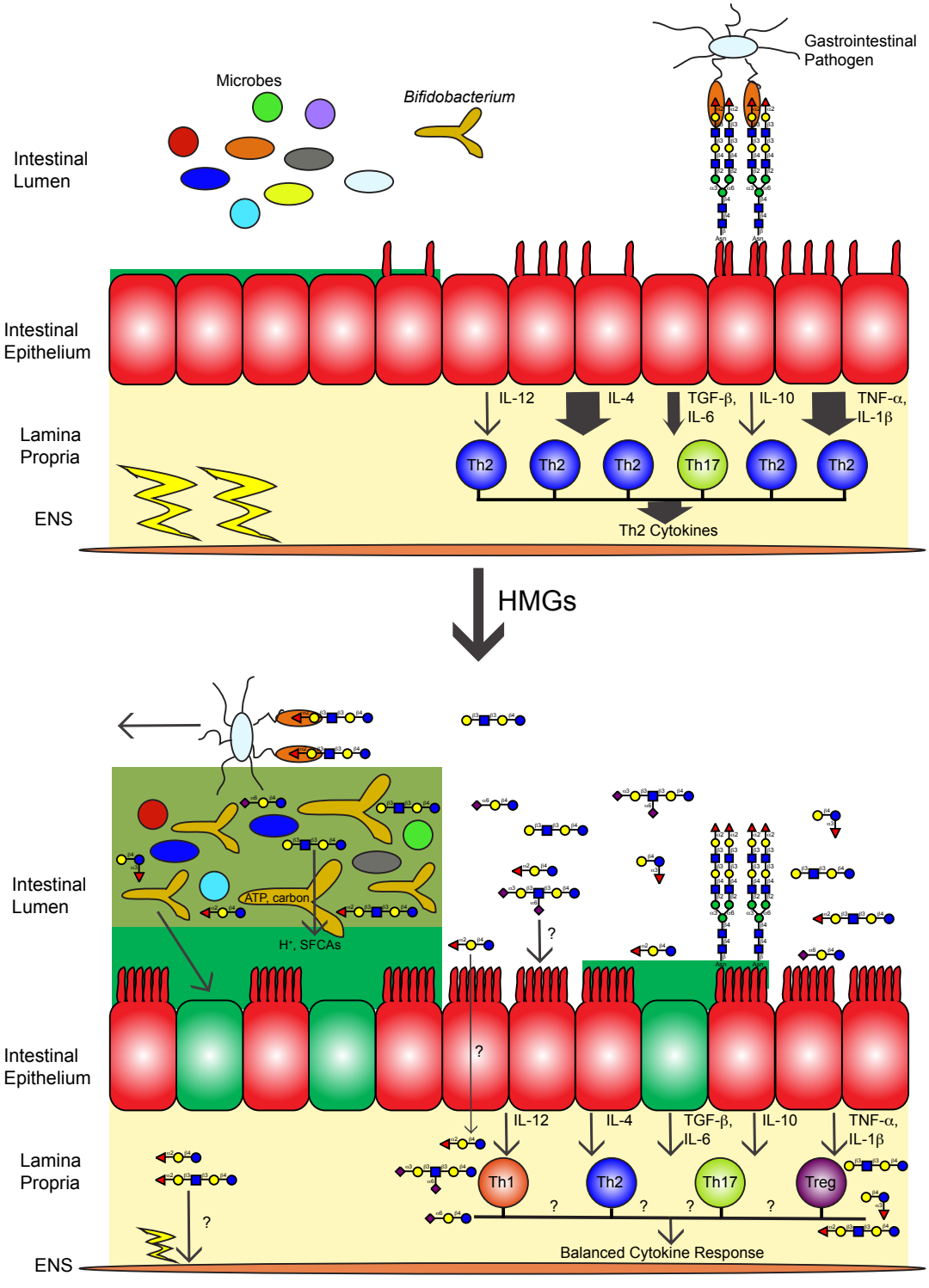

Gastrointestinal Pathogen
Microbes
Bifidobacterium
Intestinal Lumen
Intestinal Epithelium
Lamina Propria
ENS
IL-12
IL-4
TGF-β, IL-6
IL-10
TNF-α, IL-1β
Th2
Th2
Th2
Th17
Th2
Th2
Th2 Cytokines
HMGs
ATP, carbon
H+, SFCAs
?
IL-12
IL-4
TGF-β, IL-6
IL-10
TNF-α, IL-1β
Th1
Th2
Th17
Treg
Balanced Cytokine Response

Figure 2. Functions of HMGs in the Gastrointestinal Tract.

Prior to HMG exposure (top of image), the newborn intestine is still developing and beginning to be exposed to exogenous microbes that may eventually colonize the intestine, including pathogens. In the intestinal epithelium, epithelial cells (red) have not yet fully differentiated into highly absorptive, ciliated cells, and goblet cells (green) are scarce. The epithelium also secretes cytokines into the lamina propria that promote a pro-inflammatory environment due to high TNF-α, IL-1β, and IL-6 levels but low IL-10 levels. Additionally, helper T cells the newborn intestine exhibit a highly Th2-biased phenotype, which is highly pro-allergenic and may result from much IL-4 but low IL-12 secretion by the intestinal epithelium. This Th2-biased environment and pro-inflammatory cytokine profile of the intestinal epithelium makes the newborn intestine highly sensitive to insults, resulting in more powerful immune responses that may ultimately result in severe inflammatory or allergenic diseases such as necrotizing enterocolitis and food allergies. The enteric nervous system (ENS) is also highly active, with more frequent and larger-intensity motor contractions. During breast-feeding (bottom of image), high concentrations of HMGs "bathe" the intestinal epithelium, and a small percentage also crosses the epithelium to enter the lamina propria and circulation. The effects of HMG exposure include selection of a specific gut microbiota, especially specific strains of *Bifidobacterium* that metabolize HMGs and other specific microbial strains such as *Bacteriodes thetaiotamicron*. The symbiotic bacteria produce beneficial short-chain fatty acids (SFCAs) as well as fermentation end products that acidify the intestinal lumen. Some strains also induce goblet cell differentiation and mucus production (green layers above cells) by these goblet cells as well as promote more

balanced immune responses. The HMGs also serve as “receptor decoys” by binding to pathogen glycan-binding adhesins, thereby preventing pathogen adhesion and subsequent infection or invasion of the intestinal epithelium. By a currently unknown mechanism, HMGs also stimulate differentiation of the developing epithelial cells into highly absorptive, ciliated epithelial cells expressing higher levels of digestive enzymes, allowing more efficient nutrient uptake by the newborn. This differentiation is likely mediated by the changes in epithelial cell gene expression induced by HMGs. The HMGs also promote a more balanced, less inflammatory T cell response by HMG stimulation of both T cells and the intestinal epithelial cells by an unknown mechanism. This stimulation, possibly in combination with gut microbiota signals, leads to a balanced pro-inflammatory:anti-inflammatory environment and Th1:Th2:Th17 cytokine ratios. ENS motor contractions are also less frequent and of lower magnitude, which may help slow down food passage through the intestine and thus promote more efficient nutrient absorption. Again though, the mechanism underlying this decrease in ENS motor contractions is unclear. Together, these effects may help promote intestinal and development, increase nutrient absorption and growth, reduce the incidence of gastrointestinal infections and diseases, and reduce the risk of severe inflammatory and allergenic diseases.

VII. References

1. World Health Organization, *Infant and young child feeding: Model Chapter*, 2009, WHO Press: Geneva, Switzerland.
2. Horta, B.L., C.G. Victora, and World Health Organization, *Short-term effects of breastfeeding: a systematic review on the benefits of breastfeeding on diarrhoea and pneumonia mortality*, 2013, WHO Press: Geneva, Switzerland.
3. Horta, B.L., et al., *Evidence on the long-term effects of breastfeeding*, 2007, WHO Press: Geneva, Switzerland.
4. Horta, B.L., C.G. Victora, and World Health Organization, *Long-term effects of breastfeeding: a systematic review*, 2013, WHO Press: Geneva, Switzerland.
5. Walker, A., *Breast milk as the gold standard for protective nutrients.* J Pediatr, 2010. **156**(2 Suppl): p. S3-7.
6. Peterson, R., et al., *Glycoconjugates in human milk: protecting infants from disease.* Glycobiology, 2013. **23**(12): p. 1425-38.
7. Kunz, C., *Historical aspects of human milk oligosaccharides.* Adv Nutr, 2012. **3**(3): p. 430S-9S.
8. Yu, Y., et al., *Human milk contains novel glycans that are potential decoy receptors for neonatal rotaviruses.* Mol Cell Proteomics, 2014. **13**(11): p. 2944-2960.
9. Kobata, A., *Structures and application of oligosaccharides in human milk.* Proc Jpn Acad Ser B Phys Biol Sci, 2010. **86**(7): p. 731-47.
10. Bode, L., *Human milk oligosaccharides: every baby needs a sugar mama.* Glycobiology, 2012. **22**(9): p. 1147-62.

11. Cummings, R.D., *The repertoire of glycan determinants in the human glycome.* Mol Biosyst, 2009. **5**(10): p. 1087-104.

12. Urashima, T., et al., *The predominance of type I oligosaccharides is a feature specific to human breast milk.* Adv Nutr, 2012. **3**(3): p. 473S-82S.

13. Kunz, C., et al., *Oligosaccharides in human milk: structural, functional, and metabolic aspects.* Annu Rev Nutr, 2000. **20**: p. 699-722.

14. Jiang, X., et al., *Human milk contains elements that block binding of noroviruses to human histo-blood group antigens in saliva.* J Infect Dis, 2004. **190**(10): p. 1850-9.

15. Kaji, H., et al., *Large-scale identification of N-glycosylated proteins of mouse tissues and construction of a glycoprotein database, GlycoProtDB.* J Proteome Res, 2012. **11**(9): p. 4553-66.

16. Urashima, T., et al., *Recent advances in studies on milk oligosaccharides of cows and other domestic farm animals.* Biosci Biotechnol Biochem, 2013. **77**(3): p. 455-66.

17. Tao, N., et al., *Bovine milk glycome.* J Dairy Sci, 2008. **91**(10): p. 3768-78.

18. Lane, J.A., et al., *Transcriptional response of HT-29 intestinal epithelial cells to human and bovine milk oligosaccharides.* Br J Nutr, 2013. **110**(12): p. 2127-37.

19. Kumazaki, T. and A. Yoshida, *Biochemical evidence that secretor gene, Se, is a structural gene encoding a specific fucosyltransferase.* Proc Natl Acad Sci U S A, 1984. **81**(13): p. 4193-7.

20. Thurl, S., et al., *Detection of four human milk groups with respect to Lewis blood group dependent oligosaccharides.* Glycoconj J, 1997. **14**(7): p. 795-9.

21. Newburg, D.S., et al., *Innate protection conferred by fucosylated oligosaccharides of human milk against diarrhea in breastfed infants.* Glycobiology, 2004. **14**(3): p. 253-63.

22. Castanys-Munoz, E., M.J. Martin, and P.A. Prieto, *2'-fucosyllactose: an abundant, genetically determined soluble glycan present in human milk.* Nutr Rev, 2013. **71**(12): p. 773-89.

23. Johnson, P.H. and W.M. Watkins, *Purification of the Lewis blood-group gene associated alpha-3/4-fucosyltransferase from human milk: an enzyme transferring fucose primarily to type 1 and lactose-based oligosaccharide chains.* Glycoconj J, 1992. **9**(5): p. 241-9.

24. Prieels, J.P., et al., *Co-purification of the Lewis blood group N-acetylglucosaminide alpha 1 goes to 4 fucosyltransferase and an N-acetylglucosaminide alpha 1 goes to 3 fucosyltransferase from human milk.* J Biol Chem, 1981. **256**(20): p. 10456-63.

25. Weston, B.W., et al., *Molecular cloning of a fourth member of a human alpha (1,3)fucosyltransferase gene family. Multiple homologous sequences that determine expression of the Lewis x, sialyl Lewis x, and difucosyl sialyl Lewis x epitopes.* J Biol Chem, 1992. **267**(34): p. 24575-84.

26. Coppa, G.V., et al., *Changes in carbohydrate composition in human milk over 4 months of lactation.* Pediatrics, 1993. **91**(3): p. 637-41.

27. Kunz, C., et al., *Lactose-derived oligosaccharides in the milk of elephants: comparison with human milk.* Br J Nutr, 1999. **82**(5): p. 391-9.

28. Thurl, S., et al., *Variation of human milk oligosaccharides in relation to milk groups and lactational periods.* Br J Nutr, 2010. **104**(9): p. 1261-71.

29. Coppa, G.V., et al., *Oligosaccharides in human milk during different phases of lactation.* Acta Paediatr Suppl, 1999. **88**(430): p. 89-94.

30. Chaturvedi, P., et al., *Fucosylated human milk oligosaccharides vary between individuals and over the course of lactation.* Glycobiology, 2001. **11**(5): p. 365-72.

31. Thurl, S., B. Muller-Werner, and G. Sawatzki, *Quantification of individual oligosaccharide compounds from human milk using high-pH anion-exchange chromatography.* Anal Biochem, 1996. **235**(2): p. 202-6.

32. Chaturvedi, P., et al., *Milk oligosaccharide profiles by reversed-phase HPLC of their perbenzoylated derivatives.* Anal Biochem, 1997. **251**(1): p. 89-97.

33. Engfer, M.B., et al., *Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract.* Am J Clin Nutr, 2000. **71**(6): p. 1589-96.

34. Gnoth, M.J., et al., *Human milk oligosaccharides are minimally digested in vitro.* J Nutr, 2000. **130**(12): p. 3014-20.

35. Rudloff, S., et al., *Urinary excretion of lactose and oligosaccharides in preterm infants fed human milk or infant formula.* Acta Paediatr, 1996. **85**(5): p. 598-603.

36. Rudloff, S., et al., *Urinary excretion of in vivo (1)(3)C-labelled milk oligosaccharides in breastfed infants.* Br J Nutr, 2012. **107**(7): p. 957-63.

37. Goehring, K.C., et al., *Direct evidence for the presence of human milk oligosaccharides in the circulation of breastfed infants.* PLoS One, 2014. **9**(7): p. e101692.

38. Gnoth, M.J., et al., *Investigations of the in vitro transport of human milk oligosaccharides by a Caco-2 monolayer using a novel high performance liquid chromatography-mass spectrometry technique.* J Biol Chem, 2001. **276**(37): p. 34363-70.

39. Gyorgy, P., *A hitherto unrecognized biochemical difference between human milk and cow's milk.* Pediatrics, 1953. **11**(2): p. 98-108.

40. Coppa, G.V., et al., *The first prebiotics in humans: human milk oligosaccharides.* J Clin Gastroenterol, 2004. **38**(6 Suppl): p. S80-3.

41. Ward, R.E., et al., *In vitro fermentation of breast milk oligosaccharides by Bifidobacterium infantis and Lactobacillus gasseri.* Appl Environ Microbiol, 2006. **72**(6): p. 4497-9.

42. Ruiz-Moyano, S., et al., *Variation in consumption of human milk oligosaccharides by infant gut-associated strains of Bifidobacterium breve.* Appl Environ Microbiol, 2013. **79**(19): p. 6040-9.

43. Ward, R.E., et al., *In vitro fermentability of human milk oligosaccharides by several strains of bifidobacteria.* Mol Nutr Food Res, 2007. **51**(11): p. 1398-405.

44. Asakuma, S., et al., *Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria.* J Biol Chem, 2011. **286**(40): p. 34583-92.

45. Yu, Z.T., C. Chen, and D.S. Newburg, *Utilization of major fucosylated and sialylated human milk oligosaccharides by isolated human gut microbes.* Glycobiology, 2013. **23**(11): p. 1281-92.

46. Marcobal, A., et al., *Consumption of human milk oligosaccharides by gut-related microbes.* J Agric Food Chem, 2010. **58**(9): p. 5334-40.

47. Marcobal, A., et al., *Bacteroides in the infant gut consume milk oligosaccharides via mucus-utilization pathways.* Cell Host Microbe, 2011. **10**(5): p. 507-14.

48. Kitaoka, M., J. Tian, and M. Nishimoto, *Novel putative galactose operon involving lacto-N-biose phosphorylase in Bifidobacterium longum.* Appl Environ Microbiol, 2005. **71**(6): p. 3158-62.

49. Xiao, J.Z., et al., *Distribution of in vitro fermentation ability of lacto-N-biose I, a major building block of human milk oligosaccharides, in bifidobacterial strains.* Appl Environ Microbiol, 2010. **76**(1): p. 54-9.

50. Sela, D.A., et al., *The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome.* Proc Natl Acad Sci U S A, 2008. **105**(48): p. 18964-9.

51. Lewis, Z.T., et al., *Maternal fucosyltransferase 2 status affects the gut bifidobacterial communities of breastfed infants.* Microbiome, 2015. **3**: p. 13.

52. Kitaoka, M., *Bifidobacterial enzymes involved in the metabolism of human milk oligosaccharides.* Adv Nutr, 2012. **3**(3): p. 422S-9S.

53. Zivkovic, A.M., et al., *Human milk glycobiome and its impact on the infant gastrointestinal microbiota.* Proc Natl Acad Sci U S A, 2011. **108 Suppl 1**: p. 4653-8.

54. Sela, D.A. and D.A. Mills, *Nursing our microbiota: molecular linkages between bifidobacteria and milk oligosaccharides.* Trends Microbiol, 2010. **18**(7): p. 298-307.

55. Penders, J., et al., *Factors influencing the composition of the intestinal microbiota in early infancy.* Pediatrics, 2006. **118**(2): p. 511-21.

56. Benno, Y., K. Sawada, and T. Mitsuoka, *The intestinal microflora of infants: composition of fecal flora in breast-fed and bottle-fed infants.* Microbiol Immunol, 1984. **28**(9): p. 975-86.

57. Donovan, S.M., et al., *Host-microbe interactions in the neonatal intestine: role of human milk oligosaccharides.* Adv Nutr, 2012. **3**(3): p. 450S-5S.

58. Langhendries, J.P., et al., *Effect of a fermented infant formula containing viable bifidobacteria on the fecal flora composition and pH of healthy full-term infants.* J Pediatr Gastroenterol Nutr, 1995. **21**(2): p. 177-81.

59. Yoshioka, H., K. Iseki, and K. Fujita, *Development and differences of intestinal flora in the neonatal period in breast-fed and bottle-fed infants.* Pediatrics, 1983. **72**(3): p. 317-21.

60. Makino, H., et al., *Mother-to-infant transmission of intestinal bifidobacterial strains has an impact on the early development of vaginally delivered infant's microbiota.* PLoS One, 2013. **8**(11): p. e78331.

61. Jost, T., et al., *Vertical mother-neonate transfer of maternal gut bacteria via breastfeeding.* Environ Microbiol, 2014. **16**(9): p. 2891-904.

62. Saavedra, J.M., et al., *Feeding of Bifidobacterium bifidum and Streptococcus thermophilus to infants in hospital for prevention of diarrhoea and shedding of rotavirus.* Lancet, 1994. **344**(8929): p. 1046-9.

63. Benno, Y. and T. Mitsuoka, *Impact of Bifidobacterium longum on human fecal microflora.* Microbiol Immunol, 1992. **36**(7): p. 683-94.

64. Martinez, F.A., et al., *Bacteriocin production by Bifidobacterium spp. A review.* Biotechnol Adv, 2013. **31**(4): p. 482-8.

65. Mazmanian, S.K., J.L. Round, and D.L. Kasper, *A microbial symbiosis factor prevents intestinal inflammatory disease.* Nature, 2008. **453**(7195): p. 620-5.

66. Mazmanian, S.K., et al., *An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system.* Cell, 2005. **122**(1): p. 107-18.

67. Wrzosek, L., et al., *Bacteroides thetaiotaomicron and Faecalibacterium prausnitzii influence the production of mucus glycans and the development of goblet cells in the colonic epithelium of a gnotobiotic model rodent.* BMC Biol, 2013. **11**: p. 61.

68. Johansson, M.E., J.M. Larsson, and G.C. Hansson, *The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions.* Proc Natl Acad Sci U S A, 2011. **108 Suppl 1**: p. 4659-65.

69. Johansson, M.E., et al., *Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis.* Gut, 2014. **63**(2): p. 281-91.

70. Puertollano, E., S. Kolida, and P. Yaqoob, *Biological significance of short-chain fatty acid metabolism by the intestinal microbiome.* Curr Opin Clin Nutr Metab Care, 2014. **17**(2): p. 139-44.

71. Fuhrer, A., et al., *Milk sialyllactose influences colitis in mice through selective intestinal bacterial colonization.* J Exp Med, 2010. **207**(13): p. 2843-54.

72. Kurakevich, E., et al., *Milk oligosaccharide sialyl(alpha2,3)lactose activates intestinal CD11c+ cells through TLR4.* Proc Natl Acad Sci U S A, 2013. **110**(43): p. 17444-9.

73. Oozeer, R., et al., *Intestinal microbiology in early life: specific prebiotics can have similar functionalities as human-milk oligosaccharides.* Am J Clin Nutr, 2013. **98**(2): p. 561S-71S.

74. Mugambi, M.N., et al., *Synbiotics, probiotics or prebiotics in infant formula for full term infants: a systematic review.* Nutr J, 2012. **11**: p. 81.

75. Mugambi, M.N., et al., *Probiotics, prebiotics infant formula use in preterm or low birth weight infants: a systematic review.* Nutr J, 2012. **11**: p. 58.

76. Karlsson, K.A., *Meaning and therapeutic potential of microbial recognition of host glycoconjugates.* Mol Microbiol, 1998. **29**(1): p. 1-11.

77. Imberty, A. and A. Varrot, *Microbial recognition of human cell surface glycoconjugates.* Curr Opin Struct Biol, 2008. **18**(5): p. 567-76.

78. Newburg, D.S., G.M. Ruiz-Palacios, and A.L. Morrow, *Human milk glycans protect infants against enteric pathogens.* Annu Rev Nutr, 2005. **25**: p. 37-58.

79. Matrosovich, M.N., et al., *Probing of the receptor-binding sites of the H1 and H3 influenza A and influenza B virus hemagglutinins by synthetic and natural sialosides.* Virology, 1993. **196**(1): p. 111-21.

80. Yu, Y., et al., *Functional glycomic analysis of human milk glycans reveals presence of virus receptors and embryonic stem cell biomarkers.* J Biol Chem, 2012. **287**(53): p. 44784-44799.

81. Chichlowski, M., et al., *The influence of milk oligosaccharides on microbiota of infants: opportunities for formulas.* Annu Rev Food Sci Technol, 2011. **2**: p. 331-51.

82. Ruiz-Palacios, G.M., et al., *Campylobacter jejuni binds intestinal H(O) antigen (Fuc alpha 1, 2Gal beta 1, 4GlcNAc), and fucosyloligosaccharides of human milk inhibit its binding and infection.* J Biol Chem, 2003. **278**(16): p. 14112-20.

83. Morrow, A.L., et al., *Human milk oligosaccharides are associated with protection against diarrhea in breast-fed infants.* J Pediatr, 2004. **145**(3): p. 297-303.

84. Newburg, D.S., et al., *Fucosylated oligosaccharides of human milk protect suckling mice from heat-stabile enterotoxin of Escherichia coli.* J Infect Dis, 1990. **162**(5): p. 1075-80.

85. Eiwegger, T., et al., *Human milk--derived oligosaccharides and plant-derived oligosaccharides stimulate cytokine production of cord blood T-cells in vitro.* Pediatr Res, 2004. **56**(4): p. 536-40.

86. Eiwegger, T., et al., *Prebiotic oligosaccharides: in vitro evidence for gastrointestinal epithelial transfer and immunomodulatory properties.* Pediatr Allergy Immunol, 2010. **21**(8): p. 1179-88.

87. Bode, L., et al., *Inhibition of monocyte, lymphocyte, and neutrophil adhesion to endothelial cells by human milk oligosaccharides.* Thromb Haemost, 2004. **92**(6): p. 1402-10.

88. Bode, L., et al., *Human milk oligosaccharides reduce platelet-neutrophil complex formation leading to a decrease in neutrophil beta 2 integrin expression.* J Leukoc Biol, 2004. **76**(4): p. 820-6.

89. He, Y., et al., *The human milk oligosaccharide 2'-fucosyllactose modulates CD14 expression in human enterocytes, thereby attenuating LPS-induced inflammation.* Gut, 2014.

90. Cederlund, A., et al., *Lactose in human breast milk an inducer of innate immunity with implications for a role in intestinal homeostasis.* PLoS One, 2013. **8**(1): p. e53876.

91. He, Y., et al., *Human colostrum oligosaccharides modulate major immunologic pathways of immature human intestine.* Mucosal Immunol, 2014. **7**(6): p. 1326-39.

92. Jantscher-Krenn, E., et al., *The human milk oligosaccharide disialyllacto-N-tetraose prevents necrotising enterocolitis in neonatal rats.* Gut, 2012. **61**(10): p. 1417-25.

93. Neu, J. and W.A. Walker, *Necrotizing enterocolitis.* N Engl J Med, 2011. **364**(3): p. 255-64.

94. Lucas, A. and T.J. Cole, *Breast milk and neonatal necrotising enterocolitis.* Lancet, 1990. **336**(8730): p. 1519-23.

95. Sullivan, S., et al., *An exclusively human milk-based diet is associated with a lower rate of necrotizing enterocolitis than a diet of human milk and bovine milk-based products.* J Pediatr, 2010. **156**(4): p. 562-7 e1.

96. Castillo-Courtade, L., et al., *Attenuation of food allergy symptoms following treatment with human milk oligosaccharides in a mouse model.* Allergy, 2015. **70**(9): p. 1091-102.

97. Velupillai, P. and D.A. Harn, *Oligosaccharide-specific induction of interleukin 10 production by B220+ cells from schistosome-infected mice: a mechanism for regulation of CD4+ T-cell subsets.* Proc Natl Acad Sci U S A, 1994. **91**(1): p. 18-22.

98. Angeloni, S., et al., *Glycoprofiling with micro-arrays of glycoconjugates and lectins.* Glycobiology, 2005. **15**(1): p. 31-41.

99. Kuntz, S., S. Rudloff, and C. Kunz, *Oligosaccharides from human milk influence growth-related characteristics of intestinally transformed and non-transformed intestinal cells.* Br J Nutr, 2008. **99**(3): p. 462-71.

100. Kuntz, S., C. Kunz, and S. Rudloff, *Oligosaccharides from human milk induce growth arrest via G2/M by influencing growth-related cell cycle genes in intestinal epithelial cells.* Br J Nutr, 2009. **101**(9): p. 1306-15.

101. Holscher, H.D., S.R. Davis, and K.A. Tappenden, *Human milk oligosaccharides influence maturation of human intestinal Caco-2Bbe and HT-29 cell lines.* J Nutr, 2014. **144**(5): p. 586-91.

102. Krug, M., et al., *Fucose and fucose-containing sugar epitopes enhance hippocampal long-term potentiation in the freely moving rat.* Brain Res, 1994. **643**(1-2): p. 130-5.

103. Matthies, H., S. Staak, and M. Krug, *Fucose and fucosyllactose enhance in-vitro hippocampal long-term potentiation.* Brain Res, 1996. **725**(2): p. 276-80.

104. Vazquez, E., et al., *Effects of a human milk oligosaccharide, 2'-fucosyllactose, on hippocampal long-term potentiation and learning capabilities in rodents.* J Nutr Biochem, 2015. **26**(5): p. 455-65.

105. Bienenstock, J., et al., *Fucosylated but not sialylated milk oligosaccharides diminish colon motor contractions.* PLoS One, 2013. **8**(10): p. e76236.

106. Barondes, S.H., et al., *Galectins: a family of animal beta-galactoside-binding lectins.* Cell, 1994. **76**(4): p. 597-8.

107. Cummings, R.D., Liu, F. T., *Galectins*, in *Essentials of Glycobiology*, A. Varki, Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W., Etzler, M.E., Editor 2009, Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY).

108. Leffler, H., et al., *Introduction to galectins.* Glycoconj J, 2004. **19**(7-9): p. 433-40.

109. Lobsanov, Y.D., et al., *X-ray crystal structure of the human dimeric S-Lac lectin, L-14-II, in complex with lactose at 2.9-A resolution.* J Biol Chem, 1993. **268**(36): p. 27034-8.

110. Teichberg, V.I., et al., *A beta-D-galactoside binding protein from electric organ tissue of Electrophorus electricus.* Proc Natl Acad Sci U S A, 1975. **72**(4): p. 1383-7.

111. Powell, J.T., *Purification and properties of lung lectin. Rat lung and human lung beta-galactoside-binding proteins.* Biochem J, 1980. **187**(1): p. 123-9.

112. Levi, G. and V.I. Teichberg, *Isolation and physicochemical characterization of electrolectin, a beta-D-galactoside binding lectin from the electric organ of Electrophorus electricus.* J Biol Chem, 1981. **256**(11): p. 5735-40.

113. Than, N.G., et al., *A primate subfamily of galectins expressed at the maternal-fetal interface that promote immune cell death.* Proc Natl Acad Sci U S A, 2009. **106**(24): p. 9731-6.

114. Hotta, K., et al., *Galectin-12, an Adipose-expressed Galectin-like Molecule Possessing Apoptosis-inducing Activity.* J Biol Chem, 2001. **276**(36): p. 34089-97.

115. Yang, R.Y., et al., *Cell cycle regulation by galectin-12, a new member of the galectin superfamily.* J Biol Chem, 2001. **276**(23): p. 20252-60.

116. Leonidas, D.D., et al., *Crystal structure of human Charcot-Leyden crystal protein, an eosinophil lysophospholipase, identifies it as a new member of the carbohydrate-binding family of galectins.* Structure, 1995. **3**(12): p. 1379-93.

117. Swaminathan, G.J., et al., *Selective recognition of mannose by the human eosinophil Charcot-Leyden crystal protein (galectin-10): a crystallographic study at 1.8 A resolution.* Biochemistry, 1999. **38**(42): p. 13837-43.

118. Dyer, K.D. and H.F. Rosenberg, *Eosinophil Charcot-Leyden crystal protein binds to beta-galactoside sugars.* Life Sci, 1996. **58**(23): p. 2073-82.

119. Yang, R.Y., et al., *Ablation of a galectin preferentially expressed in adipocytes increases lipolysis, reduces adiposity, and improves insulin sensitivity in mice.* Proc Natl Acad Sci U S A, 2011. **108**(46): p. 18696-701.

120. Than, N.G., et al., *Functional analyses of placental protein 13/galectin-13.* Eur J Biochem, 2004. **271**(6): p. 1065-78.

121. Hirabayashi, J. and K. Kasai, *The family of metazoan metal-independent beta-galactoside-binding lectins: structure, function and molecular evolution.* Glycobiology, 1993. **3**(4): p. 297-304.

122. Stowell, S.R., et al., *Ligand reduces galectin-1 sensitivity to oxidative inactivation by enhancing dimer formation.* J Biol Chem, 2009. **284**(8): p. 4989-99.

123. Stowell, S.R., et al., *Human galectin-1, -2, and -4 induce surface exposure of phosphatidylserine in activated human neutrophils but not in activated T cells.* Blood, 2007. **109**(1): p. 219-27.

124. Stowell, S.R., et al., *Galectin-1 induces reversible phosphatidylserine exposure at the plasma membrane.* Mol Biol Cell, 2009. **20**(5): p. 1408-18.

125. Stowell, S.R., et al., *Differential roles of galectin-1 and galectin-3 in regulating leukocyte viability and cytokine secretion.* J Immunol, 2008. **180**(5): p. 3091-102.

126. Hirabayashi, J. and K. Kasai, *Effect of amino acid substitution by sited-directed mutagenesis on the carbohydrate recognition and stability of human 14-kDa beta-galactoside-binding lectin.* J Biol Chem, 1991. **266**(35): p. 23648-53.

127. Barondes, S.H., et al., *Galectins. Structure and function of a large family of animal lectins.* J Biol Chem, 1994. **269**(33): p. 20807-10.

128. Stowell, S.R., et al., *Galectin-1, -2, and -3 exhibit differential recognition of sialylated glycans and blood group antigens.* J Biol Chem, 2008. **283**(15): p. 10109-23.

129. Leffler, H. and S.H. Barondes, *Specificity of binding of three soluble rat lung lectins to substituted and unsubstituted mammalian beta-galactosides.* J Biol Chem, 1986. **261**(22): p. 10119-26.

130. Knibbs, R.N., et al., *Carbohydrate-binding protein 35. II. Analysis of the interaction of the recombinant polypeptide with saccharides.* J Biol Chem, 1993. **268**(20): p. 14940-7.

131. Huflejt, M.E. and H. Leffler, *Galectin-4 in normal tissues and cancer.* Glycoconj J, 2004. **20**(4): p. 247-55.

132. Meynier, C., et al., *NMR and MD investigations of human galectin-1/oligosaccharide complexes.* Biophys J, 2009. **97**(12): p. 3168-77.

133. Henrick, K., et al., *Evidence for subsites in the galectins involved in sugar binding at the nonreducing end of the central galactose of oligosaccharide ligands: sequence analysis, homology modeling and mutagenesis studies of hamster galectin-3.* Glycobiology, 1998. **8**(1): p. 45-57.

134. Lotz, M.M., et al., *Decreased expression of Mac-2 (carbohydrate binding protein 35) and loss of its nuclear localization are associated with the neoplastic progression of colon carcinoma.* Proc Natl Acad Sci U S A, 1993. **90**(8): p. 3466-70.

135. Liu, F.T., R.J. Patterson, and J.L. Wang, *Intracellular functions of galectins.* Biochim Biophys Acta, 2002. **1572**(2-3): p. 263-73.

136. Villeneuve, C., et al., *Mitochondrial proteomic approach reveals galectin-7 as a novel BCL-2 binding protein in human cells.* Mol Biol Cell, 2011. **22**(7): p. 999-1013.

137. Thurston, T.L., et al., *Galectin 8 targets damaged vesicles for autophagy to defend cells against bacterial invasion.* Nature, 2012. **482**(7385): p. 414-8.

138. Dagher, S.F., J.L. Wang, and R.J. Patterson, *Identification of galectin-3 as a factor in pre-mRNA splicing.* Proc Natl Acad Sci U S A, 1995. **92**(4): p. 1213-7.

139. Hughes, R.C., *Secretion of the galectin family of mammalian carbohydrate-binding proteins.* Biochim Biophys Acta, 1999. **1473**(1): p. 172-85.

140. Magnaldo, T., F. Bernerd, and M. Darmon, *Galectin-7, a human 14-kDa S-lectin, specifically expressed in keratinocytes and sensitive to retinoic acid.* Dev Biol, 1995. **168**(2): p. 259-71.

141. Matsumoto, R., et al., *Human ecalectin, a variant of human galectin-9, is a novel eosinophil chemoattractant produced by T lymphocytes.* J Biol Chem, 1998. **273**(27): p. 16976-84.

142. Liu, F.T., et al., *Expression and function of galectin-3, a beta-galactoside-binding lectin, in human monocytes and macrophages.* Am J Pathol, 1995. **147**(4): p. 1016-28.

143. Hadari, Y.R., et al., *Galectin-8 binding to integrins inhibits cell adhesion and induces apoptosis.* J Cell Sci, 2000. **113 (Pt 13)**: p. 2385-97.

144. Cho, M. and R.D. Cummings, *Galectin-1, a beta-galactoside-binding lectin in Chinese hamster ovary cells. II. Localization and biosynthesis.* J Biol Chem, 1995. **270**(10): p. 5207-12.

145. Sato, M., et al., *Quantification of galectin-7 and its localization in adult mouse tissues.* J Biochem, 2002. **131**(2): p. 255-60.

146. Sacchettini, J.C., L.G. Baum, and C.F. Brewer, *Multivalent protein-carbohydrate interactions. A new paradigm for supermolecular assembly and signal transduction.* Biochemistry, 2001. **40**(10): p. 3009-15.

147. Elola, M.T., et al., *Assembly, organization and regulation of cell-surface receptors by lectin-glycan complexes.* Biochem J, 2015. **469**(1): p. 1-16.

148. Gupta, D. and C.F. Brewer, *Homogeneous aggregation of the 14-kDa beta-galactoside specific vertebrate lectin complex with asialofetuin in mixed systems.* Biochemistry, 1994. **33**(18): p. 5526-30.

149. Pace, K.E., et al., *Restricted receptor segregation into membrane microdomains occurs on human T cells during apoptosis induced by galectin-1.* J Immunol, 1999. **163**(7): p. 3801-11.

150. Nieminen, J., et al., *Visualization of galectin-3 oligomerization on the surface of neutrophils and endothelial cells using fluorescence resonance energy transfer.* J Biol Chem, 2007. **282**(2): p. 1374-83.

151. Dias-Baruffi, M., et al., *Dimeric galectin-1 induces surface exposure of phosphatidylserine and phagocytic recognition of leukocytes without inducing apoptosis.* J Biol Chem, 2003. **278**(42): p. 41282-93.

152. Cao, Z., et al., *Galectins-3 and -7, but not galectin-1, play a role in re-epithelialization of wounds.* J Biol Chem, 2002. **277**(44): p. 42299-305.

153. Kariya, Y., et al., *Bisecting GlcNAc residues on laminin-332 down-regulate galectin-3-dependent keratinocyte motility.* J Biol Chem, 2010. **285**(5): p. 3330-40.

154. Amano, M., et al., *The ST6Gal I sialyltransferase selectively modifies N-glycans on CD45 to negatively regulate galectin-1-induced CD45 clustering, phosphatase modulation, and T cell death.* J Biol Chem, 2003. **278**(9): p. 7469-75.

155. Demetriou, M., et al., *Negative regulation of T-cell activation and autoimmunity by Mgat5 N-glycosylation.* Nature, 2001. **409**(6821): p. 733-9.

156. Earl, L.A., S. Bi, and L.G. Baum, *N- and O-glycans modulate galectin-1 binding, CD45 signaling, and T cell death.* J Biol Chem, 2010. **285**(4): p. 2232-44.

157. Laderach, D.J., et al., *Dissecting the signal transduction pathways triggered by galectin-glycan interactions in physiological and pathological settings.* IUBMB Life, 2010. **62**(1): p. 1-13.

158. Hughes, R.C., *Galectins as modulators of cell adhesion.* Biochimie, 2001. **83**(7): p. 667-76.

159. Yang, R.Y., G.A. Rabinovich, and F.T. Liu, *Galectins: structure, function and therapeutic potential.* Expert Rev Mol Med, 2008. **10**: p. e17.

160. Liu, F.T., R.Y. Yang, and D.K. Hsu, *Galectins in acute and chronic inflammation.* Ann N Y Acad Sci, 2012. **1253**: p. 80-91.

161. Dumic, J., S. Dabelic, and M. Flogel, *Galectin-3: an open-ended story.* Biochim Biophys Acta, 2006. **1760**(4): p. 616-35.

162. Perillo, N.L., et al., *Apoptosis of T cells mediated by galectin-1.* Nature, 1995. **378**(6558): p. 736-9.

163. Hsu, D.K., R.Y. Yang, and F.T. Liu, *Galectins in apoptosis.* Methods Enzymol, 2006. **417**: p. 256-73.

164. Stowell, S.R., et al., *Dimeric Galectin-8 induces phosphatidylserine exposure in leukocytes through polylactosamine recognition by the C-terminal domain.* J Biol Chem, 2008. **283**(29): p. 20547-59.

165. Karmakar, S., R.D. Cummings, and R.P. McEver, *Contributions of Ca2+ to galectin-1-induced exposure of phosphatidylserine on activated neutrophils.* J Biol Chem, 2005. **280**(31): p. 28623-31.

166. Matsumoto, R., et al., *Biological activities of ecalectin: a novel eosinophil-activating factor.* J Immunol, 2002. **168**(4): p. 1961-7.

167. Sano, H., et al., *Human galectin-3 is a novel chemoattractant for monocytes and macrophages.* J Immunol, 2000. **165**(4): p. 2156-64.

168. Cooper, D., L.V. Norling, and M. Perretti, *Novel insights into the inhibitory effects of Galectin-1 on neutrophil recruitment under flow.* J Leukoc Biol, 2008. **83**(6): p. 1459-66.

169. Chen, I.J., H.L. Chen, and M. Demetriou, *Lateral compartmentalization of T cell receptor versus CD45 by galectin-N-glycan binding and microfilaments coordinate basal and activation signaling.* J Biol Chem, 2007. **282**(48): p. 35361-72.

170. MacKinnon, A.C., et al., *Regulation of alternative macrophage activation by galectin-3.* J Immunol, 2008. **180**(4): p. 2650-8.

171. Frigeri, L.G., R.I. Zuberi, and F.T. Liu, *Epsilon BP, a beta-galactoside-binding animal lectin, recognizes IgE receptor (Fc epsilon RI) and activates mast cells.* Biochemistry, 1993. **32**(30): p. 7644-9.

172. Zuberi, R.I., L.G. Frigeri, and F.T. Liu, *Activation of rat basophilic leukemia cells by epsilon BP, an IgE-binding endogenous lectin.* Cell Immunol, 1994. **156**(1): p. 1-12.

173. Niki, T., et al., *Galectin-9 is a high affinity IgE-binding lectin with anti-allergic effect by blocking IgE-antigen complex formation.* J Biol Chem, 2009. **284**(47): p. 32344-52.

174. Carlsson, S., et al., *Affinity of galectin-8 and its carbohydrate recognition domains for ligands in solution and at the cell surface.* Glycobiology, 2007. **17**(6): p. 663-76.

175. Yamaoka, A., et al., *A human lectin, galectin-3 (epsilon bp/Mac-2), stimulates superoxide production by neutrophils.* J Immunol, 1995. **154**(7): p. 3479-87.

176. Jeng, K.C., L.G. Frigeri, and F.T. Liu, *An endogenous lectin, galectin-3 (epsilon BP/Mac-2), potentiates IL-1 production by human monocytes.* Immunol Lett, 1994. **42**(3): p. 113-6.

177. van der Leij, J., et al., *Dimeric galectin-1 induces IL-10 production in T-lymphocytes: an important tool in the regulation of the immune response.* J Pathol, 2004. **204**(5): p. 511-8.

178. Ilarregui, J.M., et al., *Tolerogenic signals delivered by dendritic cells to T cells through a galectin-1-driven immunoregulatory circuit involving interleukin 27 and interleukin 10.* Nat Immunol, 2009. **10**(9): p. 981-91.

179. Chen, H.Y., et al., *Galectins as bacterial sensors in the host innate response.* Curr Opin Microbiol, 2014. **17**: p. 75-81.

180. Baum, L.G., et al., *Microbe-Host Interactions are Positively and Negatively Regulated by Galectin-Glycan Interactions.* Front Immunol, 2014. **5**: p. 284.

181. Vasta, G.R., *Roles of galectins in infection.* Nat Rev Microbiol, 2009. **7**(6): p. 424-38.

182. Cerliani, J.P., et al., *Expanding the universe of cytokines and pattern recognition receptors: galectins and glycans in innate immunity.* J Clin Immunol, 2011. **31**(1): p. 10-21.

183. Mirelman, D., et al., *Inhibition of fungal growth by wheat germ agglutinin.* Nature, 1975. **256**(5516): p. 414-6.

184. Kohatsu, L., et al., *Galectin-3 induces death of Candida species expressing specific beta-1,2-linked mannans.* J Immunol, 2006. **177**(7): p. 4718-26.

185. Farnworth, S.L., et al., *Galectin-3 reduces the severity of pneumococcal pneumonia by augmenting neutrophil function.* Am J Pathol, 2008. **172**(2): p. 395-405.

186. Stowell, S.R., et al., *Microbial glycan microarrays define key features of host-microbial interactions.* Nat Chem Biol, 2014. **10**(6): p. 470-6.

187. Stowell, S.R., et al., *Innate immune lectins kill bacteria expressing blood group antigen.* Nat Med, 2010. **16**(3): p. 295-301.

188. Fowler, M., et al., *Galectin-3 binds to Helicobacter pylori O-antigen: it is upregulated and rapidly secreted by gastric epithelial cells in response to H. pylori adhesion.* Cell Microbiol, 2006. **8**(1): p. 44-54.

189. Cooper, D.N. and S.H. Barondes, *Evidence for export of a muscle lectin from cytosol to extracellular matrix and for a novel secretory mechanism.* J Cell Biol, 1990. **110**(5): p. 1681-91.

190. Cooper, D.N., S.M. Massa, and S.H. Barondes, *Endogenous muscle lectin inhibits myoblast adhesion to laminin.* J Cell Biol, 1991. **115**(5): p. 1437-48.

191. Nakajima, K., et al., *Galectin-3 inhibits osteoblast differentiation through notch signaling.* Neoplasia, 2014. **16**(11): p. 939-49.

192. Tanikawa, R., et al., *Galectin-9 induces osteoblast differentiation through the CD44/Smad signaling pathway.* Biochem Biophys Res Commun, 2010. **394**(2): p. 317-22.

193. Tanikawa, R., et al., *Interaction of galectin-9 with lipid rafts induces osteoblast proliferation through the c-Src/ERK signaling pathway.* J Bone Miner Res, 2008. **23**(2): p. 278-86.

194. Kiwaki, K., et al., *Galectin-3 stimulates preadipocyte proliferation and is up-regulated in growing adipose tissue.* Obesity (Silver Spring), 2007. **15**(1): p. 32-9.

195. Levy, Y., et al., *Galectin-8 functions as a matricellular modulator of cell adhesion.* J Biol Chem, 2001. **276**(33): p. 31285-95.

196. Panjwani, N., *Role of galectins in re-epithelialization of wounds.* Ann Transl Med, 2014. **2**(9): p. 89.

197. Funasaka, T., A. Raz, and P. Nangia-Makker, *Galectin-3 in angiogenesis and metastasis.* Glycobiology, 2014. **24**(10): p. 886-91.

198. Troncoso, M.F., et al., *Galectin-8: a matricellular lectin with key roles in angiogenesis.* Glycobiology, 2014. **24**(10): p. 907-14.

199. Thijssen, V.L. and A.W. Griffioen, *Galectin-1 and -9 in angiogenesis: a sweet couple.* Glycobiology, 2014. **24**(10): p. 915-20.

200. Moisan, S., et al., *Upregulation of galectin-7 in murine lymphoma cells is associated with progression toward an aggressive phenotype.* Leukemia, 2003. **17**(4): p. 751-9.

201. Magnaldo, T., D. Fowlis, and M. Darmon, *Galectin-7, a marker of all types of stratified epithelia.* Differentiation, 1998. **63**(3): p. 159-68.

202. Demers, M., T. Magnaldo, and Y. St-Pierre, *A novel function for galectin-7: promoting tumorigenesis by up-regulating MMP-9 gene expression.* Cancer Res, 2005. **65**(12): p. 5205-10.

203. Lalancette, M., et al., *Resistance of ICAM-1-deficient mice to metastasis overcome by increased aggressiveness of lymphoma cells.* Blood, 2000. **95**(1): p. 314-9.

204. Fischer, C., et al., *Galectin-1 interacts with the {alpha}5{beta}1 fibronectin receptor to restrict carcinoma cell growth via induction of p21 and p27.* J Biol Chem, 2005. **280**(44): p. 37266-77.

205. Liu, F.T. and G.A. Rabinovich, *Galectins as modulators of tumour progression.* Nat Rev Cancer, 2005. **5**(1): p. 29-41.

206. Rizqiawan, A., et al., *Autocrine galectin-1 promotes collective cell migration of squamous cell carcinoma cells through up-regulation of distinct integrins.* Biochem Biophys Res Commun, 2013. **441**(4): p. 904-10.

207. Zhao, Q., et al., *Interaction between circulating galectin-3 and cancer-associated MUC1 enhances tumour cell homotypic aggregation and prevents anoikis.* Mol Cancer, 2010. **9**: p. 154.

208. Sanchez-Ruderisch, H., et al., *Tumor suppressor p16 INK4a: Downregulation of galectin-3, an endogenous competitor of the pro-anoikis effector galectin-1, in a pancreatic carcinoma model.* FEBS J, 2010. **277**(17): p. 3552-63.

209. Chen, C., et al., *Increased circulation of galectin-3 in cancer induces secretion of metastasis-promoting cytokines from blood vascular endothelium.* Clin Cancer Res, 2013. **19**(7): p. 1693-704.

210. Khaldoyanidi, S.K., et al., *MDA-MB-435 human breast carcinoma cell homo- and heterotypic adhesion under flow conditions is mediated in part by Thomsen-Friedenreich antigen-galectin-3 interactions.* J Biol Chem, 2003. **278**(6): p. 4127-34.

211. Barrow, H., et al., *Serum galectin-2, -4, and -8 are greatly increased in colon and breast cancer patients and promote cancer cell adhesion to blood vascular endothelium.* Clin Cancer Res, 2011. **17**(22): p. 7035-46.

212. Nobumoto, A., et al., *Galectin-9 suppresses tumor metastasis by blocking adhesion to endothelium and extracellular matrices.* Glycobiology, 2008. **18**(9): p. 735-44.

213. Nishi, N., et al., *Galectin-8 modulates neutrophil function via interaction with integrin alphaM.* Glycobiology, 2003. **13**(11): p. 755-63.

214. Fortuna-Costa, A., et al., *Extracellular galectin-3 in tumor progression and metastasis.* Front Oncol, 2014. **4**: p. 138.

215. Poirier, F. and E.J. Robertson, *Normal development of mice carrying a null mutation in the gene encoding the L14 S-type lectin.* Development, 1993. **119**(4): p. 1229-36.

216. Colnot, C., et al., *Embryonic implantation in galectin 1/galectin 3 double mutant mice.* Dev Dyn, 1998. **211**(4): p. 306-13.

217. Gendronneau, G., et al., *Galectin-7 in the control of epidermal homeostasis after injury.* Mol Biol Cell, 2008. **19**(12): p. 5541-9.

218. van den Brule, F.A., et al., *Galectin-1 modulates human melanoma cell adhesion to laminin.* Biochem Biophys Res Commun, 1995. **209**(2): p. 760-7.

219. Jiang, K., et al., *Galectin-3 regulates desmoglein-2 and intestinal epithelial intercellular adhesion.* J Biol Chem, 2014. **289**(15): p. 10510-7.

220. Granucci, F., M. Foti, and P. Ricciardi-Castagnoli, *Dendritic cell biology.* Adv Immunol, 2005. **88**: p. 193-233.

221. Iwasaki, A. and R. Medzhitov, *Toll-like receptor control of the adaptive immune responses.* Nat Immunol, 2004. **5**(10): p. 987-95.

222. Ueno, H., et al., *Dendritic cell subsets in health and disease.* Immunol Rev, 2007. **219**: p. 118-42.

223. Mascanfroni, I.D., et al., *Endogenous lectins shape the function of dendritic cells and tailor adaptive immunity: mechanisms and biomedical applications.* Int Immunopharmacol, 2011. **11**(7): p. 833-41.

224. Boltjes, A. and F. van Wijk, *Human dendritic cell functional specialization in steady-state and inflammation.* Front Immunol, 2014. **5**: p. 131.

225. Niess, J.H. and H.C. Reinecker, *Lamina propria dendritic cells in the physiology and pathology of the gastrointestinal tract.* Curr Opin Gastroenterol, 2005. **21**(6): p. 687-91.

226. Arques, J.L., et al., *Salmonella induces flagellin- and MyD88-dependent migration of bacteria-capturing dendritic cells into the gut lumen.* Gastroenterology, 2009. **137**(2): p. 579-87, 587 e1-2.

227. Rescigno, M., et al., *Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria.* Nat Immunol, 2001. **2**(4): p. 361-7.

228. Rescigno, M., et al., *Dendritic cells shuttle microbes across gut epithelial monolayers.* Immunobiology, 2001. **204**(5): p. 572-81.

229. Chieppa, M., et al., *Dynamic imaging of dendritic cell extension into the small bowel lumen in response to epithelial cell TLR engagement.* J Exp Med, 2006. **203**(13): p. 2841-52.

230. Lelouard, H., et al., *Peyer's patch dendritic cells sample antigens by extending dendrites through M cell-specific transcellular pores.* Gastroenterology, 2012. **142**(3): p. 592-601 e3.

231. Pasare, C. and R. Medzhitov, *Toll-like receptors: linking innate and adaptive immunity.* Microbes Infect, 2004. **6**(15): p. 1382-7.

232. Geijtenbeek, T.B. and S.I. Gringhuis, *Signalling through C-type lectin receptors: shaping immune responses.* Nat Rev Immunol, 2009. **9**(7): p. 465-79.

233. Dietz, A.B., et al., *Maturation of human monocyte-derived dendritic cells studied by microarray hybridization.* Biochem Biophys Res Commun, 2000. **275**(3): p. 731-8.

234. Drickamer, K., *Two distinct classes of carbohydrate-recognition domains in animal lectins.* J Biol Chem, 1988. **263**(20): p. 9557-60.

235. Drickamer, K. and M.E. Taylor, *Biology of animal lectins.* Annu Rev Cell Biol, 1993. **9**: p. 237-64.

236. Willment, J.A., S. Gordon, and G.D. Brown, *Characterization of the human beta -glucan receptor and its alternatively spliced isoforms.* J Biol Chem, 2001. **276**(47): p. 43818-23.

237. Higashi, N., et al., *The macrophage C-type lectin specific for galactose/N-acetylgalactosamine is an endocytic receptor expressed on monocyte-derived immature dendritic cells.* J Biol Chem, 2002. **277**(23): p. 20686-93.

238. van Vliet, S.J., J.J. Garcia-Vallejo, and Y. van Kooyk, *Dendritic cells and C-type lectin receptors: coupling innate to adaptive immune responses.* Immunol Cell Biol, 2008. **86**(7): p. 580-7.

239. Osorio, F. and C. Reis e Sousa, *Myeloid C-type lectin receptors in pathogen recognition and host defense.* Immunity, 2011. **34**(5): p. 651-64.

240. Dambuza, I.M. and G.D. Brown, *C-type lectins in immunity: recent developments.* Curr Opin Immunol, 2015. **32**: p. 21-7.

241. Curtis, B.M., S. Scharnowske, and A.J. Watson, *Sequence and expression of a membrane-associated C-type lectin that exhibits CD4-independent binding of human immunodeficiency virus envelope glycoprotein gp120.* Proc Natl Acad Sci U S A, 1992. **89**(17): p. 8356-60.

242. Geijtenbeek, T.B., et al., *Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses.* Cell, 2000. **100**(5): p. 575-85.

243. Geijtenbeek, T.B., et al., *DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells.* Cell, 2000. **100**(5): p. 587-97.

244. Jameson, B., et al., *Expression of DC-SIGN by dendritic cells of intestinal and genital mucosae in humans and rhesus macaques.* J Virol, 2002. **76**(4): p. 1866-75.

245. Mitchell, D.A., A.J. Fadden, and K. Drickamer, *A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. Subunit organization and binding to multivalent ligands.* J Biol Chem, 2001. **276**(31): p. 28939-45.

246. Appelmelk, B.J., et al., *Cutting edge: carbohydrate profiling identifies new pathogens that interact with dendritic cell-specific ICAM-3-grabbing nonintegrin on dendritic cells.* J Immunol, 2003. **170**(4): p. 1635-9.

247. van Die, I., et al., *The dendritic cell-specific C-type lectin DC-SIGN is a receptor for Schistosoma mansoni egg antigens and recognizes the glycan antigen Lewis x.* Glycobiology, 2003. **13**(6): p. 471-8.

248. van Liempt, E., et al., *Schistosoma mansoni soluble egg antigens are internalized by human dendritic cells through multiple C-type lectins and suppress TLR-induced dendritic cell activation.* Mol Immunol, 2007. **44**(10): p. 2605-15.

249. Svajger, U., et al., *C-type lectin DC-SIGN: an adhesion, signalling and antigen-uptake molecule that guides dendritic cells in immunity.* Cell Signal, 2010. **22**(10): p. 1397-405.

250. Engering, A., et al., *The dendritic cell-specific adhesion receptor DC-SIGN internalizes antigen for presentation to T cells.* J Immunol, 2002. **168**(5): p. 2118-26.

251. Geijtenbeek, T.B., et al., *Mycobacteria target DC-SIGN to suppress dendritic cell function.* J Exp Med, 2003. **197**(1): p. 7-17.

252. Guo, Y., et al., *Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR.* Nat Struct Mol Biol, 2004. **11**(7): p. 591-8.

253. Gringhuis, S.I., et al., *Carbohydrate-specific signaling through the DC-SIGN signalosome tailors immunity to Mycobacterium tuberculosis, HIV-1 and Helicobacter pylori.* Nat Immunol, 2009. **10**(10): p. 1081-8.

254. Gringhuis, S.I., et al., *C-type lectin DC-SIGN modulates Toll-like receptor signaling via Raf-1 kinase-dependent acetylation of transcription factor NF-kappaB.* Immunity, 2007. **26**(5): p. 605-16.

255. Gringhuis, S.I., et al., *Fucose-specific DC-SIGN signalling directs T helper cell type-2 responses via IKKepsilon- and CYLD-dependent Bcl3 activation.* Nat Commun, 2014. **5**: p. 3898.

256. Gringhuis, S.I., et al., *Fucose-based PAMPs prime dendritic cells for follicular T helper cell polarization via DC-SIGN-dependent IL-27 production.* Nat Commun, 2014. **5**: p. 5074.

257. Merad, M., F. Ginhoux, and M. Collin, *Origin, homeostasis and function of Langerhans cells and other langerin-expressing dendritic cells.* Nat Rev Immunol, 2008. **8**(12): p. 935-47.

258. Turville, S.G., et al., *Diversity of receptors binding HIV on dendritic cell subsets.* Nat Immunol, 2002. **3**(10): p. 975-83.

259. Valladeau, J., et al., *The monoclonal antibody DCGM4 recognizes Langerin, a protein specific of Langerhans cells, and is rapidly internalized from the cell surface.* Eur J Immunol, 1999. **29**(9): p. 2695-704.

260. Valladeau, J., et al., *Langerin, a novel C-type lectin specific to Langerhans cells, is an endocytic receptor that induces the formation of Birbeck granules.* Immunity, 2000. **12**(1): p. 71-81.

261. Stambach, N.S. and M.E. Taylor, *Characterization of carbohydrate recognition by langerin, a C-type lectin of Langerhans cells.* Glycobiology, 2003. **13**(5): p. 401-10.

262. Hunger, R.E., et al., *Langerhans cells utilize CD1a and langerin to efficiently present nonpeptide antigens to T cells.* J Clin Invest, 2004. **113**(5): p. 701-8.

263. de Witte, L., et al., *Langerin is a natural barrier to HIV-1 transmission by Langerhans cells.* Nat Med, 2007. **13**(3): p. 367-71.

264. Holla, A. and A. Skerra, *Comparative analysis reveals selective recognition of glycans by the dendritic cell receptors DC-SIGN and Langerin.* Protein Eng Des Sel, 2011. **24**(9): p. 659-69.

265. de Jong, M.A., et al., *C-type lectin Langerin is a beta-glucan receptor on human Langerhans cells that recognizes opportunistic and pathogenic fungi.* Mol Immunol, 2010. **47**(6): p. 1216-25.

266. Feinberg, H., et al., *Structural basis for langerin recognition of diverse pathogen and mammalian glycans through a single binding site.* J Mol Biol, 2011. **405**(4): p. 1027-39.

267. Feinberg, H., et al., *Common polymorphisms in human langerin change specificity for glycan ligands.* J Biol Chem, 2013. **288**(52): p. 36762-71.

268. Suzuki, N., et al., *Molecular cloning and expression of cDNA encoding human macrophage C-type lectin. Its unique carbohydrate binding specificity for Tn antigen.* J Immunol, 1996. **156**(1): p. 128-35.

269. van Vliet, S.J., et al., *Carbohydrate profiling reveals a distinctive role for the C-type lectin MGL in the recognition of helminth parasites and tumor antigens by dendritic cells.* Int Immunol, 2005. **17**(5): p. 661-9.

270. van Kooyk, Y., J.M. Ilarregui, and S.J. van Vliet, *Novel insights into the immunomodulatory role of the dendritic cell and macrophage-expressed C-type lectin MGL.* Immunobiology, 2015. **220**(2): p. 185-92.

271. Valladeau, J., et al., *Immature human dendritic cells express asialoglycoprotein receptor isoforms for efficient receptor-mediated endocytosis.* J Immunol, 2001. **167**(10): p. 5767-74.

272. Graham, L.M. and G.D. Brown, *The Dectin-2 family of C-type lectins in immunity and homeostasis.* Cytokine, 2009. **48**(1-2): p. 148-55.

273. Saijo, S. and Y. Iwakura, *Dectin-1 and Dectin-2 in innate immunity against fungi.* Int Immunol, 2011. **23**(8): p. 467-72.

274. Kerscher, B., J.A. Willment, and G.D. Brown, *The Dectin-2 family of C-type lectin-like receptors: an update.* Int Immunol, 2013. **25**(5): p. 271-7.

275. Rizzetto, L., et al., *Systems biology of host-mycobiota interactions: dissecting Dectin-1 and Dectin-2 signalling in immune cells with DC-ATLAS.* Immunobiology, 2013. **218**(11): p. 1428-37.

276. Kanazawa, N., et al., *Molecular cloning of human dectin-2.* J Invest Dermatol, 2004. **122**(6): p. 1522-4.

277. Ariizumi, K., et al., *Cloning of a second dendritic cell-associated C-type lectin (dectin-2) and its alternatively spliced isoforms.* J Biol Chem, 2000. **275**(16): p. 11957-63.

278. Gavino, A.C., et al., *Identification and expression profiling of a human C-type lectin, structurally homologous to mouse dectin-2.* Exp Dermatol, 2005. **14**(4): p. 281-8.

279. Gringhuis, S.I., et al., *Selective C-Rel activation via Malt1 controls anti-fungal T(H)-17 immunity by dectin-1 and dectin-2.* PLoS Pathog, 2011. **7**(1): p. e1001259.

280. Loures, F.V., et al., *Recognition of Aspergillus fumigatus hyphae by human plasmacytoid dendritic cells is mediated by dectin-2 and results in formation of extracellular traps.* PLoS Pathog, 2015. **11**(2): p. e1004643.

281. Brudner, M., et al., *Lectin-dependent enhancement of Ebola virus infection via soluble and transmembrane C-type lectin receptors.* PLoS One, 2013. **8**(4): p. e60838.

282. Powell, L.D. and A. Varki, *I-type lectins.* J Biol Chem, 1995. **270**(24): p. 14243-6.

283. Crocker, P.R., et al., *Siglecs: a family of sialic-acid binding lectins.* Glycobiology, 1998. **8**(2): p. v.

284. Crocker, P.R., J.C. Paulson, and A. Varki, *Siglecs and their roles in the immune system.* Nat Rev Immunol, 2007. **7**(4): p. 255-66.

285. Pillai, S., et al., *Siglecs and immune regulation.* Annu Rev Immunol, 2012. **30**: p. 357-92.

286. Blixt, O., et al., *Sialoside specificity of the siglec family assessed using novel multivalent probes: identification of potent inhibitors of myelin-associated glycoprotein.* J Biol Chem, 2003. **278**(33): p. 31007-19.

287. Campanero-Rhodes, M.A., et al., *Carbohydrate microarrays reveal sulphation as a modulator of siglec binding.* Biochem Biophys Res Commun, 2006. **344**(4): p. 1141-6.

288. Powell, L.D. and A. Varki, *The oligosaccharide binding specificities of CD22 beta, a sialic acid-specific lectin of B cells.* J Biol Chem, 1994. **269**(14): p. 10628-36.

289. Bochner, B.S., et al., *Glycan array screening reveals a candidate ligand for Siglec-8.* J Biol Chem, 2005. **280**(6): p. 4307-12.

290. Cornish, A.L., et al., *Characterization of siglec-5, a novel glycoprotein expressed on myeloid cells related to CD33.* Blood, 1998. **92**(6): p. 2123-32.

291. Bax, M., et al., *Campylobacter jejuni lipooligosaccharides modulate dendritic cell-mediated T cell polarization in a sialic acid linkage-dependent manner.* Infect Immun, 2011. **79**(7): p. 2681-9.

292. Powell, L.D., et al., *Characterization of sialyloligosaccharide binding by recombinant soluble and native cell-associated CD22. Evidence for a minimal structural recognition motif and the potential importance of multisite binding.* J Biol Chem, 1995. **270**(13): p. 7523-32.

293. Zou, Z., et al., *Siglecs facilitate HIV-1 infection of macrophages through adhesion with viral sialic acids.* PLoS One, 2011. **6**(9): p. e24559.

294. Hanasaki, K., L.D. Powell, and A. Varki, *Binding of human plasma sialoglycoproteins by the B cell-specific lectin CD22. Selective recognition of immunoglobulin M and haptoglobin.* J Biol Chem, 1995. **270**(13): p. 7543-50.

295. Thomas, R., L.S. Davis, and P.E. Lipsky, *Isolation and characterization of human peripheral blood dendritic cells.* J Immunol, 1993. **150**(3): p. 821-34.

296. Crocker, P.R. and S. Gordon, *Properties and distribution of a lectin-like hemagglutinin differentially expressed by murine stromal tissue macrophages.* J Exp Med, 1986. **164**(6): p. 1862-75.

297. Crocker, P.R., et al., *Purification and properties of sialoadhesin, a sialic acid-binding receptor of murine tissue macrophages.* EMBO J, 1991. **10**(7): p. 1661-9.

298. Kirchberger, S., et al., *Human rhinoviruses inhibit the accessory function of dendritic cells by inducing sialoadhesin and B7-H1 expression.* J Immunol, 2005. **175**(2): p. 1145-52.

299. van den Berg, T.K., et al., *Cutting edge: CD43 functions as a T cell counterreceptor for the macrophage adhesion receptor sialoadhesin (Siglec-1).* J Immunol, 2001. **166**(6): p. 3637-40.

300. Izquierdo-Useros, N., et al., *Siglec-1 is a novel dendritic cell receptor that mediates HIV-1 trans-infection through recognition of viral membrane gangliosides.* PLoS Biol, 2012. **10**(12): p. e1001448.

301. Nicoll, G., et al., *Identification and characterization of a novel siglec, siglec-7, expressed by human natural killer cells and monocytes.* J Biol Chem, 1999. **274**(48): p. 34089-95.

302. Angata, T. and A. Varki, *Cloning, characterization, and phylogenetic analysis of siglec-9, a new member of the CD33-related group of siglecs. Evidence for co-evolution with sialic acid synthesis pathways.* J Biol Chem, 2000. **275**(29): p. 22127-35.

303. Li, N., et al., *Cloning and characterization of Siglec-10, a novel sialic acid binding member of the Ig superfamily, from human dendritic cells.* J Biol Chem, 2001. **276**(30): p. 28106-12.

304. Lock, K., et al., *Expression of CD33-related siglecs on human mononuclear phagocytes, monocyte-derived dendritic cells and plasmacytoid dendritic cells.* Immunobiology, 2004. **209**(1-2): p. 199-207.

305. Jones, C., M. Virji, and P.R. Crocker, *Recognition of sialylated meningococcal lipopolysaccharide by siglecs expressed on myeloid cells leads to enhanced bacterial uptake.* Mol Microbiol, 2003. **49**(5): p. 1213-25.

306. Carlin, A.F., et al., *Group B Streptococcus suppression of phagocyte functions by protein-mediated engagement of human Siglec-5.* J Exp Med, 2009. **206**(8): p. 1691-9.

307. Kawasaki, N., et al., *Targeted delivery of mycobacterial antigens to human dendritic cells via Siglec-7 induces robust T cell activation.* J Immunol, 2014. **193**(4): p. 1560-6.

308. Ohta, M., et al., *Immunomodulation of monocyte-derived dendritic cells through ligation of tumor-produced mucins to Siglec-9.* Biochem Biophys Res Commun, 2010. **402**(4): p. 663-9.

309. Chen, G.Y., et al., *CD24 and Siglec-10 selectively repress tissue damage-induced immune responses.* Science, 2009. **323**(5922): p. 1722-5.

310. Stephenson, H.N., et al., *Pseudaminic acid on Campylobacter jejuni flagella modulates dendritic cell IL-10 expression via Siglec-10 receptor: a novel flagellin-host interaction.* J Infect Dis, 2014. **210**(9): p. 1487-98.

311. Angata, T. and A. Varki, *Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective.* Chem Rev, 2002. **102**(2): p. 439-69.

312. McEver, R.P., *Adhesive interactions of leukocytes, platelets, and the vessel wall during hemostasis and inflammation.* Thromb Haemost, 2001. **86**(3): p. 746-56.

313. Hong, P., et al., *Human milk oligosaccharides reduce HIV-1-gp120 binding to dendritic cell-specific ICAM3-grabbing non-integrin (DC-SIGN).* Br J Nutr, 2009. **101**(4): p. 482-6.

314. Brinkman-Van der Linden, E.C. and A. Varki, *New aspects of siglec binding specificities, including the significance of fucosylation and of the sialyl-Tn epitope. Sialic acid-binding immunoglobulin superfamily lectins.* J Biol Chem, 2000. **275**(12): p. 8625-32.

315. Koliwer-Brandl, H., et al., *Lectin inhibition assays for the analysis of bioactive milk sialoglycoconjugates.* International Dairy Journal, 2011. **21**(6): p. 413-420.

316. May, A.P., et al., *Crystal structure of the N-terminal domain of sialoadhesin in complex with 3' sialyllactose at 1.85 A resolution.* Mol Cell, 1998. **1**(5): p. 719-28.

317. Hirabayashi, J., et al., *Oligosaccharide specificity of galectins: a search by frontal affinity chromatography.* Biochim Biophys Acta, 2002. **1572**(2-3): p. 232-54.

318. Bachhawat-Sikder, K., C.J. Thomas, and A. Surolia, *Thermodynamic analysis of the binding of galactose and poly-N-acetyllactosamine derivatives to human galectin-3.* FEBS Lett, 2001. **500**(1-2): p. 75-9.

319. Ideo, H., et al., *High-affinity binding of recombinant human galectin-4 to SO(3)(-)-->3Galbeta1-->3GalNAc pyranoside.* Glycobiology, 2002. **12**(3): p. 199-208.

320. Bode, L., *Recent advances on structure, metabolism, and function of human milk oligosaccharides.* J Nutr, 2006. **136**(8): p. 2127-30.

321. Cummings, R.D. and J.M. Pierce, *The challenge and promise of glycomics.* Chem Biol, 2014. **21**(1): p. 1-15.

322. Smith, D.F., X. Song, and R.D. Cummings, *Use of glycan microarrays to explore specificity of glycan-binding proteins.* Methods Enzymol, 2010. **480**: p. 417-44.

323. Rillahan, C.D. and J.C. Paulson, *Glycan microarrays for decoding the glycome.* Annu Rev Biochem, 2011. **80**: p. 797-823.

324. Porter, A., et al., *A motif-based analysis of glycan array data to determine the specificities of glycan-binding proteins.* Glycobiology, 2010. **20**(3): p. 369-80.

Chapter 2: Galectins are Human Milk Glycan Receptors

Alexander J. Noll[1], Jean-Philippe Gourdine[1], Ying Yu[1], Yi Lasanajak[1], David F. Smith[1], and Richard D. Cummings[1]

[1]Department of Biochemistry, School of Medicine, Emory University, Atlanta, GA

AJN performed all experiments listed, with assistance from YY and YL for the glycan microarray experiments. The exception was Supplemental Figure 1, which was performed by JPG. AJN also wrote the manuscript, generated all figures, and analyzed all the data with input from all the other authors.

This is a pre-copyedited, author-produced PDF of an article accepted for publication in *Glycobiology* following peer review. The version of record [Noll AJ et al. *Glycobiology* 2016 Jan 7; pii: cww002.] is available online at:

http://glycob.oxfordjournals.org/content/early/2016/01/07/glycob.cww002.long

Abstract

The biological recognition of human milk glycans (HMGs) is poorly understood. Because HMGs are rich in galactose we explored whether they might interact with human galectins, which bind galactose-containing glycans and are highly expressed in epithelial cells and other cell types. We screened a number of human galectins for their binding to HMGs on a shotgun glycan microarray consisting of 247 HMGs derived from human milk, as well as to a defined HMG microarray. Recombinant human galectins (hGal)-1, -3, -4, -7, -8, and -9 bound selectively to glycans, with each galectin recognizing a relatively unique binding motif; by contrast hGal-2 did not recognize HMGs, but did bind to the human blood group A Type 2 determinants on other microarrays. Unlike other galectins, hGal-7 preferentially bound to glycans expressing a terminal Type 1 (Galβ1-3GlcNAc) sequence, a motif that had eluded detection on non-HMG glycan microarrays. Interactions with HMGs were confirmed in a solution setting by isothermal titration microcalorimetry and hapten inhibition experiments. These results demonstrate that galectins selectively bind to HMGs and suggest the possibility that galectin-HMG interactions may play a role in infant immunity.

Introduction

Human milk provides infants with all essential nutrients, including proteins, lipids, and the digestible carbohydrate lactose [1]. Human milk glycans (HMGs), which contain lactose at their reducing end and are further modified to contain N-acetylglucosamine (GlcNAc), galactose (Gal), fucose (Fuc), and/or sialic acid (as N-acetylneuraminic acid; Neu5Ac), are a major component of human milk [2, 3]. HMGs function as prebiotics that help shape the infant's gut microflora, glycan receptor decoys against pathogenic microbes, regulators of immune responses, and even regulators of gene expression in intestinal epithelial cell cultures as well as other cell types [3-7]. These regulatory functions of HMGs may contribute to the infant health benefits associated with breast-feeding for the first six months of life [1]. Despite the known roles of HMGs in infants, the mechanism(s) by which HMGs regulate immune responses and intestinal epithelial cell gene expression are unknown.

Unlike lactose, HMGs are not appreciably digested in the infant GI tract based on *in vitro* studies [8, 9], although the gut microflora (ex-certain *Bifidobacteria* species [10, 11]) catabolizes HMGs to some degree. This lack of digestion may allow HMGs to act as physiological and/or immunological regulators in the GI tract. A key family of glycan-binding proteins implicated in immune regulation are the galectins, which are expressed by gut epithelial cells and are known for binding to galactose-rich glycans [12, 13]. Thus, we explored whether HMGs may interact in a selective manner with specific galectins.

The Human Protein Atlas project (http://www.proteinatlas.org) [14] and other studies [15-17] have shown that hGal-2, -3, -4, -7, -8, and -9, but interestingly not hGal-1,

are all expressed in epithelial cells of the esophagus, stomach, duodenum, small intestine, and/or large intestine under normal conditions, with some extracellular localization. Recent studies also show that a small percentage (~1%) of HMGs enters the infant's circulation and urine [18-20]. Thus, HMGs in either the GI tract or blood have the potential to contact galectins *in vivo*, which may modulate their activity and functions in breast-fed infants.

To explore these interactions of galectins with HMGs, we have exploited the availability of a human milk shotgun glycan microarray containing natural glycans purified from human milk, termed the HM-SGM-v2 array, [21, 22], as well as an array containing defined, simple HMG structures, and the extensive non-HMG glycan microarray from the Consortium for Functional Glycomics. Studies of galectin binding to glycans both on microarrays and free in solution demonstrate that human galectins, except for hGal-2, bind a unique subset of human milk glycans. This is the first systematic study of the binding of a lectin family to a specific metaglycome [23]. The results of this study suggest that galectin-HMG interactions might be relevant to infant health.

Results

Binding of Human Galectins to the HM-SGM-v2 Glycan Microarray and the Array from the Consortium for Functional Glycomics

Recombinant hGal-1 (C2S mutant, refer to Materials and methods for more information), hGal-2, hGal-3, hGal-4, hGal-7, hGal-8, and hGal-9 were screened on the HM-SGM-v2 array at three concentrations: 2, 20, and 200μg/ml. The results for one concentration of each galectin are shown in Figure 1 (refer to Supplementary File 1 for the results for all three concentrations of each galectin as well as the measured values). The results showed that all the galectins tested, with the exception of hGal-2, bound to glycans on the HM-SGM-v2 array. Most glycans that were strongly bound by galectins demonstrated binding in a dose-dependent manner (Supplementary File 1). Non-specific binding was minimal since binding of the galectins to the defined structure 3-fucosyllactose, a known non-binder of most galectins [24], was only at background levels. Moreover, each galectin had a more or less unique binding profile on the HM-SGM-v2 array, suggesting that each galectin appeared to recognize a structural motif within the collection of HMGs. However, there were some general similarities among galectins in binding. For example, only neutral HMG samples were bound, while the sialylated HMGs were typically not bound; the few that were bound likely have a nonsialylated branch.

The unexpected lack of hGal-2 binding to HMGs was not due to inactivity of the hGal-2 preparation; we concurrently screened the hGal-2 on the glycan microarray from the Consortium for Functional Glycomics (CFG), which contains 610 defined glycan structures, and found that it bound to several glycans (Figure 2 and Supplementary File

2). On the CFG array, hGal-2 at both 20μg/ml and 200μg/ml bound strongly to five glycans (Figure 2a & 2b); no significant binding was detected at 2μg/ml hGal-2. Manual inspection and Glycopattern analysis [25] showed that the binding motif for hGal-2 on the CFG array was the Blood Group A Type 2 determinant (GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ-, Figure 2c), a determinant that is mainly restricted to expression on erythrocytes and epithelial tissues and not known to occur on HMGs. Indeed, we did not observe any binding of a Blood Group A-specific antibody on the HM-SGM-v2 array, further confirming the absence of Blood Group A determinants on HMGs. Therefore, the lack of hGal-2 binding to the HM-SGM-v2 was due to the absence of high affinity hGal-2 determinants on HMGs.

For the remaining galectins, the glycans bound on the HM-SGM-v2 array were manually examined. Using previous glycan sequencing data for these samples [21, 22], binding motifs were defined for each galectin and found to be relatively unique (Table I). hGal-1 binding to the HM-SGM-v2 array was weak and broad, with a slight preference for branched glycans terminating in Type 1 LacNAc (Galβ1-3GlcNAc) or Type 2 LacNAc (Galβ1-4GlcNAc), although linear structures were also bound. hGal-3 only bound glycans containing at least three repeating Type 2 LacNAc/lactose structures that lacked branched features, which is consistent with previous studies [26]. hGal-8 had a similar preference for structures containing at least three linear repeating LacNAc structures without branching, similarly to hGal-3, although the actual specificity of hGal-8 was somewhat different as can be seen in Figure 1, including the weak binding of hGal-8 to sialylated glycans. hGal-9 bound only a relatively restricted panel of glycans unlike the other galectins. The structure of the major glycan (HMO-35) bound by hGal-9 is

predicted to be a nonfucosylated, biantennary, neutral HMG structure containing terminating Type 1 LacNAc and Type 2 LacNAc determinantss based on lectin and antibody screening profiles [21]. Due to limitations in sample material, the actual structures of HMO-35 and the sialylated glycans bound by hGal-8 were not determined by MS^n at this point in time. hGal-4 gave a broad pattern of binding, with LNFPI (Fucα1-2Galβ1-3GlcNAcβ1-3Galβ1-4Glc) and samples containing LNFPI-like determinants being the major structures bound, although nonfucosylated HMGs including LNT (Galβ1-3GlcNAcβ1-3Galβ1-4Glc) and LNnT (Galβ1-4GlcNAcβ1-3Galβ1-4Glc) were also recognized.

The binding pattern of hGal-7 was interesting as this galectin bound many glycans containing at least one terminal Type 1 LacNAc determinant. Branched glycans containing one or more terminal Type 1 LacNAc determinants were generally slightly preferred over non-branched structures, similarly to hGal-1. The presence of α1-2 fucosylation did not seem to increase or decrease binding to this determinant. Glycans have a non-reducing Type 2 LacNAc-terminating sequence were typically bound much more weakly, especially linear Type 2 LacNAc-terminating glycans. This binding motif was interesting when compared to hGal-7 binding motif on the CFG Array; a comparison of glycans bound on the HM-SGM-v2 and CFG arrays by hGal-7 is shown in Table II. Screening on the CFG Array showed that hGal-7 was relatively specific for glycans containing Blood Group H Type 2 (Blood Group H2) determinants and also expressing two Type 2 LacNAc units (Fucα1-2Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ-) (Supplementary File 2). Type 1 LacNAc-terminating structures are very limited on the CFG array, unlike the HM-SGM-v2 array where this determinant is very abundant based

on screenings with an antibody specific for terminal Type 1 LacNAc [21]. Therefore, the HM-SGM-v2 array helped to further refine the glycan specificity for hGal-7, because such branched Type 1 LacNAc determinant are not present on the CFG microarray.

Binding of Galectins to a Defined HMG Microarray

To complement the HM-SGM-v2 shotgun microarray, a second microarray, termed the "defined HMG microarray", was generated that contained chemically defined HMG structures (as opposed to HMG fractions purified from human milk) that are commercially available. Additionally, galactooligosaccharides (GOS), a prebiotic oligosaccharide mixture that has been proposed as an HMG alternative for infant formula supplementation [27], was also included on the defined HMG microarray both before and after fractionation to have semi-purified GOS fractions. Galectins were then screened on this array at three concentrations in the presence or absence of lactose, a specific inhibitor of galectins. Figure 3 shows the results of the screening of galectins on the defined HMG array at one concentration with and without lactose (see Supplementary File 3 for data from all screenings). As seen on the HM-SGM-v2 array, all galectins except hGal-2 bound glycans on the defined HMG microarray. Binding was typically observed with LNT, LNnT, LNFPI, and 2'-FL but not 6'-sialyllactose (6'-SL), 3-FL, or any of the GOS samples. Lactose was poorly bound, and 3'-sialyllactose (3'-SL) was only bound (albeit weakly) by hGal-1 and hGal-8. hGal-4, -7, and -9 bound well at relatively low concentrations to glycans on the defined HMG array, while hGal-1, -3 and -8 required much higher protein concentrations for detectable binding. Binding of hGal-8 binding was weak. The HMGs bound by the galectins were largely bound in a dose-dependent

manner (Supplementary File 3). Co-incubation of galectins with 0.1M lactose during screening greatly reduced galectin binding (Figure 3 and Supplementary File 3), indicating that binding required carbohydrate recognition.

Influence of the Reducing End of HMGs on Binding to Galectins

Due to the relatively small mass of the HMGs on the defined HMG array (mostly 2-5 monosaccharides), we hypothesized that the reducing end glucose of these small glycans might contribute to binding by galectins. The glycans on the HM-SGM-v2 were derivatized with the bifunctional linker 2-amino-N-(2-aminoethyl)-benzamide (AEAB) [28] by reductive amination, which converts the reducing end glucose into a sugar alcohol, an "open-ring" structure. We considered that glucose might possibly be part of the binding determinant, in which case the "open-ring" glucose may reduce or even eliminate binding. Thus, the HMGs on the defined HMG array were also derivatized with AEAB in a manner that maintained the reducing end glucose in a cyclic ("closed-ring") conformation [29] (α/β mixture) and simultaneously printed. Comparison of the corresponding "open-ring" and "closed-ring" glycan derivatives is highlighted in Table III, which shows a clear preference for most galectins to bind to the "closed-ring" conformation of HMGs. In fact, this "closed-ring" reducing end was required for hGal-9 binding on the defined HMG array; hGal-9 bound LNT, LNnT, and LNFPI, but only when the reducing end glucose ring was intact. Another dramatic example is hGal-4 and hGal-7 binding to 2'-FL, which was almost completely dependent on the reducing end glucose ring being intact. Therefore, these results suggest that the HM-SGM-v2 average rank data should be interpreted with caution because the "open-ring" glucose structures

may bias the results to longer HMG structures, where the reducing end glucose is no longer a part of the galectin binding determinant. A caveat to this interpretation is that the derivatization method maintaining the reducing end glucose in a "closed-ring" conformation also introduces an additional acryloyl group not found in the reductively aminated structure, and reducing end linkers have been found to directly participate in some galectin binding [30]. However, while this linker may partially increase affinity, the loss of part of the binding determinant should produce a much more dramatic effect. Additionally, it has been noted by others that reductively aminated lactose greatly reduces affinity [26]. Thus, the major explanation for the preference for "closed-ring" vs. "open-ring" HMG binding is most likely the reducing end glucose conformation, not the longer linker in the "closed-ring" glucose.

Galectin binding to Free, Underivatized HMGs

While the glycan microarray results demonstrate galectin binding to HMGs, this solid-phase presentation of HMGs is not the "natural" form of HMGs regardless of the linker strategy or reducing status. HMGs are unique because they naturally exist as free, reducing glycans in solution. To ensure that the results seen for galectin-HMG interactions also applied to the more natural solution-based setting, two experimental approaches were taken. In one approach we examined the ability of free, underivatized HMGs to inhibit galectin binding to the defined HMG array, and in the second approach we measured the binding of galectins to free, underivatized HMGs by isothermal titration microcalorimetry (ITC). hGal-7 and hGal-4 were used for the hapten inhibition experiments as the model prototypical and tandem-repeat galectins, respectively, due to

their relatively unique glycan specificities versus other galectins and robust binding to the HMG microarrays. hGal-7 was used as the model galectin for the ITC studies because hGal-7 only contains a single carbohydrate-binding site and no ITC data exists for hGal-7 with HMGs.

For the free HMG inhibition studies, 20μg/ml of hGal-4 or hGal-7 were used, which was determined to be the approximate apparent K_d for hGal-4 and hGal-7 binding to most of the major bound glycans on the defined HMG array by screening multiple galectin concentrations on the defined HMG microarray. hGal-4 or hGal-7 was preincubated with 0.05, 0.5, or 5mM of free HMGs prior to screening on the defined HMG array. An example of these results is shown in Figure 4 (see Supplementary File 4 for the total results for hGal-4 and hGal-7). The results show that LNT, LNnT, 2'-FL, and LNFPI, but not 3-FL, inhibited hGal-4 and hGal-7 binding to the defined HMG array in a dose-dependent manner; these results are consistent with the binding of these galectins to the defined HMG array where LNT, LNnT, LNFPI, and 2'-FL but not 3-FL were bound. For both hGal-4 and hGal-7, the free HMGs caused little or no inhibition at 50μM, about 50% inhibition at 500μM, and >95% inhibition at 5mM. These percent inhibition values mirror the measured K_d of hGal-7 for these HMGs by ITC (see below) and those reported by surface plasmon resonance for hGal-4 [31]. Additionally, these results show that the free HMGs not only inhibited binding to the same HMG structure printed on the array but also different structures; indeed, all bound HMGs on the defined HMG array were inhibited by a single free HMG. Therefore, these experiments indicate that hGal-4 and hGal-7 bind these HMG structures in solution and that binding to galectins is specific (i.e. via the lactose-binding CRD) since the inhibition profiles looked similar to galectins

screened on the defined HMG array in the presence of the specific inhibitor 0.1M lactose. These results can likely be extended to other galectins as well.

Thermodynamics of Galectin Binding to Free, Underivatized HMGs by Isothermal Titration Microcalorimetry (ITC)

To further corroborate solution binding of galectins to HMGs, hGal-7 was tested with the HMGs LNT, LNnT, LNFPI, 2'-FL, and 3-FL by ITC. Lactose was also included in this experiment as a reference because ITC data for human hGal-7 with lactose has been published [32]. Due to the relatively low affinity of hGal-7 for these HMGs (>50μM) in pilot ITC experiments, the "low c-value" method [33] was used to perform ITC, with n fixed to 1.00 based on previous knowledge [32]. The ITC results are shown in Figure 5 and the measured thermodynamic parameters (with associated uncertainties) are presented in Table IV. The results show that all tested glycans except 3-FL showed measurable binding to hGal-7, as seen in the microarrays and free HMG inhibition experiments. The hGal-7-lactose parameters measured at 298K ($K_a = 2.880\times10^3$ M^1 and ΔH = -10.7kcal/mol) were highly similar to the previous ITC data that did not use the low c-value method ($K_a = 2.2\times10^3$ M^{-1} and ΔH = -10.6kcal/mol at 300K). This finding was important because it not only validated the use of "low c-value" ITC in our hands, but also demonstrated that the GST fusion tag on the recombinant hGal-7 protein used in this study did not significantly affect the solution-based binding studies. An upward trend for 3-FL heat generation was seen in the thermogram and curve, suggesting that hGal-7 may still bind 3-FL but only at very high concentrations (>1mM), which is significantly higher than the K_d for lactose and 2'-FL. Thus, fucosylation of the 3 –OH group of

glucose is disruptive to hGal-7 binding, a feature common to most galectins due to the requirement of a free 3 –OH group on Glc/GlcNAc for binding to most galectins [26, 34]. α1-2 fucosylation of Type 1 LacNAc or Type 2 LacNAc, however, did not affect binding (comparing the results of lactose to 2'-FL and LNT to LNFPI), although a trend towards decreased enthalpic favorability and increased entropic favorability was seen for the α1-2 fucosylated HMGs. Additionally, only a 1.5-fold difference was seen in binding to the Type 1 LacNAc-terminating LNT versus Type 2 LacNAc-terminating LNnT structures, which is most likely insignificant from a receptor-ligand interaction standpoint, in contrast to the results seen by glycan microarray studies (discussed below). Therefore, the ITC results further confirmed solution binding of galectins to HMGs and also provided previously untested thermodynamic data for hGal-7 binding to glycans, which has helped to better define the glycan specificity of hGal-7.

Absence of Detectable Galectins in Human Milk

We also tested whether human milk itself might contain galectins. For this, we utilized dialyzed, defatted human milk, recombinant galectins as standards, and defined rabbit anti-sera to the galectins. In Western blot analyses we did not detect any galectins in human milk (Supplementary Figure 1), although standard galectins were easily detectable. Using recombinant galectins as standards at different amounts we established that ~5ng per 300 μg milk protein loaded onto the gels was the limit of sensitivity by this approach.

Discussion

A major finding in our study is that all but one of the human galectins tested interact with specific HMGs at their physiologically relevant concentrations. While such interactions have been predicted to occur [35], this represents the first systematic study to directly test interactions of human galectins with a large variety of HMGs other than a few relatively simple glycans, e.g. lactose, LNnT, 2'-FL [24, 26, 30, 31, 36]. The results of our study extend these earlier observations and also identify more complex HMGs as additional targets of specific galectins.

While lactose is present at sufficiently high concentrations in human milk (~0.2M) to inhibit galectin activity [37], lactose is utilized as a carbohydrate source by the infant and is thus metabolized in the proximal small intestine of the infant by lactase. With the possible exception of newborns and some preterm infants who may not quantitatively digest lactose [38, 39], it is predicted that lactose would only be an efficient galectin binder in the upper GI tract prior to reaching the small intestine. By contrast, *in vitro* studies suggest that HMGs are not significantly digested by the conditions and human digestive enzymes of the GI tract [8, 9], although digestion can occur by the colonic microflora. Therefore, HMGs may be relatively intact within most of the infant GI tract.

Previous studies have demonstrated galectin expression by human GI tract epithelial cells [15-17], indicating that galectins are expressed in anatomical regions that may come in contact with HMGs. It is unknown, however, whether galectins are properly positioned for contact with HMGs in all cases. For example, some galectins may lack extracellular localization or may only be expressed in deeper levels of the tissue. In

addition, whether or not the galectin expression and localization in newborn, infant, and toddler tissues mimics that of adult human tissues is unclear. Nonetheless, the current literature on GI tract cell/tissue expression along with membrane localization and secretion of galectins suggest that galectins are likely to be exposed to HMGs and could directly interact physiologically. For example, the K_d of hGal-7 (and other galectins) for simple HMGs is in the high micromolar range, which is below or near the concentration of ~0.5-5mM of these simple HMGs (2'-FL, LNT, LNnT, and LNFPI) in human milk [40]. On the other hand, the ~1% of HMGs that enters the circulation is likely not at a sufficiently high enough concentration to bind HMGs [18-20]. Clearly, more studies are need in the future to explore the positioning and exposure of specific human galectins in the infant gut and the potential of physiological interactions there between HMGs and galectins. As certain galectin-glycan interactions have been found to promote beneficial health effects such as gut homeostasis and oral tolerance [41], the importance of galectin-HMG interactions in GI tract physiology will be addressed in future studies.

In addition to HMG binding by galectins, a striking result of our study was the unique binding signature of each galectin on the HMG arrays. This result is consistent with the data observed with galectins screened on the CFG glycan microarray (microarray data available from http://www.functionalglycomics.org/fg/; see also [42, 43]) and other shotgun glycan microarrays [28]. However, this finding was important because this was the first study to reveal differential glycan specificity of galectins for a human metaglycome [23], in this case the human milk metaglycome. These and previous results demonstrate that human galectins have relatively unique glycan specificities, including in the context of a natural metaglycome, despite the ability of all of these

galectins to bind, albeit typically with lower affinity, to lactose and LacNAc. This glycan specificity may relate to the fact that each galectin has more or less unique physiological activities [44-48]. In other words, most human galectins do not appear to be redundant in their activities, which may be at least partially explained by their non-redundant glycan specificities. HMGs might be superior ligands for galectins compared to other oligosaccharides such as galactooligosaccharides (GOS) and fructooligosaccharides (FOS), which are proposed infant formula additives as an "HMG substitute" particularly because of their prebiotic properties [27]. Interestingly, we found no binding of galectins to GOS on the defined HMG array and thus such components are unlikely to modulate galectin activities.

For hGal-2, no binding was seen on the HM-SGM-v2 or defined HMG microarrays, even though binding to the CFG Glycan Microarray was observed. Although the lack of binding to the HM-SGM-v2 may be due to the absence of high affinity ligands such as the Type 2 Blood Group A determinant, this result was still surprising because rat galectin-2 was previously shown to bind the simple HMG structures including LNnT, LNT, and LNFPI, with micromolar affinity (K_d's of 130, 68, and 23μM, respectively) by frontal affinity chromatography with reductively aminated (and thus "open-ring") glycans [26]. Thus, human and rat galectin-2 may differ in glycan-binding specificity, but in any case these results raise the question as to the physiological relevance of hGal-2 interactions with HMGs *in vivo*.

Compared to previous studies on HMG-GBP interactions, our study used a rather unique high-throughput approach to explore binding to glycans on multiple microarrays. A key new technology to make this study possible was the development of the HMG

shotgun microarray [21, 49]. We could identify a novel branched Type 1 LacNAc binding motif for hGal-7 that was not seen on the CFG microarray because, with the exception of simple HMGs, such HMG structures are not present on the CFG glycan microarray. On the other hand, the Type 2 Blood Group H (H2) antigen with at least two LacNAc motifs (i.e.-Fucα1-2Galβ1-4GlcNAcGalβ1-4GlcNAcβ-) was the only motif recognized on the CFG glycan microarray, although binding to other glycan structures, including non-HMG structures with terminal Type 1 LacNAc, could be detected when hGal-7 was screened at a much higher concentration of 200μg/ml (Supplementary File 3). This was an interesting result because, while α1-2 fucosylation of Type 2 LacNAc-terminating glycans greatly improved binding, α1-2 fucosylation of Type 1 LacNAc-terminating glycans and glycans with only one lactose or Type 2 LacNAc repeat had no effect on binding by ITC (Table IV). The mechanism of how α1-2 fucosylation only improves binding to a specific subset of glycans is currently unclear, but this will be addressed in future studies examining the thermodynamics of binding along with co-crystallization of hGal-7 with defined glycan structures with or without α1-2 fucosylation.

Based on the ITC data, there was a trend towards decreased enthalpic favorability and increased entropic favorability for binding to LNT and LNnT (Type 1 LacNAc- and Type 2 LacNAc-terminating structures, respectively), suggesting slightly different mechanisms for hGal-7 binding to Type 1 LacNAc- and Type 2 LacNAc-terminating structures. However, the K_d values for LNT and LNnT binding were <2-fold different. This result was in contrast to the defined HMG array studies, where LNT and LNFPI were consistently bound 3-5-fold better than LNnT. This suggests that hGal-7 may have

greater avidity for Type 1 LacNAc vs. Type 2 LacNAc structures, which further suggests a different mechanism of Type 1 LacNAc vs. Type 2 LacNAc binding despite similar affinity. While likely not important from an HMG binding standpoint, this concept may be important in hGal-7 cell surface glycan interactions and warrants further study. Previous ITC studies with hGal-1 and hGal-3 [32, 36] show that these two galectins can also bind some simple HMGs. Due to the complexity of tandem-repeat galectins (two non-identical CRDs), these were not tested by ITC with HMGs. Additionally, the contribution of the individual CRDs to HMG binding by tandem-repeat galectins was of interest but beyond the scope of this study. Current efforts are underway in the laboratory to measure the affinities of tandem-repeat galectins to HMGs as well as the glycan-binding specificity, affinity, and binding mechanism of individual CRDs towards HMGs.

The binding of HMGs to galectins raises another interesting point: are galectins themselves present in human milk? Galectins are present in most human tissues, but no studies have explored their presence in human milk. Using a Western blot approach, we did not detect galectins in human milk (<5ng per 300μg milk protein, or <0.002% of total milk protein by mass; Supplementary Figure 1). This also result corroborates previous proteomics studies that did not identify galectins in human milk, although hGal-7 was detected but only at trace levels in one of these studies [50, 51]. Therefore, despite their antimicrobial properties [43, 52, 53] and other activities that might otherwise be thought to be beneficial to infants, human milk lacks significant quantities of galectins. Such a result might suggest that endogenous galectins in the infant gut might be privileged to interact with HMGs in those locations. An additional issue to be considered in the future is the degree to which HMGs or endogenous galectins in the GI tract might alter the

microbiota independently of the HMGs' prebiotic functions. Finally, the expression/localization of galectins and/or galectin glycan ligands in the neonatal and infant GI epithelium are also of interest to determine galectin expression and localization. In any case, our results suggest that differential interactions of HMGs with human galectins might impact infant health and immune development.

Materials and Methods

Recombinant Human Galectin Expression and Purification

The recombinant human galectins used in this study were hGal-1, -2, -3, -4, -7, -8, and -9. The hGal-1 used had a C2S substitution; this substitution greatly improves stability but does not alter affinity for lactose [54]. Recombinant hGal-3 cloning was previously described [55]. Recombinant hGal-9 protein was purchased from R&D. The hGal-2, -4, and -7 CDS were PCR-amplified from human genomic DNA, and hGal-8 short isoform CDS were generated by gene synthesis with OptimumGene™ Codon Optimization (Genescript) based on the NCBI reference sequences. Recombinant hGal-1 C2S was generated by PCR-based mutagenesis of wild-type hGal-1 CDS and cloned into pQE-50 at the BamHI and HindIII cut sites. The hGal-2 CDS was cloned into pET11b (Novagen) at the NdeI and BamHI cut sites, hGal-4 CDS into a modified pET29a vector (Novagen), hGal-7 CDS into pGEX-4T-1 (GE Healthcare) at the BamHI and XhoI sites, and hGal-8 CDS into pET22b (Novagen) at the NdeI and HindIII cut sites. The galectins were expressed from *E. coli* M15 (Qiagen), BL21(DE3) (Life Technologies), or BL21 star (DE3) (Life Technologies). hGal-1, -2, -3, -4 and -8 were expressed as untagged, native proteins. hGal-7 was expressed as a GST-tagged fusion protein, and the GST tag was not removed prior to glycan microarray screening. Sanger sequencing (Genewiz) was used to confirm the nucleotide sequence of each galectin.

Galectins except hGal-9 were expressed for 4 hours after induction with 0.1-1mM IPTG (USB) when the OD_{600} of cultures was 0.5-0.7 with a UV-1700 UV-Vis Spectrophotometer (Shimadzu). The cells were then pelleted by centrifugation and frozen at -20^{o}C overnight. Cell pellets were lysed with CellLytic B Buffer (Sigma) containing

14mM β-mercaptoethanol (Fisher Scientific), 1mM lysozyme (Sigma), 10U/ml Benzonase Nuclease (Novagen), and Complete, EDTA-free protease inhibitor cocktail (Roche). For hGal-2, CellLytic B was replaced with PBS (6.7mM KH_2PO_4 pH 7.5, 150mM NaCl) and sonication was performed. Cell debris was removed by centrifugation at 15,000×g for 30min at 4°C. The supernatants were applied to columns containing 10-25ml lactosyl-Sepharose gel (prepared as previously described [56]) that was equilibrated with PBS containing 14mM β-mercaptoethanol. The columns were washed with PBS + 14mM β-mercaptoethanol and then eluted with lactose elution buffer (PBS containing 14mM β-mercaptoethanol and 0.1M lactose (Fisher Scientific)). Elution fractions positive for protein by absorbance at 280nm were pooled, and aliquots were stored at -80°C in the lactose elution buffer until immediately before use. SDS-PAGE, Coomassie Brilliant Blue R-250 staining, and densitometry analysis using the Gel Analysis feature in ImageJ (http://imagej.nih.gov/ij/) were used to confirm that all galectin preparations were >90% pure.

Generation and Printing of on Human Milk Shotgun Glycan Microarray and Defined HMG Microarray

The Human Milk Shotgun Glycan Microarray Version 2 (HM-SGM-v2) was generated and printed as previously described [21]. The HMGs and other glycans used for generating the defined HMG microarray were all purchased from V-Labs, except galactooligosaccharides (GOS), which was a gift from Abbott. All structures (except GOS, which is a mixture) are shown in Supplementary Table I. Each HMG was derivatized with 2-ethyl-N-(aminoethyl)-benzamide (AEAB) [28] by reductive

amination, or with a procedure that maintains the reducing end ring structure, as previously described [57]. GOS was further fractionated into six fractions (F1-F6) with a Shimadzu CBM-20A HPLC system using a Zorbax NH_2 normal-phase column, which were detected by absorbance at 330nm with an SPD-20A UV detector. All the AEAB-labeled glycans shown in Supplemental Table I, as well as the crude GOS mixture and six chromatography fractions, were printed on N-hydroxylsuccinamide (NHS)-activated slides (Schott) as previously described [58].

Screening of Galectins on Human Milk Glycan-260 Shotgun Microarray and Defined HMG Microarray

Prior to screening, all galectins except hGal-9 were biotinylated with EZ-Link Sulfo-NHS-LC-Biotin (Thermo Scientific) according to manufacturer's instructions. Biotinylation was performed in lactose elution buffer as the solvent. Excess biotinylation reagent, free NHS, and lactose were removed by passing the biotinylated galectins over a PD-10 desalting column (GE Healthcare) according to manufacturer's instructions, and eluted with PBS. This biotinylation procedure has been found to retain >95% galectin activity as measured by lactosyl-Sepharose binding and not compromise binding specificity when compared to antibody-based detection (refer to the CFG website, http://www.functionalglycomics.org/fg/, for galectin glycan microarray data with anti-galectin antibody detection). Galectins were quantitated by measuring the absorbance at 280nm and comparing to the theoretical molar absorptivity of each galectin (calculated using the ExPASY Protparam tool, http://web.expasy.org/protparam/). β-mercaptoethanol at ~14mM final was then added to each biotinylated galectin preparation.

For glycan microarray screening, the biotinylated galectins and full-length hGal-9 were diluted in TSM binding buffer (20mM Tris-HCl pH 7.4, 150mM NaCl, 2mM $CaCl_2$, 2mM $MgCl_2$, 0.05% v/v Tween-20, and 1% w/v BSA) containing 14mM β-mercaptoethanol. The galectins were screened on the human milk glycan microarrays and, for hGal-2 and -7, the CFG glycan microarray as previously described [49, 58]. All biotinylated galectins were screened at 2, 20, and 200μg/ml and detected with Cy5-labeled streptavidin (Molecular Probes) at 0.5μg/ml on both the defined HMG and HM-SGM-v2 microarrays. Biotinylated hGal-2 was also similarly screened on the Consortium for Functional Glycomics (CFG) Glycan Microarray Version 5.1, but was only screened at 200μg/ml on the HM-SGM-v2 and defined HMG arrays. hGal-9 was screened at 0.2 and 2μg/ml and detected with goat anti-human galectin-9 affinity purified polyclonal antibody (R&D) at a final concentration of 20μg/ml, followed by incubation with Alexa Fluor 488-labeled rabbit anti-goat IgG (Molecular Probes) at 5μg/ml. hGal-7 was screened on the CFG Glycan Microarray at 0.5, 2.0, 5.0, and 200μg/ml; only the first three concentrations were used for rank analysis due to detector saturation by the high affinity binders at 200μg/ml. As controls for specificity, the 200ug/ml concentration of each galectin (and sometimes the 2 and 20μg/ml concentrations) were screened in the presence of 0.1M lactose in TSM binding buffer, but only on the defined HMG microarray since HM-SGM-v2 arrays were highly limited. For hGal-9, only the 2μg/ml concentration was screened on the defined HMG microarray in the presence or absence of 0.1M lactose.

Rank analysis was performed as previously described [59]. Briefly, the glycan structure with the highest RFU at a given concentration was ranked 100 and all other

RFU's were normalized to this value. The average rank was calculated as the average of the ranks at all concentrations screened for a given structure. Rank analysis of glycan binding was performed using all three concentrations of galectins screened, if possible. However, the lowest concentration(s) tended to show weak or no binding and thus were excluded from the rank analysis. For this reason, the 2μg/ml hGal-3 and 2μg/ml hGal-8 screens on the HM-SGM-v2 array were excluded from rank analysis. For the defined HMG microarray, only the 200μg/ml concentration was used for rank analysis of hGal-3. For hGal-2 and -7 screening on the CFG Glycan microarray, rank analysis as well as analysis of high binding structures and motifs was performed using Glycopattern (https://glycopattern.emory.edu/) [25]. For the HM-SGM-v2 results, all the known structures referred to in this manuscript have been previously sequenced by MS^n [21, 22].

Free HMG Inhibition of Defined HMG Microarray binding by Galectins

The HMGs 2'-fucosyllactose (2'-FL), 3-fucosyllactose (3-FL), lacto-N-tetraose (LNT), lacto-N-neotetraose (LNnT), and lacto-N-fucopentaose I (LNFPI) were purchased from V-Labs. hGal-4 and hGal-7 at 1-4mg/ml was biotinylated as described above. The biotinylated hGal-4 or hGal-7 was then desalted on a Bio-Gel P10 column to quantitatively remove lactose; the absence of detectable lactose in the desalted galectin preparations was confirmed by phenol-sulfuric acid assay analysis. The biotinylated hGal-4 or hGal-7 at 20μg/ml (0.55μM or 0.49μM, respectively) was preincubated with 50μM, 500μM, or 5mM of 2'-FL, 3-FL, LNT, LNnT, or LNFPI (V-Labs) in TSM binding buffer containing 14mM β-mercaptoethanol for 30 minutes prior to glycan microarray screening. In the case of hGal-7, 3-FL was only used at 5mM. 20μg/ml hGal-

4 or hGal-7 preincubated with TSM binding buffer containing 14mM β-mercaptoethanol was used as the mock control. Each sample was then incubated on a separate subarray on a single slide containing 14 arrays. In the case of hGal-4, two slides had to be used; a mock control was included on both slides as the reference control. Glycan microarray screening was otherwise performed as described above, except that the slides were scanned and analyzed using a GenePix 4300A scanner and GenePix Pro 7 software (Molecular Devices).

Isothermal Titration Microcalorimetry (ITC)

The HMGs 2'-fucosyllactose (2'-FL), 3-fucosyllactose (3-FL), lacto-N-tetraose (LNT), lacto-N-neotetraose (LNnT), and lacto-N-fucopentaose I (LNFPI) were purchased from V-Labs. Lactose was purchased from Fisher Scientific. All these HMGs were dissolved in PBS (6.7mM KH_2PO_4 pH 7.4, 150mM NaCl) and accurately quantitated by the phenol-sulfuric acid assay as previously described [60] with slight modifications; a monosaccharide mixture that best represented each HMG's monosaccharide composition was used as the standard. The measured concentrations of each HMG were: 8.6mM lactose, 9.1mM 2'-FL, 4.89mM 3-FL, 4.28mM LNT, 4.11mM LNFPI, and 4.6 & 9.1mM LNnT.

Prior to beginning the ITC, hGal-7 was passed over a Bio-Gel P10 column to quantitatively remove lactose; the absence of detectable lactose in the desalted galectin preparations was confirmed by phenol-sulfuric acid assay analysis. PBS + 14mM β-mercaptoethanol was the final buffer. The hGal-7 concentration was then measured by the absorbance at 280nm using the theoretical molar absorptivity calculated using the

ExPASY Protparam tool, http://web.expasy.org/protparam/). hGal-7 was measured to be 27.9μM in the first experiment and 28.1μM in the second experiment.

ITC was performed using a MicroCal auto-iTC$_{200}$ instrument (GE Healthcare). An initial water-water titration was performed to ensure background titration heats were negligible and the noise was low. Four galectin-glycan titrations were performed in a given experiment; lactose and LNnT were included in both experiments to examine inter-experimental variability, which was determined to be minimal. LNnT was also used at two different concentrations in these two experiments (4.6mM and 9.1mM); no significant differences in the curve-fitting parameters were observed. Additionally, the corresponding buffer-glycan titrations were also performed and subtracted from the corresponding galectin-glycan titration data. PBS + 14mM β-mercaptoethanol eluted from the Bio-Gel P10 column (prior to sample application but after column washing) was used for the buffer-glycan titrations. The data was analyzed with Origin 7 software with manual adjustment of the integration ranges and baseline when deemed appropriate due to the relatively low heats generated. The data was fit to a One Site model with n fixed at 1.00; initial parameters for K_a and ΔH ranged from 1000-10000 and 1000-20000, respectively. One independent experiment of hGal-7 with LNT, LNFPI, 2'-FL, and 3-FL was performed, with the reported uncertainties representing the Origin 7-calculated curve-fit uncertainties. Two independent experiments of hGal-7 with lactose and LNnT were performed (although LNnT was used at two different concentrations, but this did not significantly alter the results), with the reported uncertainties representing the standard error of the mean (SEM).

Acknowledgements

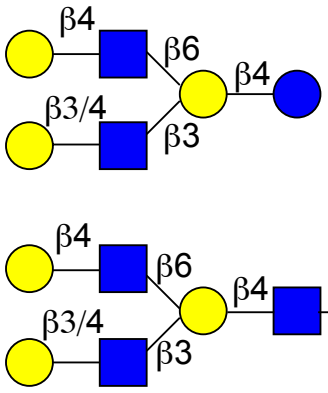
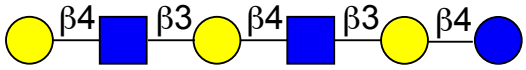
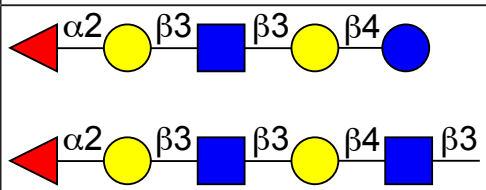
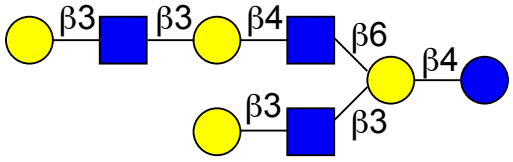
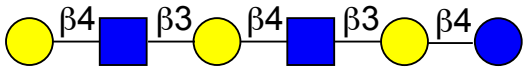
This work was supported by the National Institutes of Health [P41GM103694, R24GM09879 to R.D.C.] and Abbott Laboratories [to D.F.S. and R.D.C.]. We thank Melinda Hanes for assistance with and training for the ITC experiments, Jamie Heimburg-Molinaro for critical review of the manuscript, Sandra Cummings and Hong Ju for technical assistance, and Oskar Laur at the Emory Custom Cloning Core (ECCC) facility for generating the pET22b-hGal-8 construct.

Abbreviations

2'-FL = 2'-fucosyllactose, 3-FL = 3-fucosyllactose, 3'-SL = 3'-sialyllactose, 6'-SL = 6'-sialyllactose, AEAB = 2-ethyl-N-(aminoethyl)-benzamide, CFG = Consortium for Functional Glycomics, FOS = fructooligosaccharides, Fuc: fucose, Gal: galactose, GalNAc: N-acetylgalactosamine, GlcNAc: N-acetylglucosamine, GOS = galactooligosaccharides, hGal: human galectin, HMG = human milk glycan, ITC = isothermal titration microcalorimetry, Lac = lactose, LacNAc = N-acetyllactosamine, LNFPI: lacto-N-fucopentaose I, LNnT: lacto-N-neotetraose, LNT = lacto-N-tetraose, Neu5Ac: 5-N-acetylneuraminic acid, PBS = phosphate-buffered saline, RFU = relative fluorescence units

Tables

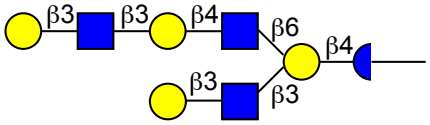
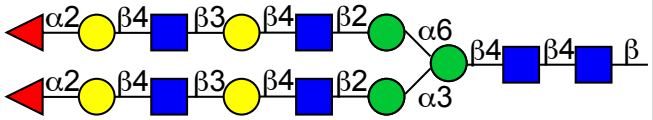
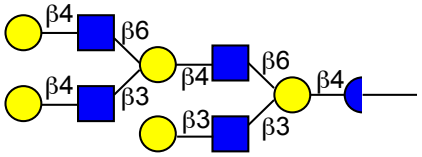
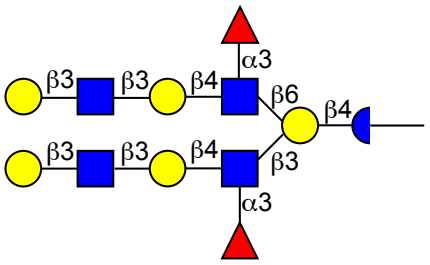
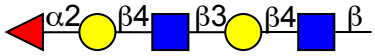
Table I. Major HM-SGM-v2 Motifs Recognized by Galectins.

Galectin	HMG Motif(s)[a, b]
Galectin-1	
Galectin-3	
Galectin-4	
Galectin-7	
Galectin-8	
Galectin-9	undefined neutral, nonfucosylated motif

[a]Motifs are proposed structures based on manual inspection of HMG-260 microarray data and the known HMG structure(s) within the bound samples

[b]Monosaccharide key: = D-glucose, = D-galactose, = L-fucose, = D-N-acetylglucosamine

Table II. Comparison of the Major Galectin-7 Binders on HM-SGM-v2 and CFG Microarrays

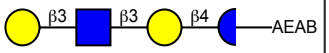
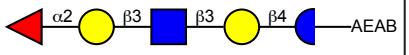
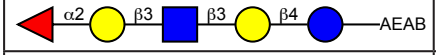
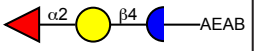
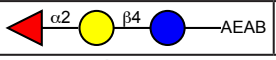
HM-SGM-v2 Microarray[b]	Consortium for Functional Glycomics Glycan Microarray v5.1[c]
	
	
	

[a]Monosaccharide key: = D-mannose, = D-galactose, = L-fucose, = D-N-acetylglucosamine, = "open-ring" glucose

[b]Shown are examples of structures found in three different samples bound on the HMG-260 Microarray. Many other samples were bound but are not shown here.

[c]These were the only three structures bound on the CFG Microarray (average rank of 10 or greater)

Table III. Comparison of Galectin Binding to "Open-Ring" vs. Corresponding "Closed-Ring" HMG Structures on Defined HMG Array

Structure	hGal-1 Average Rank	hGal-3 Average Rank	hGal-4 Average Rank	hGal-7 Average Rank	hGal-9 Average Rank
	83	n.b.[b]	28	66	n.b.
	72	17	62	100	52
	85	n.b.	10	9	n.b.
	87	30	89	20	70
	100	29	57	77	n.b.
	72	100	99	96	100
	17	n.b.	4	n.b.	n.b.
	41	n.b.	61	12	n.b.

[a]Monosaccharide key: = D-Glucose, = D-galactose, = L-fucose, = D-N-acetylglucosamine, = "open-ring" glucose

[b]n.b. = not bound

Table IV. ITC Measurements of Human Galectin-7 with HMGs

HMG	K_a (M^{-1} x 10^{-3})	K_d (mM)	ΔH (kcal/ mol)	ΔS (cal/ mol/K)	ΔG (kcal/ mol)
Lactose	2.88 (0.075)[a]	0.348	-10.695 (0.0050)	-20.1	-4.72
LNnT	6.6 (0.40)	0.15	-11.8 (0.31)	-22.3	-5.2
LNT	10.5 (0.32)	0.0952	-9.7 (0.15)	-14.1	-5.48
LNFPI	9.8 (0.31)	0.10	-8.1 (0.13)	-8.8	-5.4
2'-FL	3.0 (0.18)	0.34	-8.7 (0.31)	-13.2	-4.7
3-FL	n/a[b]	>1	n/a	n/a	n/a

[a]Numbers in parentheses represent the uncertainty. For Lactose and LNnT, the uncertainty is the standard error of the mean (SEM) of two independent experiments. For all other HMGs, the uncertainties are the Origin 7-calculated curve-fitting uncertainties from a single experiment.
[b]n/a = data not available due to lack of measurable binding

Legends to figures

Figure 1. Summary of HM-SGM-v2 Microarray Binding by Galectins.

Data are examples of one concentration of each biotinylated galectin tested screened on the HM-SGM-v2 shotgun microarray, with Streptavidin-Cy5 used for detection. The concentrations fell within the approximate linear range of binding to highlight the strongest bound samples. Error bars represent the standard deviation of binding to four technical replicates printed on the array. Refer to Supplementary File 1 for the total data from these screenings at all concentrations of all galectins.

Figure 2. CFG Glycan Microarray Version 5.1 Results for Galectin-2 Binding.

(a) Galectin-2 binding to the CFG Microarray at 20μg/ml and 200μg/ml. Galectin-2 was also screened at 2μg/ml (not shown here; Supplementary File 2) but showed no binding. (b) A list of the top five structures bound by galectin-2. These five structures were bound at both 20μg/ml and 200μg/ml. The fifth structure shown in this table was bound in a dose-independent manner, suggesting this was a non-specific binder. Note that additional structures were bound at 200μg/ml galectin-2 only (Supplementary File 2) but are not shown here. (c) Proposed glycan binding motif of galectin-2 based on manual inspection of the structures in (b) and Glycopattern analysis. This structure represents the Histo-Blood Group Antigen A Type 2 (i.e. Blood Group A2) determinant.

Figure 3. Summary of Defined HMG Microarray Binding by Galectins.

One concentration of each biotinylated galectin screened on the HM-SGM-v2 shotgun microarray (left panels) along with the same concentration of galectin screened in the

presence of 0.1M lactose (right panels). Streptavidin-Cy5 was used for detection. Error bars represent the standard deviation of binding to four technical replicates printed on the array after removing the highest and lowest RFU value of six total technical replicates. Refer to Supplementary File 3 for the total data from these screenings at all concentrations of all galectins in the presence and absence of 0.1M lactose.

Figure 4. Summary of Inhibition of Galectin Binding to the Defined HMG Microarray by Free HMGs.

Biotinylated galectin-7 (a) and biotinylated galectin-4 (b) were screened on the defined HMG microarray in the presence or absence of 50μM 2-‘FL, 500μM 2’-FL, 5mM 2’-FL, or 5mM 3-FL. Streptavidin-Cy5 was used for detection. Error bars represent the standard deviation of binding to four technical replicates printed on the array after removing the highest and lowest RFU value of six total technical replicates. The y-axis is set to the same scale for all graphs. Similar results were seen for 50μM, 500μM, and 5mM LNT, LNnT, and LNFPI as 2’-FL but are not shown here (refer to Supplementary Figure 4 for the total inhibition data for hGal-4 and hGal-7).

Figure 5. ITC thermograms for galectin-7 with HMGs and curve-fitting results after subtraction of buffer-HMG titration data.

Thermograms are 28μM hGal-7 with 8.6mM lactose (top left and top center), 9.1mM 2’-FL (top right), 9.1mM LNnT (middle left), 4.6mM LNnT (center), 4.28mM LNT (middle right), 4.11mM LNFPI (bottom left), and 4.89mM 3-FL (bottom center). One-Site Model curve-fitting results are shown in the box below each thermogram.

Figures

Figure 1.

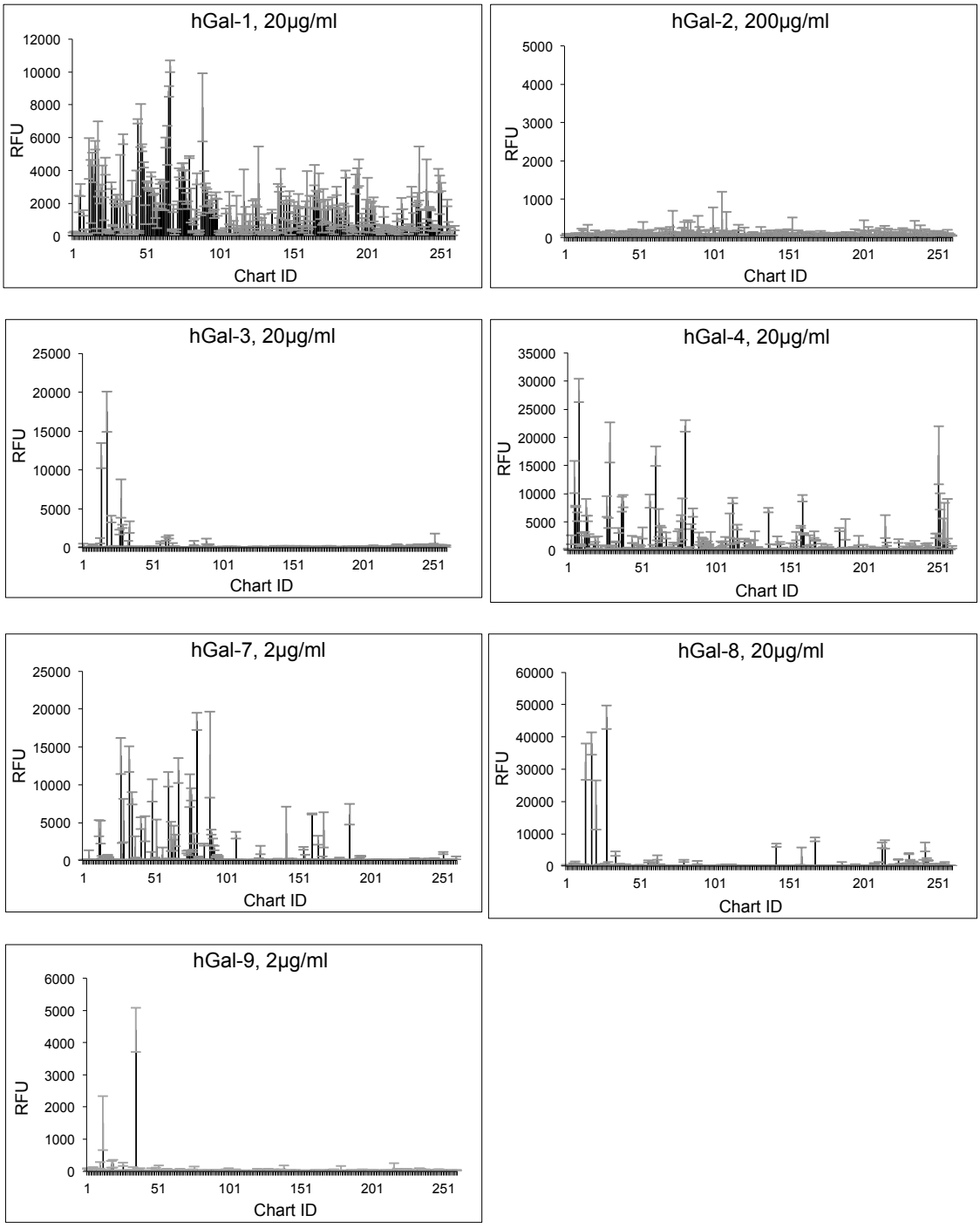
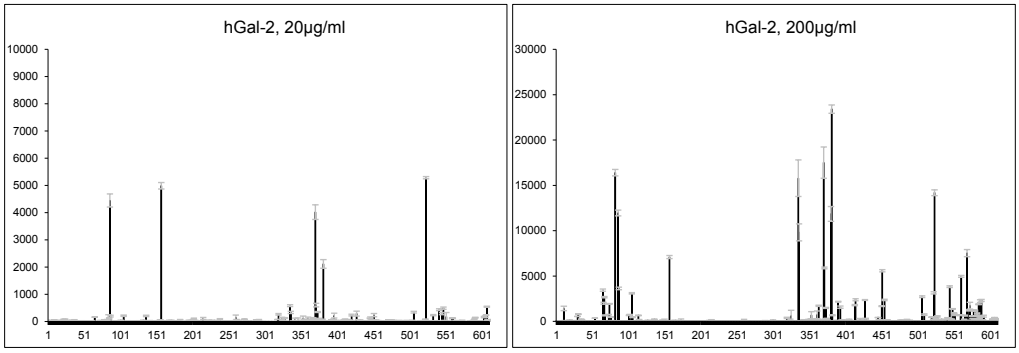


Figure 2.

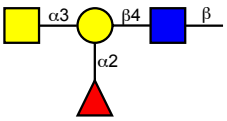
a)



b)

Structure[a]	Average Rank
α3 β4 β3 α Sp0; α2	80
α3 β4 β3 α6; α2; β4 β4 β Sp20; α3 α3 β4 β3; α2	75
β3 α3 β4 β Sp14; α2	70
α3 β4 β Sp8; α2	68
β4 α3 β4 β Sp8; α2	62[b]

c)



[a]Monosaccharide key: = D-mannose, = D-galactose, = L-fucose, = D-N-acetylglucosamine, = D-N-acetylgalactosamine

[b]This sample was bound in a dose-independent manner (i.e.-fold-change in average RFU between 20µg/ml and 200µg/ml human galectin-2 screenings was < 2-fold)

Figure 3.

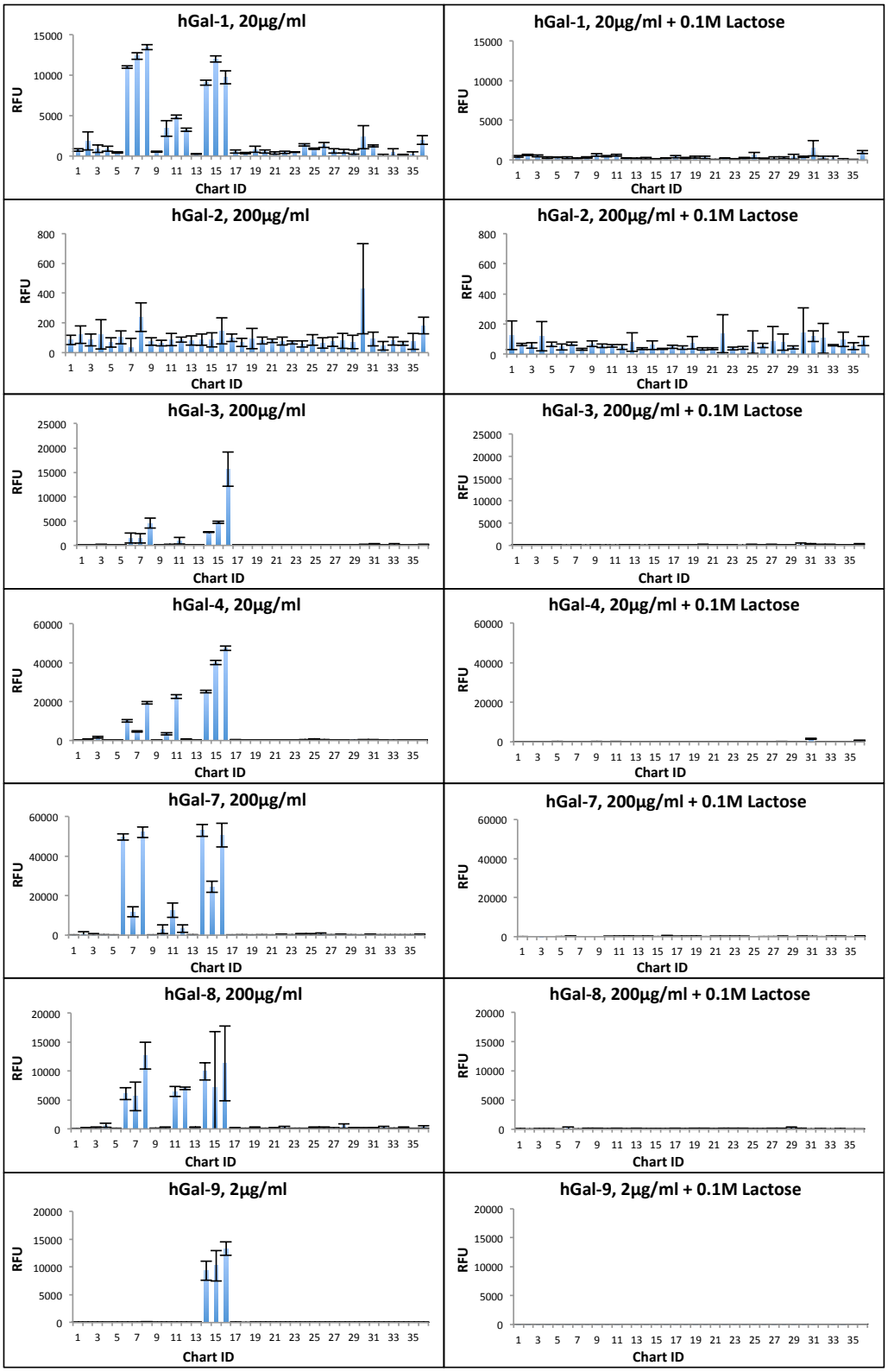

hGal-1, 20µg/ml
hGal-1, 20µg/ml + 0.1M Lactose
hGal-2, 200µg/ml
hGal-2, 200µg/ml + 0.1M Lactose
hGal-3, 200µg/ml
hGal-3, 200µg/ml + 0.1M Lactose
hGal-4, 20µg/ml
hGal-4, 20µg/ml + 0.1M Lactose
hGal-7, 200µg/ml
hGal-7, 200µg/ml + 0.1M Lactose
hGal-8, 200µg/ml
hGal-8, 200µg/ml + 0.1M Lactose
hGal-9, 2µg/ml
hGal-9, 2µg/ml + 0.1M Lactose
RFU
Chart ID

Figure 4.

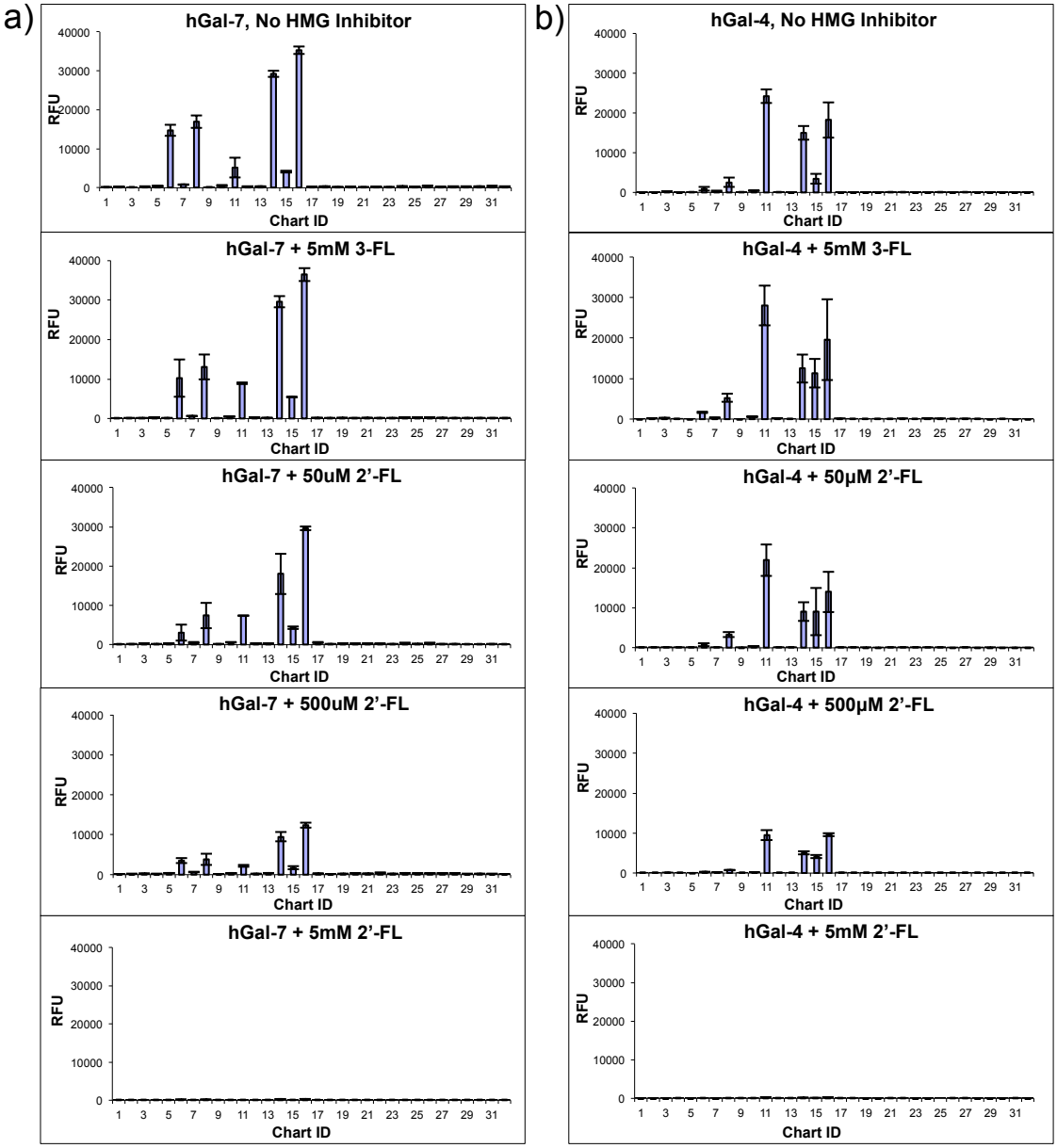
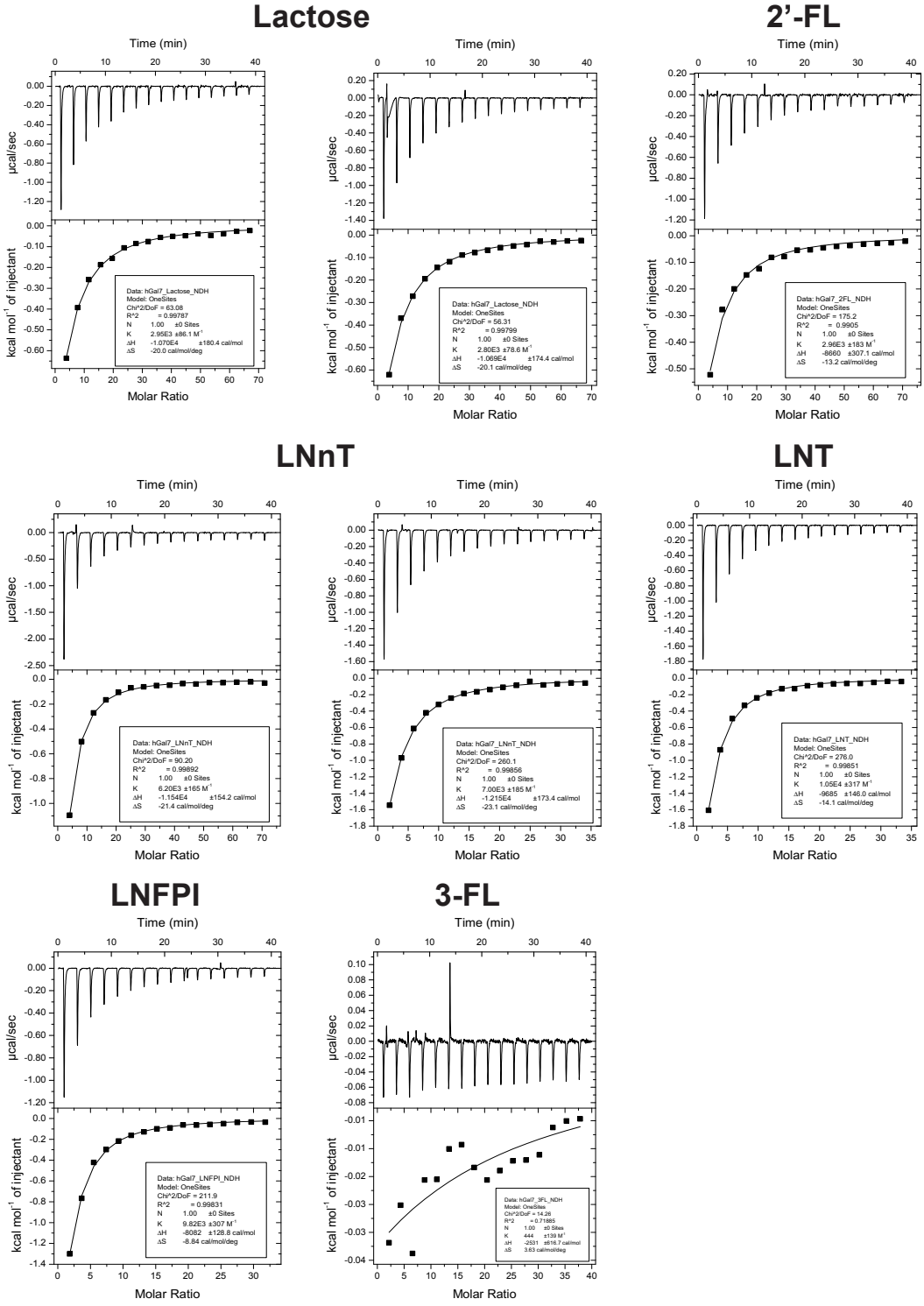

a)
hGal-7, No HMG Inhibitor
hGal-7 + 5mM 3-FL
hGal-7 + 50uM 2'-FL
hGal-7 + 500uM 2'-FL
hGal-7 + 5mM 2'-FL
b)
hGal-4, No HMG Inhibitor
hGal-4 + 5mM 3-FL
hGal-4 + 50µM 2'-FL
hGal-4 + 500µM 2'-FL
hGal-4 + 5mM 2'-FL
RFU
Chart ID

Figure 5.


Lactose
2'-FL
LNnT
LNT
LNFPI
3-FL
Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio
Data: hGal7_Lactose_NDH
Model: OneSites
Chi^2/DoF = 63.08
R^2 = 0.99787
N 1.00 ±0 Sites
K 2.95E3 ±86.1 M-1
ΔH -1.070E4 ±180.4 cal/mol
ΔS -20.0 cal/mol/deg
Data: hGal7_Lactose_NDH
Model: OneSites
Chi^2/DoF = 56.31
R^2 = 0.99799
N 1.00 ±0 Sites
K 2.80E3 ±78.6 M-1
ΔH -1.069E4 ±174.4 cal/mol
ΔS -20.1 cal/mol/deg
Data: hGal7_2FL_NDH
Model: OneSites
Chi^2/DoF = 175.2
R^2 = 0.9905
N 1.00 ±0 Sites
K 2.96E3 ±183 M-1
ΔH -8660 ±307.1 cal/mol
ΔS -13.2 cal/mol/deg
Data: hGal7_LNnT_NDH
Model: OneSites
Chi^2/DoF = 90.20
R^2 = 0.99892
N 1.00 ±0 Sites
K 6.20E3 ±165 M-1
ΔH -1.154E4 ±154.2 cal/mol
ΔS -21.4 cal/mol/deg
Data: hGal7_LNnT_NDH
Model: OneSites
Chi^2/DoF = 260.1
R^2 = 0.99856
N 1.00 ±0 Sites
K 7.00E3 ±185 M-1
ΔH -1.215E4 ±173.4 cal/mol
ΔS -23.1 cal/mol/deg
Data: hGal7_LNT_NDH
Model: OneSites
Chi^2/DoF = 276.0
R^2 = 0.99851
N 1.00 ±0 Sites
K 1.05E4 ±317 M-1
ΔH -9685 ±146.0 cal/mol
ΔS -14.1 cal/mol/deg
Data: hGal7_LNFPI_NDH
Model: OneSites
Chi^2/DoF = 211.9
R^2 = 0.99831
N 1.00 ±0 Sites
K 9.82E3 ±307 M-1
ΔH -8082 ±128.8 cal/mol
ΔS -8.84 cal/mol/deg
Data: hGal7_3FL_NDH
Model: OneSites
Chi^2/DoF = 14.28
R^2 = 0.71885
N 1.00 ±0 Sites
K 444 ±139 M-1
ΔH -2531 ±616.7 cal/mol
ΔS 3.63 cal/mol/deg

Supplementary Material List

Supplementary material to this manuscript is accessible online from the publisher's website and includes:

Supplementary Information (includes Supplementary Materials & Methods, Supplementary Figure 1, and Supplementary Table 1)

Supplementary File 1. HM-SGM-v2 Microarray Galectin Binding Data

This Excel spreadsheet, containing 7 worksheets, includes the average RFU data and graphs of binding data for 2.0, 20, and 200µg/ml hGal-1 C2S, hGal-3, hGal-4, hGal-7, and hGal-8; 0.2, 2.0, and 20µg/ml hGal-9; and 200µg/ml hGal-2 screened on the HM-SGM-v2. The results for each galectin are presented in different worksheets.

Supplementary File 2. Human Galectin-2 and -7 CFG Glycan Microarray Binding Data

This Excel spreadsheet, containing 9 worksheets, includes the average RFU data and graphs of binding data for 2.0µg/ml, 20µg/ml, and 200µg/ml hGal-2 as well as 0.5µg/ml, 2.0µg/ml, 5.0µg/ml, and 200µg/ml hGal-7. Each concentration is presented in different worksheets. The rank analysis data and calculations for hGal-2 and hGal-7 are also included as their own individual worksheets.

Supplementary File 3. Defined HMG Microarray Galectin Binding Data

This Excel spreadsheet, containing 7 worksheets, includes the average RFU data and graphs of binding data for 2.0, 20, and 200µg/ml hGal-1 C2S, hGal-3, hGal-4, hGal-7, and hGal-8; 0.2 and 2.0µg/ml hGal-9; and 200µg/ml hGal-2 screened on the defined HMG microarray. The data includes the results for the galectins screened in the presence or absence of 0.1M lactose. The results for each galectin are presented in different worksheets.

Supplementary File 4. Inhibition of Galectin-4 and -7 Defined HMG Array Binding

This Excel spreadsheet contains 4 worksheets of the total data for hGal-4 and hGal-7 binding to the defined HMG microarray in the presence or absence of 50µM, 500µM, and 5.0mM of the free HMGs LNnT, LNT, LNFPI, 2'-FL, and 3-FL. The four worksheets contain all the hGal-4 average RFU data, hGal-4 binding data graphs, hGal-7 average RFU data, and hGal-7 binding data graphs, respectively.

Supplementary Information

Anti-Galectin Antiserum Generation

Recombinant galectins were purified as described in Materials & Methods. For galectin-9, a construct expressing only the C-terminus of galectin-9 (hGal-9C) was used by cloning the sequence encoding the final 151 amino acids of human galectin-9 into pET32b. Prior to immunization, any fusion tags were removed and separated from the galectin protein by repurification on lactosyl-Sepharose. The purified, tagless galectin proteins were then sent to ProSci for immunization in rabbits. The antiserum was tested by dot blotting and Western blotting to confirm the absence of cross-reactivity with other recombinant galectins.

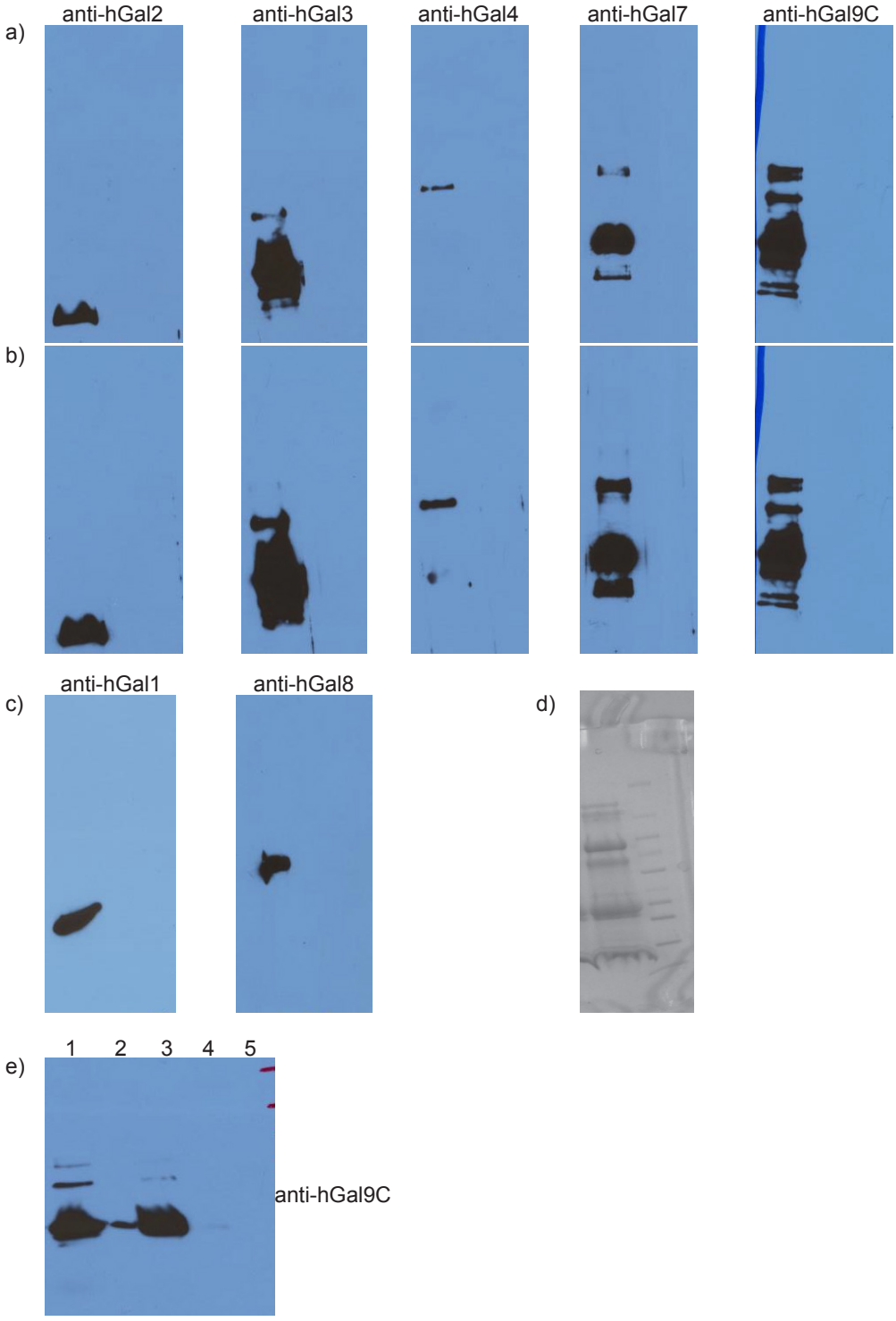
Detection of Galectins in Human Milk

Human milk was obtained from the Mother's Milk Bank (Austin, TX). Whole human milk was first dialyzed into PBS + 14mM β-mercaptoethanol using a 1000Da molecular weight cutoff tubing (Spectrum). Next, the dialyzed sample was centrifuged a low speed and the cream layer removed. The total protein content of the defatted, dialyzed milk was measured by BCA Assay (Pierce). The presence of protein was further assayed by analyzing the defatted, dialyzed milk by SDS-PAGE on a Mini-PROTEAN TGX 4-20% SDS-PAGE gel (Bio-Rad) and Coomassie Brilliant Blue R-250 staining (Bio-Rad).

To analyze the milk preparation for galectins, 300μg of human milk protein was then loaded on a Mini-PROTEAN TGX 4-20% SDS-PAGE gel (Bio-Rad) in one lane. A second lane contained 10μg of galectin as a positive control. Following electrophoresis,

the proteins were transferred to a nitrocellulose membrane using an iBlot (Invitrogen). Western Blotting was then performed using the corresponding anti-galectin antiserum at 1:10,000-1:50,000 dilution with 1:5000 goat anti-rabbit IgG-HRP as the detection reagent. Blots were developed with SuperSignal West Pico Reagent (Pierce).

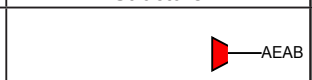
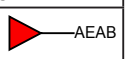
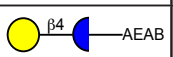
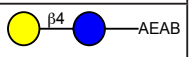
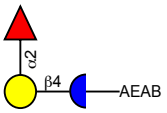
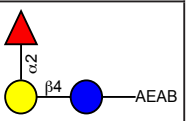
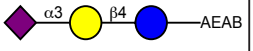
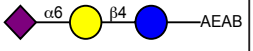
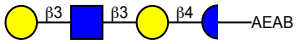
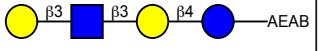
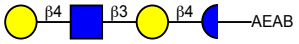
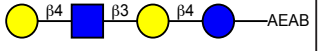
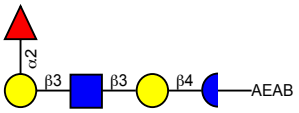
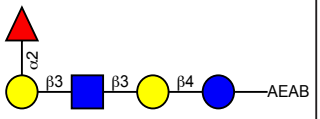
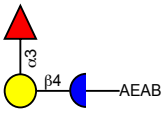
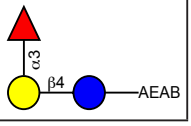
To determine the limit of galectin detection in human milk, 10μg or 5ng of recombinant hGal-9C was run on SDS-PAGE in the presence or absence of 300μg milk protein. 300μg milk protein without added galectin was also run. SDS-PAGE, Western Transfer, and Western Blotting were performed as described above using anti-Gal-9C antiserum.


a)
anti-hGal2
anti-hGal3
anti-hGal4
anti-hGal7
anti-hGal9C
b)
c)
anti-hGal1
anti-hGal8
d)
e)
1 2 3 4 5
anti-hGal9C

Supplementary Figure 1. Lack of Detection of Galectins in Human Milk

a-c) 10μg of recombinant galectin (left lane) and 300μg milk protein (right lane) were assayed by Western Blotting with the corresponding anti-galectin antiserum. The anti-hGal-2, -3, -4, -7, and -9C blots were simultaneously developed using either a) 3 minutes or b) 15 minutes of exposure. c) The anti-hGal-1 and -8 blots were simultaneously developed (but separately from the other galectins) using 10 minutes of exposure. d) Coomassie Brilliant Blue R-250 staining of 20μl (~200μg) of defatted, dialyzed milk protein. Left lane: milk, right lane: Spectra Multicolor Broad Range Protein Ladder. e) Limit of detection of galectin-9 in human milk. 10μg or 5ng of recombinant hGal-9C was loaded on an SDS-PAGE gel in the presence or absence of 300μg of milk protein. 20μl milk (~300μg milk protein) was also assayed. The blot was overlaid with anti-hGal-9C antiserum. The lanes contained the following samples: 1: 10μg hGal9C, 2: 5ng hGal-9C, 3: 10μg hGal-9C + 300μg milk protein, 4: 5ng hGal-9C + 300μg milk protein, 5: ~300μg milk protein.

Supplementary Table I. List of HMG structures found on the defined HMG microarray

Open-Ring Glycan Structures		Closed-Ring Glycan Structures	
Name (Abbreviation)	Structure	Name (Abbreviation)	Structure
Fucose (Fuc)		Fucose (GG-Fuc)	
Lactose (Lac)		Lactose (GG-Lac)	
2'-fucosyllactose (2'-FL)		2'-fucosyllactose (GG-2'-FL)	
3'-sialyllactose (3'-SL)		3'-sialyllactose (GG-3'-SL)	
6'-sialyllactose (6'-SL)		6'-sialyllactose (GG-6'-SL)	
Lacto-N-Tetraose (LNT)		Lacto-N-Tetraose (GG-LNT)	
Lacto-N-neoTetraose (LNnT)		Lacto-N-neoTetraose (GG-LNnT)	
Lacto-N-Fucopentaose I (LNFPI)		Lacto-N-Fucopentaose I (GG-LNFPI)	
3-fucosyllactose (3-FL)		3-fucosyllactose (GG-3-FL)	

[a]Monosaccharide key: = D-Glucose, = D-galactose, = L-fucose, = D-N-acetylglucosamine, = N-acetylneuraminic acid, = "open-ring" glucose, = "open-ring" fucose

[b]Galactooligosaccharides (GOS) are also found on this microarray but are not depicted here because of the complexity and heterogeneity of the GOS sample used in this study

References

1. World Health Organization, *Infant and young child feeding: Model Chapter*, 2009, WHO Press: Geneva, Switzerland.
2. Kunz, C., et al., *Oligosaccharides in human milk: structural, functional, and metabolic aspects.* Annu Rev Nutr, 2000. **20**: p. 699-722.
3. Bode, L., *Human milk oligosaccharides: every baby needs a sugar mama.* Glycobiology, 2012. **22**(9): p. 1147-62.
4. Newburg, D.S., G.M. Ruiz-Palacios, and A.L. Morrow, *Human milk glycans protect infants against enteric pathogens.* Annu Rev Nutr, 2005. **25**: p. 37-58.
5. He, Y., et al., *Human colostrum oligosaccharides modulate major immunologic pathways of immature human intestine.* Mucosal Immunol, 2014. **7**(6): p. 1326-39.
6. Bhargava, P., et al., *Immunomodulatory glycan LNFPIII alleviates hepatosteatosis and insulin resistance through direct and indirect control of metabolic pathways.* Nat Med, 2012. **18**(11): p. 1665-72.
7. Bienenstock, J., et al., *Fucosylated but not sialylated milk oligosaccharides diminish colon motor contractions.* PLoS One, 2013. **8**(10): p. e76236.
8. Engfer, M.B., et al., *Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract.* Am J Clin Nutr, 2000. **71**(6): p. 1589-96.
9. Gnoth, M.J., et al., *Human milk oligosaccharides are minimally digested in vitro.* J Nutr, 2000. **130**(12): p. 3014-20.

10. Ward, R.E., et al., *In vitro fermentation of breast milk oligosaccharides by Bifidobacterium infantis and Lactobacillus gasseri.* Appl Environ Microbiol, 2006. **72**(6): p. 4497-9.

11. Asakuma, S., et al., *Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria.* J Biol Chem, 2011. **286**(40): p. 34583-92.

12. Cummings, R.D., Liu, F. T., *Galectins*, in *Essentials of Glycobiology*, A. Varki, Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W., Etzler, M.E., Editor 2009, Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY).

13. Barondes, S.H., et al., *Galectins: a family of animal beta-galactoside-binding lectins.* Cell, 1994. **76**(4): p. 597-8.

14. Uhlen, M., et al., *Proteomics. Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.

15. Magnaldo, T., D. Fowlis, and M. Darmon, *Galectin-7, a marker of all types of stratified epithelia.* Differentiation, 1998. **63**(3): p. 159-68.

16. Huflejt, M.E. and H. Leffler, *Galectin-4 in normal tissues and cancer.* Glycoconj J, 2004. **20**(4): p. 247-55.

17. Saal, I., et al., *Human galectin-2: expression profiling by RT-PCR/immunohistochemistry and its introduction as a histochemical tool for ligand localization.* Histol Histopathol, 2005. **20**(4): p. 1191-208.

18. Rudloff, S., et al., *Urinary excretion of lactose and oligosaccharides in preterm infants fed human milk or infant formula.* Acta Paediatr, 1996. **85**(5): p. 598-603.

19. Rudloff, S., et al., *Urinary excretion of in vivo (1)(3)C-labelled milk oligosaccharides in breastfed infants.* Br J Nutr, 2012. **107**(7): p. 957-63.

20. Goehring, K.C., et al., *Direct evidence for the presence of human milk oligosaccharides in the circulation of breastfed infants.* PLoS One, 2014. **9**(7): p. e101692.

21. Yu, Y., et al., *Human milk contains novel glycans that are potential decoy receptors for neonatal rotaviruses.* Mol Cell Proteomics, 2014. **13**(11): p. 2944-2960.

22. Ashline, D.J., et al., *Structural characterization by MSn of human milk glycans recognized by human rotaviruses.* Mol Cell Proteomics, 2014. **13**(11): p. 2961-2974.

23. Cummings, R.D. and J.M. Pierce, *The challenge and promise of glycomics.* Chem Biol, 2014. **21**(1): p. 1-15.

24. Sparrow, C.P., H. Leffler, and S.H. Barondes, *Multiple soluble beta-galactoside-binding lectins from human lung.* J Biol Chem, 1987. **262**(15): p. 7383-90.

25. Agravat, S.B., et al., *GlycoPattern: a web platform for glycan array mining.* Bioinformatics, 2014. **30**(23): p. 3417-3418.

26. Hirabayashi, J., et al., *Oligosaccharide specificity of galectins: a search by frontal affinity chromatography.* Biochim Biophys Acta, 2002. **1572**(2-3): p. 232-54.

27. Oozeer, R., et al., *Intestinal microbiology in early life: specific prebiotics can have similar functionalities as human-milk oligosaccharides.* Am J Clin Nutr, 2013. **98**(2): p. 561S-71S.

28. Song, X., et al., *Novel fluorescent glycan microarray strategy reveals ligands for galectins.* Chem Biol, 2009. **16**(1): p. 36-47.

29. Song, X., et al., *Fluorescent glycosylamides produced by microscale derivatization of free glycans for natural glycan microarrays.* ACS Chem Biol, 2009. **4**(9): p. 741-50.

30. Carlsson, S., et al., *Affinity of galectin-8 and its carbohydrate recognition domains for ligands in solution and at the cell surface.* Glycobiology, 2007. **17**(6): p. 663-76.

31. Ideo, H., et al., *High-affinity binding of recombinant human galectin-4 to SO(3)(-)-->3Galbeta1-->3GalNAc pyranoside.* Glycobiology, 2002. **12**(3): p. 199-208.

32. Ahmad, N., et al., *Thermodynamic binding studies of cell surface carbohydrate epitopes to galectins-1,-3, and-7: Evidence for differential binding specificities.* Canadian Journal of Chemistry-Revue Canadienne De Chimie, 2002. **80**(8): p. 1096-1104.

33. Turnbull, W.B. and A.H. Daranas, *On the value of c: can low affinity systems be studied by isothermal titration calorimetry?* J Am Chem Soc, 2003. **125**(48): p. 14859-66.

34. Lobsanov, Y.D., et al., *X-ray crystal structure of the human dimeric S-Lac lectin, L-14-II, in complex with lactose at 2.9-A resolution.* J Biol Chem, 1993. **268**(36): p. 27034-8.

35. Bode, L., *Recent advances on structure, metabolism, and function of human milk oligosaccharides.* J Nutr, 2006. **136**(8): p. 2127-30.

36. Bachhawat-Sikder, K., C.J. Thomas, and A. Surolia, *Thermodynamic analysis of the binding of galactose and poly-N-acetyllactosamine derivatives to human galectin-3.* FEBS Lett, 2001. **500**(1-2): p. 75-9.

37. Saarela, T., J. Kokkonen, and M. Koivisto, *Macronutrient and energy contents of human milk fractions during the first six months of lactation.* Acta Paediatr, 2005. **94**(9): p. 1176-81.

38. Kien, C.L., R.E. McClead, and L. Cordero, Jr., *In vivo lactose digestion in preterm infants.* Am J Clin Nutr, 1996. **64**(5): p. 700-5.

39. Commare, C.E. and K.A. Tappenden, *Development of the infant intestine: implications for nutrition support.* Nutr Clin Pract, 2007. **22**(2): p. 159-73.

40. Urashima, T., et al., *The predominance of type I oligosaccharides is a feature specific to human breast milk.* Adv Nutr, 2012. **3**(3): p. 473S-82S.

41. Shan, M., et al., *Mucus enhances gut homeostasis and oral tolerance by delivering immunoregulatory signals.* Science, 2013. **342**(6157): p. 447-53.

42. Stowell, S.R., et al., *Galectin-1, -2, and -3 exhibit differential recognition of sialylated glycans and blood group antigens.* J Biol Chem, 2008. **283**(15): p. 10109-23.

43. Stowell, S.R., et al., *Innate immune lectins kill bacteria expressing blood group antigen.* Nat Med, 2010. **16**(3): p. 295-301.

44. Stowell, S.R., et al., *Differential roles of galectin-1 and galectin-3 in regulating leukocyte viability and cytokine secretion.* J Immunol, 2008. **180**(5): p. 3091-102.

45. Matsushita, N., et al., *Requirement of divalent galactoside-binding activity of ecalectin/galectin-9 for eosinophil chemoattraction.* J Biol Chem, 2000. **275**(12): p. 8355-60.

46. Paclik, D., et al., *Galectin-2 and -4, but not galectin-1, promote intestinal epithelial wound healing in vitro through a TGF-beta-independent mechanism.* Inflamm Bowel Dis, 2008. **14**(10): p. 1366-72.

47. Cerliani, J.P., et al., *Expanding the universe of cytokines and pattern recognition receptors: galectins and glycans in innate immunity.* J Clin Immunol, 2011. **31**(1): p. 10-21.

48. Di Lella, S., et al., *When galectins recognize glycans: from biochemistry to physiology and back again.* Biochemistry, 2011. **50**(37): p. 7842-57.

49. Yu, Y., et al., *Functional glycomic analysis of human milk glycans reveals presence of virus receptors and embryonic stem cell biomarkers.* J Biol Chem, 2012. **287**(53): p. 44784-44799.

50. Molinari, C.E., et al., *Proteome mapping of human skim milk proteins in term and preterm milk.* J Proteome Res, 2012. **11**(3): p. 1696-714.

51. Coscia, A., et al., *Detection of cow's milk proteins and minor components in human milk using proteomics techniques.* J Matern Fetal Neonatal Med, 2012. **25 Suppl 4**: p. 54-6.

52. Kohatsu, L., et al., *Galectin-3 induces death of Candida species expressing specific beta-1,2-linked mannans.* J Immunol, 2006. **177**(7): p. 4718-26.

53. Stowell, S.R., et al., *Microbial glycan microarrays define key features of host-microbial interactions.* Nat Chem Biol, 2014. **10**(6): p. 470-6.

54. Hirabayashi, J. and K. Kasai, *Effect of amino acid substitution by sited-directed mutagenesis on the carbohydrate recognition and stability of human 14-kDa beta-galactoside-binding lectin.* J Biol Chem, 1991. **266**(35): p. 23648-53.

55. Massa, S.M., et al., *L-29, an endogenous lectin, binds to glycoconjugate ligands with positive cooperativity.* Biochemistry, 1993. **32**(1): p. 260-7.

56. Levi, G. and V.I. Teichberg, *Isolation and physicochemical characterization of electrolectin, a beta-D-galactoside binding lectin from the electric organ of Electrophorus electricus.* J Biol Chem, 1981. **256**(11): p. 5735-40.

57. Song, X., et al., *Derivatization of free natural glycans for incorporation onto glycan arrays: derivatizing glycans on the microscale for microarray and other applications (ms# CP-10-0194).* Curr Protoc Chem Biol, 2011. **3**(2): p. 53-63.

58. Heimburg-Molinaro, J., et al., *Preparation and analysis of glycan microarrays.* Curr Protoc Protein Sci, 2011. **64**(12.10): p. 12.10.1-12.10.29.

59. Smith, D.F., X. Song, and R.D. Cummings, *Use of glycan microarrays to explore specificity of glycan-binding proteins.* Methods Enzymol, 2010. **480**: p. 417-44.

60. DuBois, M., et al., *Colorimetric Method for Determination of Sugars and Related Substances.* Analytical Chemistry, 1956. **28**(3): p. 350-356.

Chapter 3: Human DC-SIGN Binds Specific Human Milk Glycans

Alexander J. Noll[1, 2, 5], Ying Yu[1, 5], Yi Lasanajak[1], Geralyn Duska-McEwen[3], Rachael H. Buck[3], David F. Smith[1], and Richard D. Cummings[1, 4]

[1]The Glycomics Center, Department of Biochemistry,
Emory University School of Medicine, Atlanta, GA, U.S.A.

[2]Program in Microbiology and Molecular Genetics, Graduate Division of Biological and Biomedical Sciences, Laney Graduate School, Emory University, Atlanta, GA

[3]Abbott Nutrition, Global Discovery R&D, Columbus, Ohio, U.S.A

[4]Current Address: Dept. of Surgery, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, U. S. A.

[5]These authors contributed equally to this work. YY performed the experiments, with assistance from YL. AJN, YY, DFS, and RDC analyzed the data. AJN wrote the manuscript and generated all the figures, with input from all the authors.

This is a pre-copyedited, author-produced PDF of an article accepted for publication in *Biochemical Journal* following peer review. The version of record [Noll AJ et al. *Biochem J* 2016 Mar 14; pii: BCJ20160046.] is available online at: http://www.biochemj.org/content/early/2016/03/14/BCJ20160046.long

Abstract

Human milk glycans (HMGs) are prebiotics, pathogen receptor decoys, and regulators of host physiology and immune responses. Mechanistically, human lectins (glycan-binding proteins, hGBPs) expressed by dendritic cells (DC) are of major interest, as these cells directly contact HMGs. To explore such interactions, we screened many C-type lectins and Siglecs expressed by DC for glycan binding on microarrays presenting over 200 HMGs. Unexpectedly, DC-SIGN showed robust binding to many HMGs, whereas other C-type lectins failed to bind, and Siglecs-5 and -9 showed weak binding to a few glycans. By contrast, most hGBPs bound to multiple glycans on other microarrays lacking HMGs. An α-linked fucose residue was characteristic of HMGs bound by DC-SIGN. Binding of DC-SIGN to the simple HMGs 2'-fucosyllactose (2'-FL) and 3-fucosyllactose (3-FL) was confirmed by flow cytometry to beads conjugated with 2'-FL or 3-FL, as well as the ability of the free glycans to inhibit DC-SIGN binding. 2'-FL had an IC_{50} of ~1 mM for DC-SIGN, which is within the physiological concentration of 2'-FL in human milk. These results demonstrate that DC-SIGN among the many hGBPs expressed by DC binds to α-fucosylated HMGs, and suggest that such interactions may be important in influencing immune responses in the developing infant.

Summary Statement

This work shows that DC-SIGN, a well-known glycan-binding protein involved in immune responses, is a major receptor of specific human milk glycans. This finding may be important in understanding the role of human milk glycans in pediatric immune development.

Short Title

DC-SIGN is a Human Milk Glycan Receptor

Key Words

DC-SIGN/Glycan-Binding Proteins/Glycan Microarrays/Glycan Recognition/Human Milk Glycans

Abbreviations list

2'-FL = 2'-fucosyllactose; 3-FL = 3-fucosyllactose; 3'-SL = 3'-sialyllactose; 6'-SL = 6'-sialyllactose; AEAB = N-aminoethyl 2-aminobenzamide; CFG = Consortium for Functional Glycomics; DC = dendritic cell; DC-SIGN = Dendritic Cell-Specific Intercellular Adhesion Molecule-3-Grabbing Non-integrin; Fuc = L-fucose; Gal = D-galactose; Glc = D-glucose, GBP = glycan-binding protein; GalNAc = D-N-acetylgalactosamine; GGAEAB = glycosylamide-glycyl-N-aminoethyl 2-aminobenzamide; GI = gastrointestinal; GlcNAc = D-N-acetylglucosamine; GOS = galactooligosaccharides; H1 = Blood Group H Type 1; hGBP = human glycan-binding protein; HMGs = human milk glycans; HM-SGM-v2 = human milk shotgun glycan microarray version 2; LNFPI = lacto-N-fucopentaose I; LNnT = lacto-N-neotetraose; LNT = lacto-N-tetraose; MAGS = Metadata-Assisted Glycan Sequencing; MALDI TOF MS = matrix-assisted laser desorption/ionization-time of flight mass spectrometry; MGL = macrophage galactose/N-acetylgalactosamine-type lectin; MS^n = multi-dimensional mass spectrometry; Neu5Ac = 5-N-acetylneuraminic acid; NHS = N-

hydroxysuccinimide; Siglec = sialic acid–binding, immunoglobulin-like lectin; SLe^a = sialyl Lewis a

Introduction

Carbohydrates are the most abundant class of biomolecules in human milk. The majority of this total carbohydrate (~70g/L) is lactose, a major source of energy for infants, and the remainder (5-20g/L) consists of non-digestible, larger-sized glycans that are derived from lactose [1-3]. These human milk glycans (HMGs) have been classically defined as prebiotics and receptor decoys that are predicted to prevent infection by blocking pathogen adherence to the infant epithelium [4, 5]. However, HMGs may have functions beyond interactions with microbes, as more recent studies suggest that HMGs may regulate multiple physiological functions in infants, including gene expression and immune and allergic responses [6, 7]. HMGs also regulate gut motility [8] and enhance learning and memory [9], suggesting their role in neuronal responses and cognition. However, the mechanisms underlying these physiological functions of HMGs are still unclear.

Human lectins (glycan-binding proteins, hGBPs) play numerous roles in physiology and immunity, including regulation of gene expression and immune responses, pathogen sensing, cell-cell interactions, and tissue homing [10-12]. The glycan specificities of many hGBPs have been explored by multiple techniques, but the most powerful new approach has utilized glycan microarrays in which hundreds of structurally defined glycan ligands are displayed on a single slide, as developed by the Consortium for Functional Glycomics (CFG) (http://www.functionalglycomics.org/glycomics/publicdata/primaryscreen.jsp). These studies have shown that each hGBP has a restricted specificity, even within a given

hGBP family [13]. The binding of different hGBPs to specific glycan determinants allows different hGBPs to regulate specific physiological functions.

There have been some recent studies broadly examining hGBP glycan specificity toward HMGs [14-16], and such general screening suggests that some HMGs may be recognized by specific hGBP. By extension, we hypothesized that HMGs might serve as general ligands for many hGBP, which could be important in modulating the hGBP downstream effector or signaling functions. The purpose of our study was to identify hGBP that bind HMGs, investigate glycan determinant specificity and the extent of the human milk metaglycome bound, and determine if binding occurs at physiologically relevant concentrations.

To address these questions, we focused on those hGBPs expressed by dendritic cells (DCs), since such cells may directly contact HMGs in the developing infant intestine via dendrite extension through the intestinal epithelium [17, 18]. We screened members of the C-type lectin [19] and I-type lectin [11, 20] families for binding to a human milk shotgun glycan microarray as well as defined glycan microarrays. The results of this study showed that from this large set of hGBPs, only Dendritic Cell-Specific Intercellular Adhesion Molecule-3-Grabbing Non-integrin (DC-SIGN) was a major binder of HMGs, with multiple α-linked fucose-containing glycans bound on an array consisting of about 250 purified HMGs structures. This binding of specific HMGs by DC-SIGN suggest that DC-SIGN may serve as an HMG receptor, which may have implications in infant immunity, physiology, and development.

Materials and methods

Preparation and Screening of Microarrays

All of the recombinant hGBP used in this study were purchased from R&D Systems (Minneapolis, MN) and are shown in Table 1, which includes information on the amino acid sequences, fusion tags, and catalog numbers used. The proteins were checked for activity by their binding to one or more glycans in various glycan microarrays. The human milk shotgun glycan microarray version 2 (HM-SGM-v2), consisting of 247 purified HMGs structures and 13 controls, has been previously described [21]. The defined HMGs microarray, consisting of simple, defined HMGs structures, was generated as described previously [22]. The recombinant hGBP were screened on the Consortium for Functional Glycomics (CFG) glycan microarray version 5.1, HM-SGM-v2, and defined HMGs microarray as previously described [23]. 5μg/ml of Alexa Fluor 488-labeled anti-human IgG (Molecular Probes, Eugene, OR) or Alexa Fluor 488-labeled anti-pentaHis antibody (Qiagen, Valencia, CA) antibodies were used for detection of recombinant hGBP carrying an F_c fusion tag or 6-9x Histidine tag, respectively. As a control we screened 10μg/ml DC-SIGN binding on the HM-SGM-v2, in which Ca^{2+} was omitted from the binding buffer and replaced with 0.2mM EDTA to confirm Ca^{2+}-dependent binding. For HMG inhibition experiments, the recombinant hGBP was preincubated with free HMG or 2-ethyl-N(aminoethyl)benzamide (AEAB)-derivatized HMG containing an "open-ring" reducing end or glycosylamide-glycyl-N-aminoethyl 2-aminobenzamide (GGAEAB)-derivatized HMG containing a "closed-ring" reducing end, generated as previously described [24, 25] for 1 hour prior to screening of the defined HMG microarrays. Detection was performed using 5μg/ml Alexa Fluor 633-labeled anti-

human IgG (Molecular Probes). Rank and average rank calculations of the microarray data was performed as previously described [26]. The microarray data was manually examined for binding motifs and, for the CFG microarray data, was further analyzed with Glycopattern (https://glycopattern.emory.edu) [27] to define the CFG glycan microarray binding motif.

Preparation and Screening of HMGs Microarray for MAGS

A panel of HMGs samples bound by DC-SIGN were printed on separate Nexterion N-hydroxysuccinimide (NHS) H slides (Schott AG, Mainz, Germany) and screened with lectins, antibodies, and DC-SIGN at three different concentrations of each sample. Slide printing and sample screening were performed as previously described [23]. The anti-SLe^a antibody was purchased from Abcam. All of the other lectins, antibodies, and glycosidases used for MAGS, as well as the concentration(s) and glycosidase treatment procedures, are the same as described in a previous study [21]. Multi-dimensional mass spectrometry on HMG-9, HMG-19, and HMG-36 was performed as previously described [28].

Preparation of HMGs-Derivatized Beads and Flow Cytometry Assessment of Binding

HMGs were first derivatized with AEAB [13] by reductive amination as previously described [24]. The HMGs were then coupled to 1.00μm diameter PolyBead® Carboxylate Microspheres using the PolyLink Protein Coupling Kit (PolySciences Inc., Warrington, PA) as follows. Beads (200μl) were pelleted by gentle centrifugation at 500-1000×g for 5 minutes and resuspended in 160μl of PolyLink coupling buffer. Twenty μl

of 200mg/ml freshly prepared EDC and 20μl of freshly prepared sulfo-NHS (Thermo Scientific) were then added and the reaction incubated at room temperature with gentle rotation for 30 minutes. The beads were then washed twice with 250μl PolyLink Wash/Storage Buffer and then resuspended in 1mM glycan-AEAB in 100mM sodium phosphate pH 8.5. The reaction was incubated at room temperature with gentle mixing for 1-2 hours. The beads were washed three times with PolyLink Wash/Storage Buffer and stored at 4^{o}C in the same buffer until use.

For measurement of DC-SIGN binding to the glycan-derivatized beads, lacto-N-tetraose (LNT)-, 2'-fucosyllactose (2'-FL)-, and 3-fucosyllactose (3-FL)-derivatized beads were incubated for 1 hour with 5μg/ml of recombinant human DC-SIGN at room temperature, washed three times with PBS, and then incubated for 1 hour with 2μg/ml of Alexa Fluor 633-labeled goat anti-human IgG. As a negative control, 2'-FL-derivatized beads were incubated with secondary antibody only (no DC-SIGN). All samples were analyzed by flow cytometry with a BD FACSCalibur with the 633nm laser. 10,000 events were counted, and the FL-4 filter was used for detection. The data was analyzed using FlowJo. Gating was assigned in FlowJo by running the beads alone vs. buffer alone on a forward- vs. side-scatter plot, with >99% of the events falling in the gated area for all samples.

Results

Binding of hGBPs to the Human Milk Shotgun Glycan Microarray

A set of eight recombinant hGBPs was tested for binding to HMGs, and this set included C-type lectins and sialic acid–binding, immunoglobulin-like lectin (Siglec) members of the I-type lectin family (refer to **Table 1** for all of the hGBP used in this study). These hGBP were selected based on their stability, availability, and known expression by dendritic cells (DCs) [10, 29-31]. The C-type lectins and Siglecs were screened on a human milk shotgun glycan microarray consisting of 247 HMGs structures purified from human milk as well as 13 control glycans. This microarray was termed the HM-SGM-v2 [21]. However, only three of these hGBP, DC-SIGN, Siglec-5, and Siglec-9 showed binding to the HM-SGM-v2 (**Fig. 1**; also refer to **Supplementary File 1** for the data for all concentrations of all hGBP screened). The binding of Langerin was considered inconclusive because high concentrations of protein were needed and the signal:noise ratio was poor (**Supplementary File 1**). All other hGBP showed no evidence of binding to the HM-SGM-v2, although most bound to other glycan microarrays.

DC-SIGN bound to many glycans on the HM-SGM-v2 (**Fig. 1c**), specifically all of the glycans containing at least one α-linked fucose residue based on the calculated composition from matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS) molecular mass measurements [21]. This binding was specific in that all binding required Ca^{2+} even at the highest DC-SIGN concentration used (**Fig. 1d**). The large number of glycans bound by DC-SIGN on the HM-SGM-v2 necessitated further examination of these bound structures in order to define the HMGs determinant recognized by DC-SIGN. To this end, a Metadata-Assisted Glycan

Sequencing (MAGS) approach was used [32], where a number of structures bound by DC-SIGN were printed on a separate microarray and screened with lectins and antibodies that have defined binding to a variety of glycan determinants including α-fucosylated structures, terminal β1-3-linked or β1-4-linked galactose, α2-6-linked sialic acid, Lewis epitopes, and Blood Group H Type 1 or Type 2 (**Supplementary File 2**). DC-SIGN was also screened on this microarray and confirmed to bind all of the printed structures (**Supplementary File 2**).

Based on this MAGS data and mass spectrometry sequencing data for some structures [33], proposed structures for the HMGs bound by DC-SIGN are shown in **Figure 2**. The key feature of all these structures is the presence of α-linked fucose, specifically terminal Lewis a (Galβ1-3(Fucα1-4)GlcNAcβ-), terminal Lewis b (Fucα1-2Galβ1-3(Fucα1-4)GlcNAcβ-), terminal Lewis y (Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ-) and/or a terminal Lewis x (Galβ1-4(Fucα1-3)GlcNAcβ-) determinant. Not all fucosylated HMGs were bound though. For example, HMO-8 and HMO-29 contained one fucose while HMO-37 and HMO-80 contained 2 fucoses but were not bound. The fucosylated glycan determinant present in these four structures was likely Blood Group H Type 1 (H1) since the anti-H1 antibody but none of the Lewis antibodies bound these four structures. Additionally, HMO-23, -31, -41, -47, -48, and -49, containing 1 or 2 fucoses, were also not bound and contain only an internal Lewis x determinant (or, in the case of HMO-31, internal Lewis x as the major structures) [21, 28]. Therefore, the Lewis x is a binding determinant of DC-SIGN only when present at the non-reducing end of HMGs. The binding of DC-SIGN to HMGs containing terminal Lewis glycan determinants but not Blood Group H determinants on HMGs also corroborates previous studies on the

glycan specificity of DC-SIGN [34, 35]. Additionally, 2'-fucosyllactose (2'-FL, HMO-3) was also weakly bound on the HM-SGM-v2 (average rank = 11), a ligand not seen in previous studies. Overall, these results suggest that DC-SIGN recognizes α-fucosylated HMGs containing Lewis glycan determinants at the non-reducing end as well as 2'-fucosyllactose, and the high abundance of these structures and determinants in the HMGs metaglycome explains why DC-SIGN binds robustly to the HM-SGM-v2.

Siglec-5 bound weakly to four sialylated HMGs: HMO-157, HMO-213, HMO-118, and HMO-237 (**Fig. 1a**). However, this binding required a high Siglec-5 concentration of 90μg/ml and the signal was detectable but weak. While the structures of these four HMGs have not been completely defined, HMO-157, HMO-213, and HMO-237 were bound by GM35 monoclonal antibody [21], which we have shown to bind to the sialyl Lewis a (SLe^a) determinant (Neu5Acα2-3Galβ1-3(Fucα1-4)GlcNAcβ-) and so-called "sialyl Lewis c" (SLe^c) (Neu5Acα2-3Galβ1-3GlcNAcβ) determinant [36]. These data suggest that Siglec-5 binds to a restricted set of HMGs structures containing the SLe^c determinant, but the reason that Siglec-5 bound to only a restricted subset of all structures containing this determinant is unclear, since several other glycans on the array also were bound by GM35 but not Siglec-5. HMO-118 is likely to be a mixture that contains 3'-sialyllactose (3'-SL) based on its predicted composition; however, 3'-SL itself was not bound on the defined HMGs microarray by Siglec-5 (described in more detail below), suggesting trace glycans within HMO-118 may have contributed to binding. Siglec-9 bound an extensive number of HMGs, all of which are sialylated, although three of the four HMGs bound by Siglec-5 (HMO-157, HMO-213, and HMO-118) were consistently the strongest Siglec-9 binders as well (**Fig. 1b**). However, the

binding of Siglec-9 was only weakly dose-dependent (**Supplementary File 1**) and oddly depended on reducing end derivatization of the glycans, as discussed below. Thus, the results indicate that DC-SIGN robustly recognizes a number of α-fucosylated HMGs, whereas Siglec binding is weak and may not be significant. In regard to Siglec binding the significance was further tested below.

Binding of hGBPs to the CFG Glycan Microarray

To confirm that all hGBPs were active, they were concurrently screened on the microarray from the Consortium for Functional Glycomics (CFG). Most hGBP tested on the CFG microarray showed binding to at least two glycans on that microarray (**Supplementary File 3**). However, the CFG microarray data for Siglec-1 was deemed relatively inconclusive as no specific candidate glycans were identified. Many of the hGBPs that did not bind to the HM-SGM-v2 bound glycan determinants on the CFG microarray that were not found on the HM-SGM-v2, verifying hGBP activity. For example, Dectin-2 is specific for mannan structures containing the motif Manα1-2Manα1-6(Manα1-3)Manα- or Manα1-2Manα1-6Manα1-6(Manα1-3)Manα-, although Dectin-2 also weakly bound $Man_{8-9}GlcNAc_2$ N-glycan structures. A common feature of all structures bound by macrophage galactose/N-acetylgalactosamine-type lectin (MGL, CLEC10A) was the presence of GalNAc, particularly at the reducing and/or non-reducing end, although not all these GalNAc-containing structures were bound. Early studies suggested that Blood Group A and B antigens and enzymes with Blood Group A and B activity may be present in human milk (reviewed in [37]), but this has not been confirmed in more recent studies [38]. Indeed, we have screened the HM-SGM-v2 microarray with

an antibody recognizing Blood Group A Types 1-3 determinants and saw no binding of this antibody (data not shown), confirming more recent studies that Blood Group A (and most likely Blood Group B) determinants are not present at detectable amounts on HMGs. Thus, since mannose and GalNAc are not found on free HMGs, it is logical that Dectin-2 and MGL-1 did not bind the HMGs microarrays.

For the Siglecs, most showed a broad binding pattern on the CFG glycan microarray but no binding to the few HMGs present on the CFG microarray. Siglec-5 bound some but not all complex N-glycans containing terminal β1-3-linked galactose; beyond that, Siglec-5 did not bind to a common motif. In contrast to the HM-SGM-v2, we did not observe Siglec-5 binding to sialylated HMGs on the CFG microarray, including a lack of binding to 3'-SL and all the non-HMG glycans containing the sialyl Le^c determinant. Siglec-5 also did not bind a microarray consisting of defined HMGs structures (as described below), suggesting that Siglec-5 may not bind well HMGs and thus was possibly binding trace contaminants on the HM-SGM-v2 or only binds to only specific glycan presentations such as glycans with specific linkers. Siglec-7 bound to a variety of sialylated structures, the strongest of which was sialyl Lewis x containing 6-O-sulfated GlcNAc; some N-glycan structures and α2-8-sialylated structures were also bound. However, no motifs found on HMGs were bound by Siglec-7. Siglec-10 showed a very broad binding pattern, including binding to both sialylated and non-siaylated glycans. The biological and biochemical significance of the Siglec-10 binding to non-sialylated glycans is currently unclear, but we believe that the binding may have been artificially induced by the presentation and/or aglycone component (refer to the Discussion section for more information). Langerin not only strongly bound mannan and

high mannose N-glycan structures ($Man_{6-9}GlcNAc_2$) but also lactose that was 6-O-sulfated on the galactose, but neither determinant is found on HMGs. Additionally, Langerin bound weakly to glycans containing terminal β-linked GlcNAc, Blood Group H Types 1 and 2, Blood Group A and B Type 2, Lewis y, and other sulfated glycan determinants. These results are in good agreement with previous glycan microarray results for Langerin [39]. Although Type 1 Blood Group H and Lewis y are found on some HMGs, Langerin binding to the HMGs microarrays was inconclusive (**Supplementary File 1)**.

The screening of DC-SIGN on the CFG microarray revealed three major motifs (**Table 2** and **Supplementary File 3**). The first motif was terminal α1-2-linked mannose on mannan backbones, including high-mannose N-glycans, although the mannans containing α-linked mannose at the reducing end were bound slightly stronger than the high mannose N-glycans. The second motif was the Lewis a determinant, including Lewis b structures. The third motif was the Lewis x determinant at the non-reducing end of glycan structures, which also included Lewis y structures. Notably, the Lewis a and non-reducing end Lewis x determinants were also the major HMG binding determinants revealed by the HM-SGM-v2 screening (Table 2). Sialyl Lewis a and especially sialyl Lewis x structures were typically poorly bound by DC-SIGN, although some sialylated, fucosylated HMGs were bound on the HM-SGM-v2 whose structures remain to be determined (**Fig. 1** and **Supplementary File 1**). Glycans containing Blood Group H Type 1 and Type 2 determinants, as well as 2'-FL, were poorly bound by DC-SIGN on the CFG microarray, a finding also seen in previous studies [35]. This further suggests that α1-2 fucosylated glycan structures are lower affinity than Lewis a- and Lewis x-

containing structures. Binding motifs for Siglec-9 on the CFG microarray were sialyl Lewis x on N-glycans as well as 3'- or 6'-sialyllactosamine (Neu5Acα2-3/6Galβ1-4GlcNAcβ-) that was 6-O-sulfated on the GlcNAc, but binding was slightly stronger to 6-O-sulfo-sialyl Lewis x as previously seen on this array (refer to the CFG website, http://www.functionalglycomics.org/glycomics/publicdata/primaryscreen.jsp); importantly, these motifs are not found on HMGs. Siglec-9 also weakly bound 3'-sialyllactosamine and 6'-siayllactosamine as well as the sialyl Lewis x tetrasaccharide (Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAc), although the HMGs 3'-SL and 6'-sialyllactose (6'-SL) were poorly if at all bound. Siglec-9 binding to the CFG microarray was also poorly dose-dependent, as seen when Siglec-9 was screened on the defined HMGs microarray, suggesting that the binding may not be specific. Therefore, only DC-SIGN was concluded to be a strong HMG receptor, while Siglec-5, Siglec-9, and Langerin are likely poor HMGs receptors.

Binding of hGBPs to a Defined HMG Microarray

To further investigate the binding of hGBP to HMGs, the hGBP were also screened on a microarray consisting of a selection of chemically defined HMGs-related glycans and galactooligosaccharides (GOS) that are commonly used or under experimental testing as supplements in infant formula. This microarray was termed the "defined HMG microarray". As expected from the HM-SGM-v2 screenings, DC-SIGN and Siglec-9 bound structures on the defined HMG microarray (**Fig. 3**) while all other hGBP showed no binding (refer to **Supplementary File 4** for the defined HMG glycan microarray data at all concentrations of all hGBP screened). In contrast to the HM-SGM-

v2 data, Siglec-5 at 90μg/ml did not bind to the defined HMG microarray even though HMO-118 (likely 3'-SL) was bound by Siglec-5 on the HM-SGM-v2. This suggests that Siglec-5 binds 3'-SL with low affinity and HMO-118 may contain trace contaminants that improved Siglec-5 binding. Siglec-9 bound to both 3'-SL and 6'-SL.

DC-SIGN bound the fucosylated HMGs 2'-fucosyllactose (2'-FL) and 3-fucosyllactose (3-FL), while very weak binding was seen towards the Blood Group H Type 1-containing glycan LNFPI. It should be noted that 2'-FL (HMO-3) but not 3-FL (HMO-2) was bound on the HM-SGM-v2, although 2'-FL was weakly bound relative to Lewis a and Lewis x structures (average rank = 11; **Supplementary Figure 1**). However, DC-SIGN binding to 3-FL was highly dependent on maintaining the ring-structure of the reducing end glucose because DC-SIGN poorly bound to reductively aminated 3-FL (**Fig. 3** and **Supplementary File 4**), which was the only ring-form of glycans on the HM-SGM-v2. Thus, the actual strength of binding to 3-FL was likely underestimated on the HM-SGM-v2. In contrast to the HMG microarray results, 2'-FL was not bound by DC-SIGN on the CFG glycan microarray (Chart ID 77, rank < 10; **Supplementary Figure 3**). The reason for this non-binding on the CFG glycan microarray, but may have to do with differences in the linker or other presentation issues vs. the HMG microarrays. 3-FL was absent from the CFG glycan microarray. This suggests that 2'-FL and 3-FL are weaker ligands than structures containing terminal Lewis a or Lewis x determinants, although more studies are needed to confirm this observation. Overall, these results suggest that 2'-FL and 3-FL are also ligands for DC-SIGN. Despite their potentially lower binding strength than Lewis a and Lewis x-containing HMGs, 2'-FL and 3-FL are much more abundant than these Lewis a- and Lewis x-containing HMG structures in

human milk, with concentrations ranging from about 0.5-5mM for these two HMGs [2]. The overall results of this experiment show that Siglec-9 and DC-SIGN, but not Siglec-5, may also bind to simple, defined HMGs structures.

Binding of hGBPs to the Beads Derivatized with Human Milk Glycans

To confirm the binding of DC-SIGN to the defined HMGs 2'-FL and 3-FL in a different format, polystyrene beads were derivatized with 2'-FL, 3-FL, or LNT and binding of DC-SIGN to these derivatized beads was measured by flow cytometry (**Fig. 4**). As expected, DC-SIGN bound to the 2'-FL- and 3-FL-derivatized beads but not the LNT-derivatized beads (non-fucosylated HMGs control), which further confirmed the binding of DC-SIGN to fucosylated HMGs.

HMG Inhibition of hGBPs Binding

Experiments using glycan microarrays and beads are useful for defining glycan specificity and potential binding of hGBP to HMGs. However, the glycan microarray screenings themselves have a few important limitations. Specifically, the glycans on the microarray are synthetically derivatized with a bifunctional linker at the reducing end and presented in a solid-phase format, which is in contrast to HMGs that occur as free, reducing glycans in human milk. To confirm that DC-SIGN and Siglec-9 can also bind to free, underivatized HMGs in solution, DC-SIGN and Siglec-9 were screened on the defined HMGs microarray in the presence or absence of various concentrations of 2'-FL and 6'-SL, respectively; lactose was used as a negative control for non-specific HMGs inhibition. DC-SIGN binding to both the defined HMGs microarray (**Fig. 5a**) and the

MAGS array (**Fig. 5b**) was inhibited by 2'-FL (refer to **Supplementary File 5** for the total data for DC-SIGN inhibition) in a dose-dependent manner and with an approximate IC_{50} of 1mM for 2'-FL, confirming that DC-SIGN specifically binds to natural 2'-FL and in solution. Lactose (1-10mM) caused little or no inhibition of DC-SIGN binding to the defined HMGs microarray (**Supplementary File 5**), confirming that the presence of α-linked fucose is required for DC-SIGN binding. The data also confirm binding to all of the HMGs on the HM-SGM-v2 and defined HMGs microarrays was specific.

On the other hand, Siglec-9 binding to the defined HMGs microarray was not inhibited by even 10mM 6'-SL (**Fig. 6a**), although binding could be inhibited by 1mM 6'-SL derivatized with the AEAB linker at the reducing end (**Fig. 6b**; also see **Supplementary File 6** for the total data for Siglec-9 inhibition). Therefore, Siglec-9 did not appear to bind the natural form of 6'-SL (and likely 3'-SL), only the chemically derivatized version; this suggests that Siglec-9 binding to the defined HMGs microarray only occurs because of this HMGs derivatization. The solution K_d of Siglec-9 for free 6'-SL and 3'-SL was determined to be >10mM, which is likely not physiologically relevant.

Discussion

A major finding of this study is that DC-SIGN is the only hGBP tested that showed specific binding to HMGs and binding was most robust toward α-fucosylated glycans. A striking observation was the proportion of HMGs bound by DC-SIGN. About half of the HMG structures on the HM-SGM-v2 were bound by 10μg/ml DC-SIGN (**Fig. 1, Supplementary File 1**), suggesting that DC-SIGN binds to nearly half of the structures in the HMG metaglycome. The strongest binding was towards HMGs containing a Lewis glycan determinant at the non-reducing end (**Figs. 1, 2, 3, Supplementary Files 1, 2, and 3**). Potentially weaker but likely physiologically significant binding of DC-SIGN to 2'-fucosyllactose and 3-fucosyllactose was also observed (**Figs. 1, 3, 5, Supplementary Files 1, 4, 5**). HMGs containing only internal Lewis x or Blood Group H Type 1 were poorly, if at all, bound by DC-SIGN. Therefore, DC-SIGN appears to be a receptor for specific fucosylated HMGs.

The approximate IC_{50} of DC-SIGN for 2'-FL inhibition of binding to the glycan microarray was 1mM (**Fig. 5**). Given the typical concentration of 1-5mM (0.5-2.5g/L) 2'-FL in Secretor-positive human milk [2], this suggests that the binding is within the physiological range. Taking into account that DC-SIGN also binds half of the total HMGs metaglycome in secretor- and Lewis-positive individuals (**Fig. 1**) as well as glycoproteins in human milk such as bile salt-stimulated lipase [40] and MUC1 [17], the actual concentration of DC-SIGN glycan determinants in human milk is probably much higher (~5-10mM), suggesting that DC-SIGN may be close to ligand saturation when exposed to human milk (assuming an average K_d of 1mM). Total HMGs have been

previously shown to block DC-SIGN binding to HIV virions [41], further suggesting that some HMGs are DC-SIGN ligands and can block DC-SIGN functions.

Human intestinal dendritic cells express DC-SIGN [42], and DC-SIGN expression is known to occur on cells (likely dendritic cells) in infant GI tract tissue [17]. DC can extend their dendrites from the lamina propria into the intestinal lumen to "sample" microbes [18]. Since HMGs are not significantly digested by the human repertoire of digestive mechanisms and enzymes in the GI tract [43, 44], DC-SIGN on dendritic cells may be exposed to and bind HMGs to near saturable levels in the small intestine of breast-fed infants. DC-SIGN is also known to modulate immune responses, though this binding is not yet known to be a direct stimulator of gene expression [10]. However, it is possible that the interaction of DC-SIGN with HMGs may cause changes in the DC-SIGN-mediated modulation of immune responses and may also help mechanistically explain how HMGs promote changes in gene expression and immune responses [6]. Notwithstanding, how such interactions occur and if the HMGs act as agonists or antagonists of DC-SIGN activity is still not fully understood. Interestingly, about 20% of individuals lack the Secretor enzyme responsible for producing α1-2-fucosylated HMGs and 2'-FL, about 10% of individuals lack the Lewis enzyme responsible for producing α1-4-fucosylated HMGs (Lewis a structures), and ~1% of individuals lack both enzymes [45]. Thus, milk from secretor-negative and/or Lewis-negative individuals may not be capable or interacting as well with DC-SIGN as milk from Secretor- and Lewis-positive mothers, although this might be at least partially compensated by the increased 3-FL concentration in non-secretor vs. secretor human milk as 3-FL is also a DC-SIGN ligand

(Figure 3, 5, Supplementary File 3, 5). The physiological consequence of lacking the Secretor and/or Lewis enzyme on DC-SIGN binding *in vivo* are thus unclear.

Unexpectedly, given the fact that sialic acid is a common residue in HMGs, the only Siglecs tested that showed some binding to HMGs were Siglecs-5 and -9 (**Fig. 1**), consistent with the possibility that glycan recognition by Siglecs is complex and the presence of sialic acid is necessary but not sufficient in most cases. Siglec-5 binding was weak and only occurred at high Siglec-5 concentrations, while Siglec-9 binding was stronger but binding to the free, underivatized HMGs 6'-SL was still weak (**Fig. 6**). Instead, Siglec-9 bound strongly to 6'-SL derivatized at the reducing end with an aglycone linker, AEAB and especially GGAEAB. This finding suggests that the aglycone component and/or multivalent presentation may be an important factor in Siglec-9 and other Siglecs for binding glycoconjugate ligands, or that specific sialylated glycans yet to be identified are strong ligands for Siglecs. This finding of the potential importance the aglycone in Siglec binding may also explain why the binding of Siglecs to the CFG microarray in this study has a generally weak, broad binding pattern. This result may be due to differences in glycan presentation, which may have positively or negatively affected by the presence of specific aglycone linker units. Thus, the weak, broad binding pattern of Siglecs to the CFG microarrays was likely because of non-preferential glycan presentation and/or aglycone components as opposed to poor Siglec activity or the recombinant Siglec construct used. Future studies in our lab are aimed at understanding the functional importance of aglycone components, especially natural aglycone components such as lipids and peptides, in Siglec binding. This future study may also unravel why some Siglecs, especially Siglec-10, bind to a few non-sialylated glycans.

The glycan presentation in multivalent forms may be most important for Siglec binding, as prior studies showed that Siglec-1, -3, -5, -7, and -9 all bound 3'-SL and 6'SL-derviatized beads with μM affinity constants by surface plasmon resonance [15]. The multivalent presentation and/or aglycone bead component may contribute to this strong binding, since an IC_{50} of ~1mM was calculated for free 6'-SL inhibition [15], which was ~100-1000-fold higher than for the K_d for 6'-SL beads. Siglec-5 binds 3'-SL and 6'-SL with a K_d of 2-4μM but to free 3'-SL and 6'-SL with a K_d of ~8mM [46], which is high relative to the concentrations of these two sialylated glycans in human milk. Based on these current and previous findings, we conclude that the binding of dendritic cell-expressed Siglecs to free, underivatized HMGs is weak and likely non-physiological.. We speculate that this low affinity binding is due to the lack of an aglycone component on and/or multivalent presentation of HMGs, which normally exist as free, reducing glycan structures in solution. These findings also stress the importance of using other methodologies besides glycan microarrays to confirm binding of samples to HMGs, which naturally exist as free, underivatized structures in solution.

In addition to C-type lectins and Siglecs used in this study, other hGBP have been screened on the HMGs microarrays. These include galectins, most of which showed binding and, in some cases, robust binding to neutral HMGs [22]. Preliminary HMGs microarray screenings of the three human selectins (P-, E-, and L-selectin), which are known to bind sialyl Lewis x and sialyl Lewis a determinants in solution with relatively low affinity [47, 48], were negative. This suggests that human selectins are poor HMGs receptors, consistent with previous studies showing that, although selectins may bind HMGs, the interaction and effects are weak [49-51]. Preliminary screenings with Siglec-

11 also revealed no binding of these hGBP to the HMGs microarrays. Future studies are aimed at examining other receptors that bind glycoconjugates, including Toll-like receptors and cytokine receptors.

This study adds DC-SIGN to the list of hGBP that may act as HMG receptors. Given the physiological concentration of HMGs binding to DC-SIGN and galectins, as well as the anatomical localization and expression patterns of these hGBP, these interactions may be important mechanisms underlying the known HMGs functions of regulating gene expression and immune responses [6]. Therefore, future studies to understand the interactions of these hGBP with HMGs and subsequent physiological effects are currently underway.

Acknowledgements

We thank Jamie Heimburg-Molinaro for critical review of the manuscript and Sandra Cummings and Hong Ju for technical assistance.

Declarations of Interest

R.D.C. and D.F.S. are consultants for Abbott Nutrition. G.D.M. and R.H.B. are employees of Abbott. The other authors declare that they have no conflict of interest in the work reported.

Funding Information

This work was support by NIH Grants P41GM103694 to RDC and a Grant from Abbott Nutrition, Columbus, OH.

Author Contribution Statement

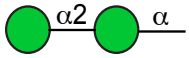
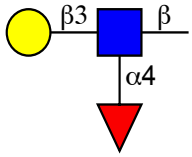
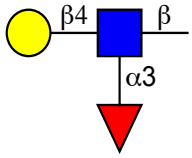
AJN, YY, GDM, RHB, DFS, and RDC proposed and designed experiments. YY performed experiments. AJN, YY, DFS, and RDC analyzed data. YL, GDM, and RHB provided critical reagents and support. AJN and RDC organized data and wrote manuscript with review, comments, and contributions from all authors.

Table 1. Recombinant Human Glycan-Binding Proteins Used in this Study

Name	Lectin Family	Genbank Accession Number	Amino Acid Sequence	Protein Fusion Tag	R & D Catalog Number
DC-SIGN	C-Type Lectin	Q9NNX6	Lys62 - Ala404	N-terminal MD-Human IgG_1 F_c-IEGR fusion tag	161-DC-050
Langerin	C-Type Lectin	Q9UJ71	Tyr64 - Pro328	N-terminal 9x His tag	2088-LN-050
Dectin-2	C-Type Lectin	Q6EIG7	Thr46 - Leu209	N-terminal 6x His	3114-DC-050
MGL (CLEC10A)	C-Type Lectin	Q8IUN9	Gln61 - His316	N-terminal 6x His	4888-CL-050
Siglec-1	I-Type Lectin	Q9BZZ2	Ser20 - Gln1641	C-terminal 6x His tag	5197-SL-050
Siglec-5	I-Type Lectin	O15389	Glu17 - Thr434	C-terminal IEGRID-Human IgG_1 F_c-fusion tag	1072-SL-050
Siglec-7	I-Type Lectin	Q9Y286	Gln19 - Gly357	C-terminal DIEGRMD-Human IgG_1 F_c fusion tag	1138-SL-050
Siglec-9	I-Type Lectin	Q9Y336	Gln18 - Gly348	C-terminal DIEGRMD-Human IgG_1 F_c fusion tag	1139-SL-050
Siglec-10	I-Type Lectin	Q96LC7	Met17 - Thr546	C-terminal IEGRMD-Human IgG_1 F_c fusion tag	2130-SL-050

[a]All recombinant proteins were purchased from R&D Biosystems and were expressed from a mouse myeloma cell line, NS0-derived

Table 2. DC-SIGN CFG Glycan Microarray Binding Motifs.

Glycan Motif	Glycan Determinant Present
	α1-2 Mannose
	Lewis a
	Lewis x

Figures

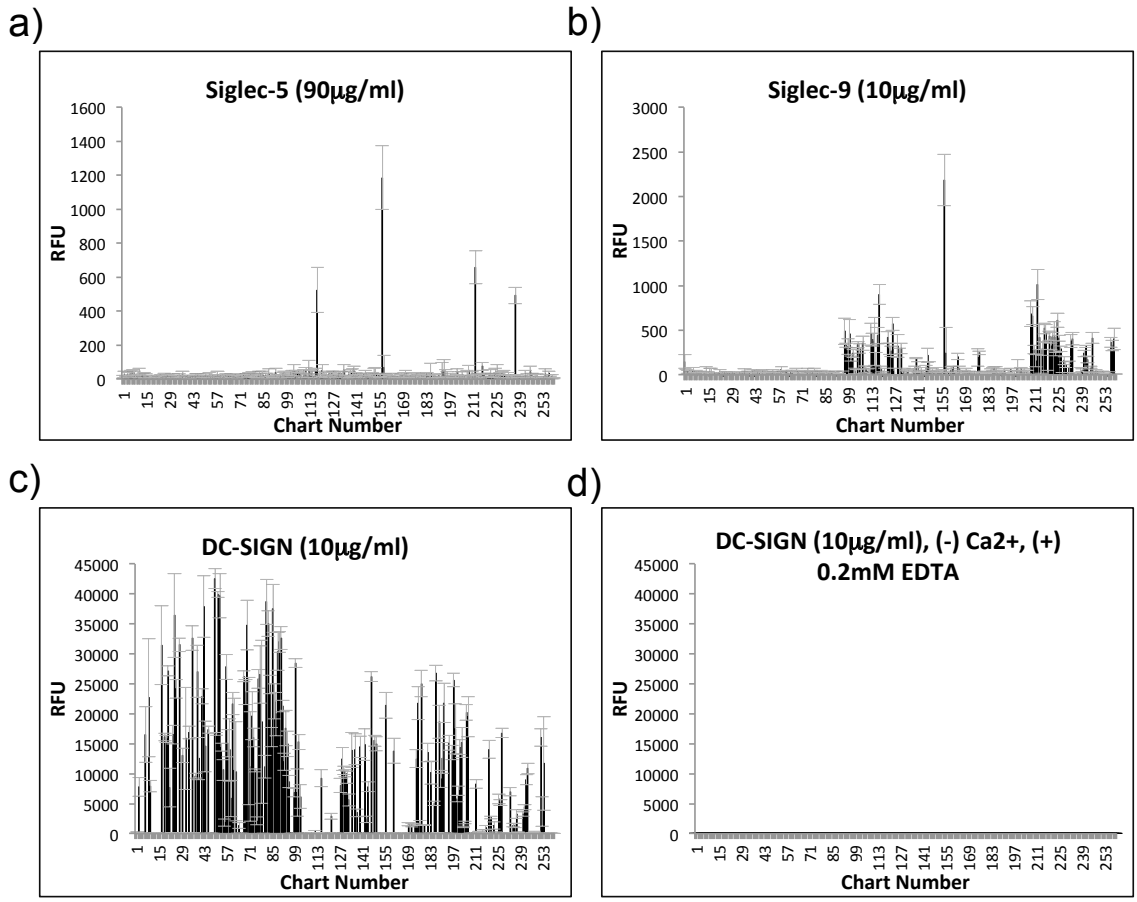


Fig. 1. HM-SGM-v2 Data for Siglec-5, Siglec-9, and DC-SIGN.

Siglec-5, Siglec-9, and DC-SIGN were screened on the HM-SGM-v2 at multiple concentrations. The results for 90μg/ml Siglec-5 (**a**), 10μg/ml Siglec-9 (**b**), and 10μg/ml DC-SIGN with (**c**) or without (**d**) Ca^{2+} are shown; Alexa Fluor 488-labeled anti-human IgG was used for detection. Refer to **Supplementary File 1** for the results at all concentrations screened. For DC-SIGN without Ca^{2+} (**d**), Ca^{2+} was omitted from the binding buffer and 0.2mM EDTA was added.

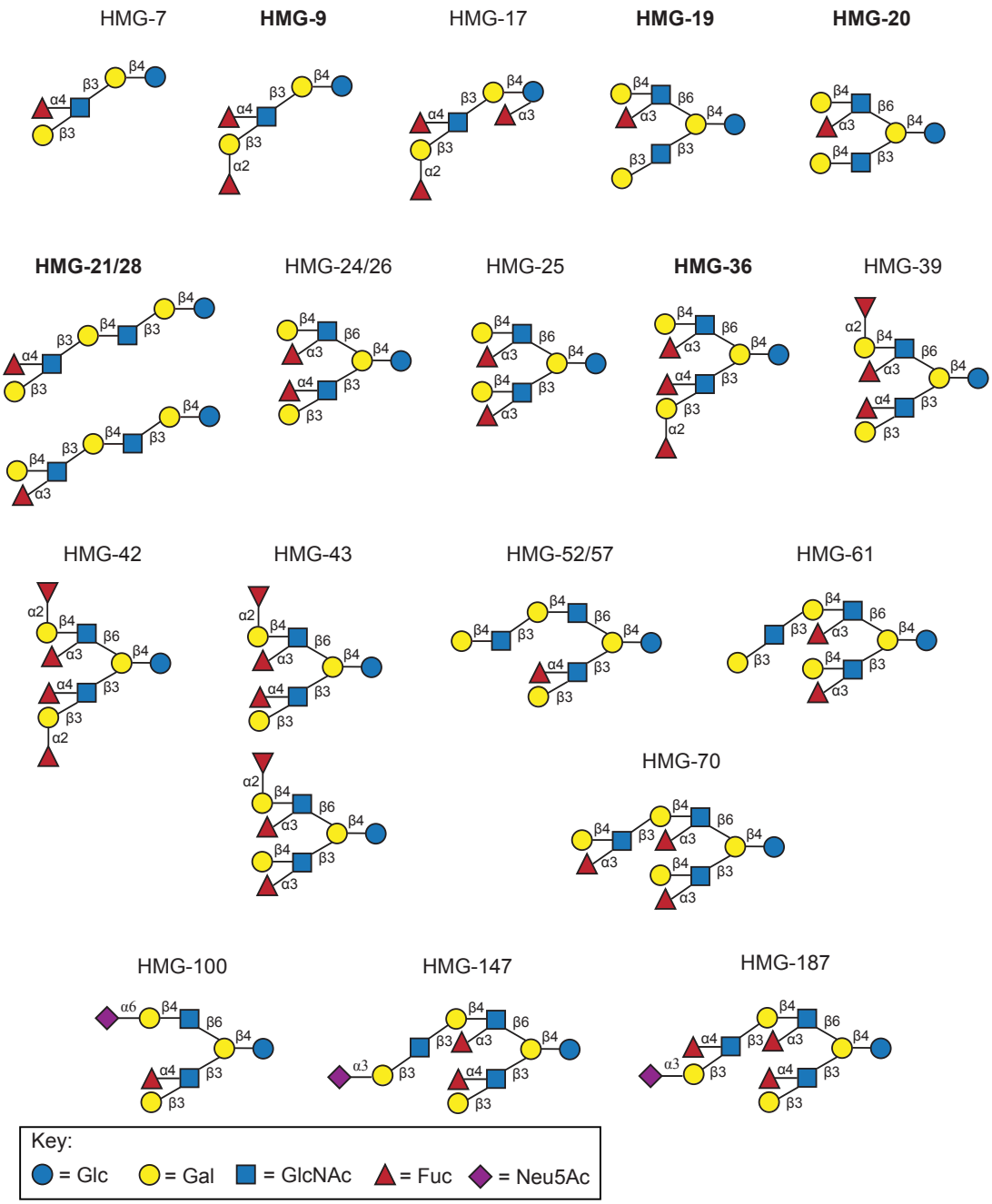

HMG-7
HMG-9
HMG-17
HMG-19
HMG-20
HMG-21/28
HMG-24/26
HMG-25
HMG-36
HMG-39
HMG-42
HMG-43
HMG-52/57
HMG-61
HMG-70
HMG-100
HMG-147
HMG-187
Key:
= Glc
= Gal
= GlcNAc
= Fuc
= Neu5Ac

Fig. 2. Proposed Structures of HMGs Bound by DC-SIGN on the HM-SGM-v2.

A portion of the HM-SGM-v2 structures bound by DC-SIGN were printed on a separate microarray and interrogated by metadata-assisted glycan sequencing (MAGS), where multiple lectins and antibodies specific for particular glycan determinants were screened. Proposed structures for these HMGs samples are shown. HMG samples in bold-face font were further analyzed by multi-dimensional mass spectrometry (MS^n) to more accurately determine the structures(s) within these samples; HMG-20, -21, and -28 were previously sequenced by MS^n [21, 28] and HMGs-9, -19, and -36 were also by sequenced by MS^n in a more recent manuscript [33]. Refer to **Supplementary File 2** for the lectin and antibody screening data.

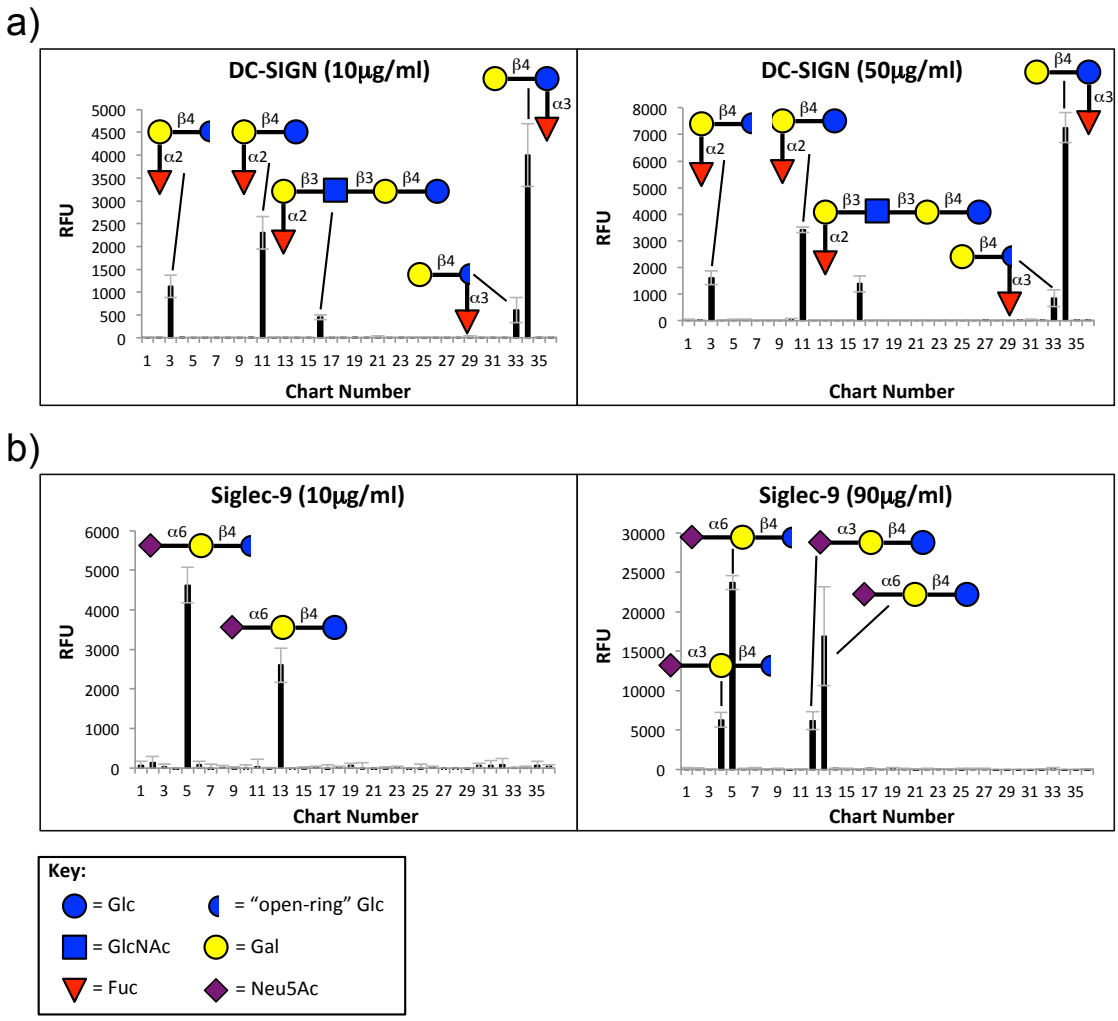


Fig. 3. Defined HMGs Microarray Screening Data for DC-SIGN and Siglec-9.

10µg/ml and 50µg/ml DC-SIGN (**a**) and 10µg/ml and 90µg/ml Siglec-9 (**b**) were screened on the defined HMGs microarray, and Alexa Fluor 488-labeled anti-human IgG was used for detection. Graphs on the left show the lower concentrations and the right graphs show the higher concentrations. Refer to **Supplementary File 4** for the raw data for these screenings.

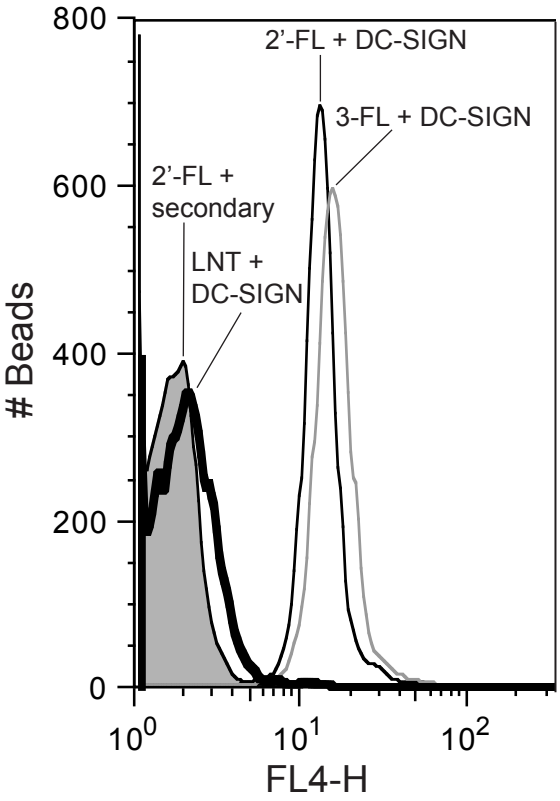


Fig. 4. DC-SIGN Binding to HMGs-Derivatized Microspheres.

DC-SIGN was incubated with microspheres (beads) derivatized with 2'-FL, 3-FL, or LNT. Alexa Fluor 633-labeled anti-human IgG was used for detection of recombinant DC-SIGN. 2'-FL microspheres incubated with the Alexa Fluor 633-labeled anti-human IgG alone was used as the negative control. All samples were analyzed by Flow Cytometry with a 633nm laser and FL-4 filter for detection. Histograms of DC-SIGN binding to LNT beads (thick line), 2'-FL beads (thin black line), and 3-FL beads (thin grey line), as well as secondary antibody alone binding to 2'-FL beads (filled line), are shown.

a) Defined HMG Microarray b) HMG MAGS Microarray

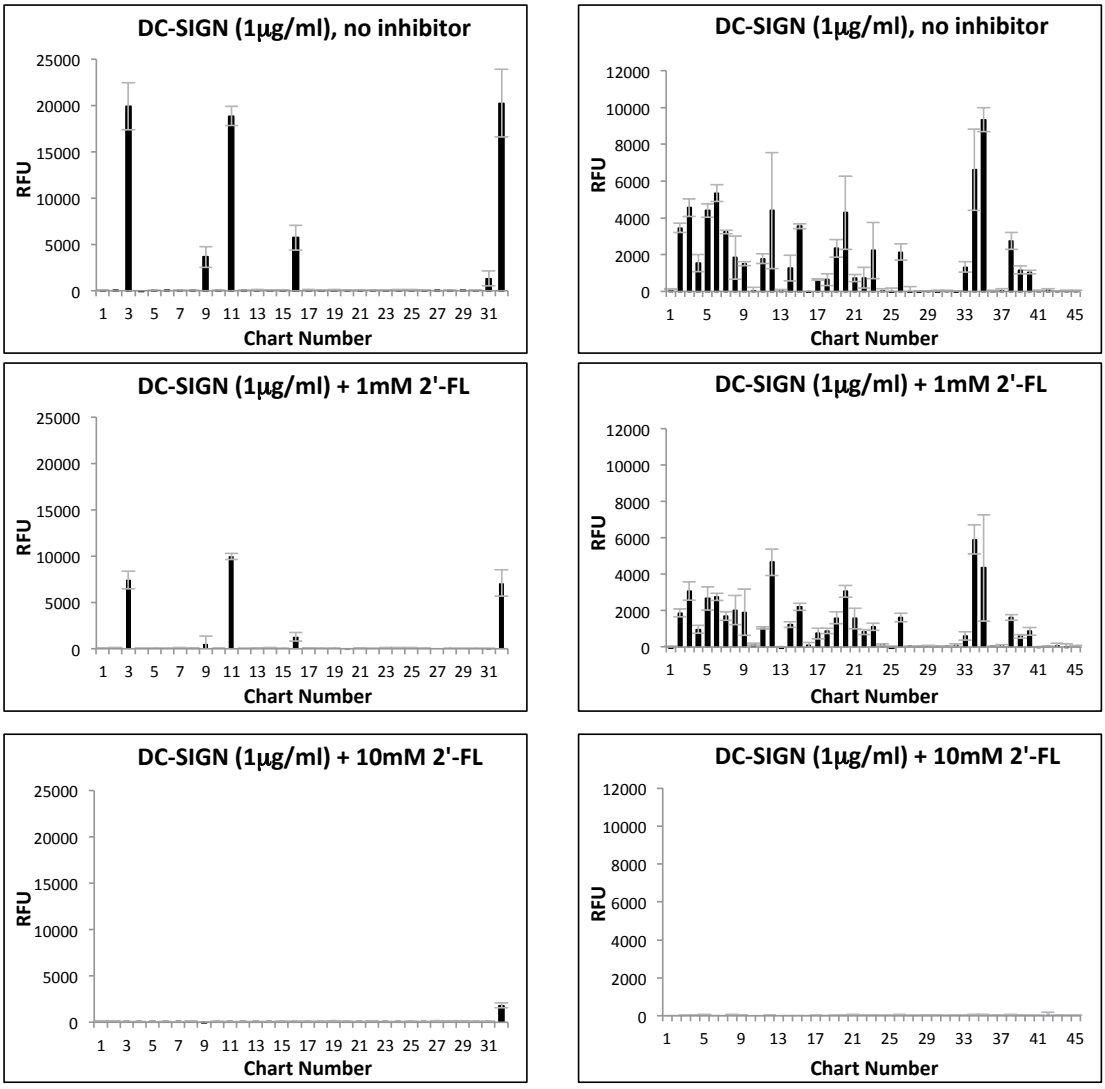
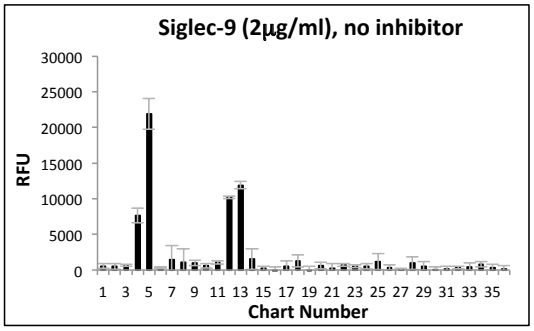
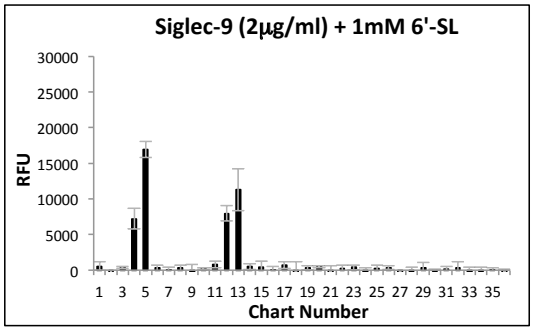


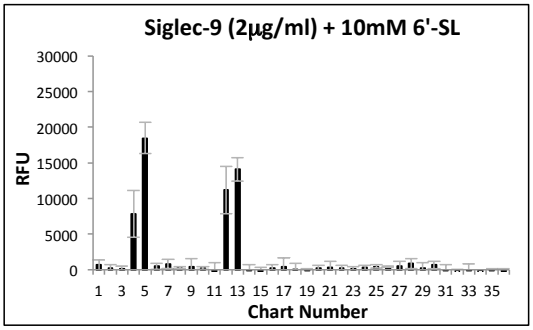
Fig. 5. Inhibition of DC-SIGN Binding to HMGs Microarrays with Free, Underivatized HMGs.

1μg/ml DC-SIGN was preincubated with or without 0.1, 1, or 10mM of free, underivatized 2'-FL and then screened on the defined HMGs microarray (**a**) or the HMGs MAGS microarray (**b**) described in **Supplementary File 2**. The results for the no inhibitor, 1mM 2'-FL, and 10mM 2'-FL on both microarrays are shown; refer to **Supplementary File 5** for the results of all other screenings, including 0.1mM 2'-FL.

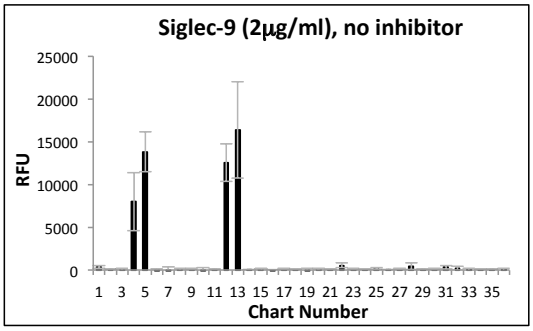
a)

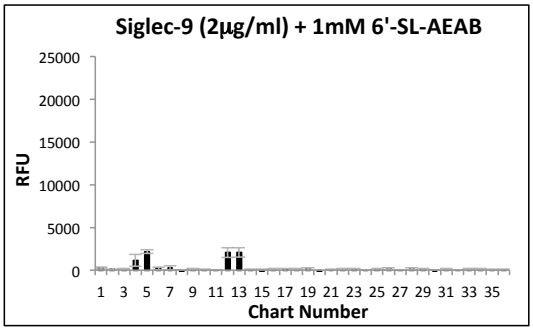

Siglec-9 (2μg/ml), no inhibitor
RFU
Chart Number


Siglec-9 (2μg/ml) + 1mM 6'-SL
RFU
Chart Number


Siglec-9 (2μg/ml) + 10mM 6'-SL
RFU
Chart Number

b)


Siglec-9 (2μg/ml), no inhibitor
RFU
Chart Number


Siglec-9 (2μg/ml) + 1mM 6'-SL-AEAB
RFU
Chart Number

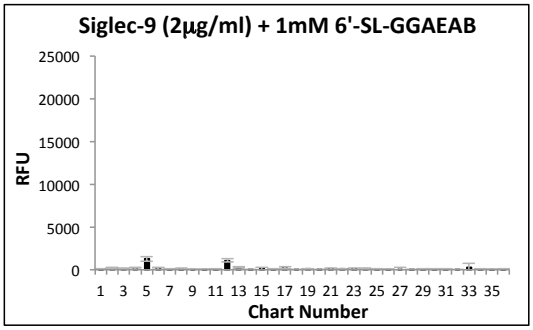

Siglec-9 (2μg/ml) + 1mM 6'-SL-GGAEAB
RFU
Chart Number

Fig. 6. Inhibition of Siglec-9 Binding to Defined HMGs Microarray with Free, Underivatized HMGs and Free, Derivatized 6'-Sialyllactose.

2μg/ml Siglec-9 was preincubated with 1mM or 10mM free, underivatized 6'-SL or no inhibitor and screened on the defined HMGs microarray (**a**). 2μg/ml Siglec-9 was preincubated with free 1mM 6'-SL-AEAB (Neu5Acα2-6Galβ1-4Glcitol-AEAB) or 6'-SL-GGAEAB (Neu5Acα2-6Galβ1-4Glc-GGAEAB) [25] or no inhibitor and screened on the defined HMGs microarray (**b**). Parts **a** and **b** of this figure were performed on separate slides but on the same day and at the same time. Refer to **Supplementary File 6** for all other free inhibition of Siglec-9 binding to the defined HMGs microarray.

Supplementary Information

Supplementary Files (.xlsx files) are accessible from the publisher's website and include the following six files:

Supplementary File 1. hGBP HM-SGM-v2 Data

This file, containing seven worksheets, shows the average RFU data as well as graphs of binding data for 100μg/ml Langerin, 100μg/ml MGL-1, 90μg/ml Siglec-1, 90μg/ml Siglec-5, 90μg/ml Siglec-7, and 10μg/ml and 50μg/ml Siglec-9 binding to the HM-SGM-v2 on separate worksheets. Additionally, the data for 0.1μg/ml, 1.0μg/ml, and 10μg/ml DC-SIGN, as well as DC-SIGN screened in a buffer lacking Ca^{2+} as well as 0.2mM EDTA as a control for non-specific interactions, is included on a single worksheet. The data for Dectin-2, which showed no binding to the HM-SGM-v2, is not shown.

Supplementary File 2. MAGS Data and DC-SIGN Binding to MAGS Array

This file includes two worksheets. The first worksheet contains a summary of the average RFU results for the screening of anti-glycan antibodies and lectins on the MAGS array to identify glycan determinants present on the HMGs samples printed. The RFU's in red and black text indicate positive and negative binding, respectively, by the lectin or antibody. The second worksheet includes the average RFU data as well as graphs of binding data for 0.5μg/ml, 1.0μg/ml, 2.0μg/ml, and 5.0μg/ml DC-SIGN to the MAGS array. Additionally, the results for 1.0μg/ml DC-SIGN binding after the array had been treated with neuraminidase are also shown.

Supplementary File 3. hGBP CFG Microarray Data

This eight-worksheet file contains the average RFU data and graphs of binding data for 10μg/ml DC-SIGN, 10μg/ml Langerin, 50μg/ml Dectin-2, 10μg/ml MGL-1, 20μg/ml Siglec-1, 20μg/ml Siglec-5, 90μg/ml Siglec-7, and 10μg/ml and 90μg/ml Siglec-9 on the glycan microarray from the Consortium for Functional Glycomics (CFG).

Supplementary File 4. DC-SIGN and Siglec-9 Defined HMGs Microarray Data

This file contains two worksheets. The worksheets contain the average RFU binding data and graphs of binding data for 10μg/ml and 50μg/ml DC-SIGN, and 10μg/ml and 90μg/ml Siglec-9, respectively, on the defined HMGs microarray. All of the other hGBP screened showed no binding to the microarray (data not shown).

Supplementary File 5. Free HMGs Inhibition of DC-SIGN Binding to Defined HMGs and MAGS Microarrays

This file contains two worksheets. The first worksheet includes the average RFU data and graphs of binding data of 1.0μg/ml DC-SIGN screened on the defined HMGs microarray in the presence or absence of 0.1mM, 1.0mM, or 10mM 2'-FL or lactose. The second worksheet includes the average RFU data and graphs of binding data of 1.0μg/ml DC-SIGN screened on the MAGS array in the presence or absence of 1.0mM or 10mM 2'-FL.

Supplementary File 6. Siglec-9 Free HMGs Inhibition Data

This file, containing two worksheets, shows the results of 2.0μg/ml Siglec-9 binding to the defined HMGs microarray in the presence of free HMGs or HMGs derivatives including 1mM and 10mM 6'-SL, 1mM and 10mM lactose, 1mM 6'-SL-AEAB, and 1mM 6'-SL-GGAEAB. The two worksheets show the data for two different slides, although these two slides were concurrently screened in the same experiment.

References

1. Kunz, C., et al., *Oligosaccharides in human milk: structural, functional, and metabolic aspects.* Annu Rev Nutr, 2000. **20**: p. 699-722.
2. Urashima, T., et al., *The predominance of type I oligosaccharides is a feature specific to human breast milk.* Adv Nutr, 2012. **3**(3): p. 473S-82S.
3. Newburg, D.S., *Glycobiology of human milk.* Biochemistry (Mosc), 2013. **78**(7): p. 771-85.
4. Newburg, D.S., G.M. Ruiz-Palacios, and A.L. Morrow, *Human milk glycans protect infants against enteric pathogens.* Annu Rev Nutr, 2005. **25**: p. 37-58.
5. Zivkovic, A.M., et al., *Human milk glycobiome and its impact on the infant gastrointestinal microbiota.* Proc Natl Acad Sci U S A, 2011. **108 Suppl 1**: p. 4653-8.
6. Bode, L., *Human milk oligosaccharides: every baby needs a sugar mama.* Glycobiology, 2012. **22**(9): p. 1147-62.
7. Castillo-Courtade, L., et al., *Attenuation of food allergy symptoms following treatment with human milk oligosaccharides in a mouse model.* Allergy, 2015. **70**(9): p. 1091-102.
8. Bienenstock, J., et al., *Fucosylated but not sialylated milk oligosaccharides diminish colon motor contractions.* PLoS One, 2013. **8**(10): p. e76236.
9. Vazquez, E., et al., *Effects of a human milk oligosaccharide, 2'-fucosyllactose, on hippocampal long-term potentiation and learning capabilities in rodents.* J Nutr Biochem, 2015. **26**(5): p. 455-65.

10. Geijtenbeek, T.B. and S.I. Gringhuis, *Signalling through C-type lectin receptors: shaping immune responses.* Nat Rev Immunol, 2009. **9**(7): p. 465-79.

11. Crocker, P.R., J.C. Paulson, and A. Varki, *Siglecs and their roles in the immune system.* Nat Rev Immunol, 2007. **7**(4): p. 255-66.

12. McEver, R.P., *Adhesive interactions of leukocytes, platelets, and the vessel wall during hemostasis and inflammation.* Thromb Haemost, 2001. **86**(3): p. 746-56.

13. Song, X., et al., *Novel fluorescent glycan microarray strategy reveals ligands for galectins.* Chem Biol, 2009. **16**(1): p. 36-47.

14. Guo, Y., et al., *Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR.* Nat Struct Mol Biol, 2004. **11**(7): p. 591-8.

15. Zou, Z., et al., *Siglecs facilitate HIV-1 infection of macrophages through adhesion with viral sialic acids.* PLoS One, 2011. **6**(9): p. e24559.

16. Brinkman-Van der Linden, E.C. and A. Varki, *New aspects of siglec binding specificities, including the significance of fucosylation and of the sialyl-Tn epitope. Sialic acid-binding immunoglobulin superfamily lectins.* J Biol Chem, 2000. **275**(12): p. 8625-32.

17. Koning, N., et al., *Human Milk Blocks DC-SIGN-Pathogen Interaction via MUC1.* Front Immunol, 2015. **6**: p. 112.

18. Rescigno, M., et al., *Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria.* Nat Immunol, 2001. **2**(4): p. 361-7.

19. Cummings, R.D. and R.P. McEver, *C-type Lectins*, in *Essentials of Glycobiology*, A. Varki, et al., Editors. 2009: Cold Spring Harbor (NY).

20. Varki, A. and P.R. Crocker, *I-type Lectins*, in *Essentials of Glycobiology*, A. Varki, et al., Editors. 2009: Cold Spring Harbor (NY).

21. Yu, Y., et al., *Human milk contains novel glycans that are potential decoy receptors for neonatal rotaviruses.* Mol Cell Proteomics, 2014. **13**(11): p. 2944-2960.

22. Noll, A.J., et al., *Galectins are human milk glycan receptors.* Glycobiology, 2016. **In press.**

23. Heimburg-Molinaro, J., et al., *Preparation and analysis of glycan microarrays.* Curr Protoc Protein Sci, 2011. **64**(12.10): p. 12.10.1-12.10.29.

24. Song, X., et al., *Derivatization of free natural glycans for incorporation onto glycan arrays: derivatizing glycans on the microscale for microarray and other applications (ms# CP-10-0194).* Curr Protoc Chem Biol, 2011. **3**(2): p. 53-63.

25. Song, X., et al., *Fluorescent glycosylamides produced by microscale derivatization of free glycans for natural glycan microarrays.* ACS Chem Biol, 2009. **4**(9): p. 741-50.

26. Smith, D.F., X. Song, and R.D. Cummings, *Use of glycan microarrays to explore specificity of glycan-binding proteins.* Methods Enzymol, 2010. **480**: p. 417-44.

27. Agravat, S.B., et al., *GlycoPattern: a web platform for glycan array mining.* Bioinformatics, 2014. **30**(23): p. 3417-3418.

28. Ashline, D.J., et al., *Structural characterization by MSn of human milk glycans recognized by human rotaviruses.* Mol Cell Proteomics, 2014. **13**(11): p. 2961-2974.

29. Lock, K., et al., *Expression of CD33-related siglecs on human mononuclear phagocytes, monocyte-derived dendritic cells and plasmacytoid dendritic cells.* Immunobiology, 2004. **209**(1-2): p. 199-207.

30. Izquierdo-Useros, N., et al., *Siglec-1 is a novel dendritic cell receptor that mediates HIV-1 trans-infection through recognition of viral membrane gangliosides.* PLoS Biol, 2012. **10**(12): p. e1001448.

31. Li, N., et al., *Cloning and characterization of Siglec-10, a novel sialic acid binding member of the Ig superfamily, from human dendritic cells.* J Biol Chem, 2001. **276**(30): p. 28106-12.

32. Smith, D.F. and R.D. Cummings, *Application of microarrays for deciphering the structure and function of the human glycome.* Mol Cell Proteomics, 2013. **12**(4): p. 902-12.

33. Agravat, S.B., et al., *Computational Approaches to Define a Human Milk Metaglycome.* Bioinformatics, 2016. **In press.**

34. Appelmelk, B.J., et al., *Cutting edge: carbohydrate profiling identifies new pathogens that interact with dendritic cell-specific ICAM-3-grabbing nonintegrin on dendritic cells.* J Immunol, 2003. **170**(4): p. 1635-9.

35. van Die, I., et al., *The dendritic cell-specific C-type lectin DC-SIGN is a receptor for Schistosoma mansoni egg antigens and recognizes the glycan antigen Lewis x.* Glycobiology, 2003. **13**(6): p. 471-8.

36. Brazil, J.C., et al., *alpha3/4 Fucosyltransferase 3-dependent synthesis of Sialyl Lewis A on CD44 variant containing exon 6 mediates polymorphonuclear*

leukocyte detachment from intestinal epithelium during transepithelial migration. J Immunol, 2013. **191**(9): p. 4804-17.

37. Kobata, A., *Structures and application of oligosaccharides in human milk.* Proc Jpn Acad Ser B Phys Biol Sci, 2010. **86**(7): p. 731-47.

38. Jiang, X., et al., *Human milk contains elements that block binding of noroviruses to human histo-blood group antigens in saliva.* J Infect Dis, 2004. **190**(10): p. 1850-9.

39. Holla, A. and A. Skerra, *Comparative analysis reveals selective recognition of glycans by the dendritic cell receptors DC-SIGN and Langerin.* Protein Eng Des Sel, 2011. **24**(9): p. 659-69.

40. Naarding, M.A., et al., *Bile salt-stimulated lipase from human milk binds DC-SIGN and inhibits human immunodeficiency virus type 1 transfer to CD4+ T cells.* Antimicrob Agents Chemother, 2006. **50**(10): p. 3367-74.

41. Hong, P., et al., *Human milk oligosaccharides reduce HIV-1-gp120 binding to dendritic cell-specific ICAM3-grabbing non-integrin (DC-SIGN).* Br J Nutr, 2009. **101**(4): p. 482-6.

42. Jameson, B., et al., *Expression of DC-SIGN by dendritic cells of intestinal and genital mucosae in humans and rhesus macaques.* J Virol, 2002. **76**(4): p. 1866-75.

43. Engfer, M.B., et al., *Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract.* Am J Clin Nutr, 2000. **71**(6): p. 1589-96.

44. Gnoth, M.J., et al., *Human milk oligosaccharides are minimally digested in vitro.* J Nutr, 2000. **130**(12): p. 3014-20.

45. Thurl, S., et al., *Detection of four human milk groups with respect to Lewis blood group dependent oligosaccharides.* Glycoconj J, 1997. **14**(7): p. 795-9.

46. Zhuravleva, M.A., K. Trandem, and P.D. Sun, *Structural implications of Siglec-5-mediated sialoglycan recognition.* J Mol Biol, 2008. **375**(2): p. 437-47.

47. Poppe, L., et al., *Conformation of sLex Tetrasaccharide, Free in Solution and Bound to E-, P-, and L-Selectin.* Journal of the American Chemical Society, 1997. **119**(7): p. 1727-1736.

48. Nelson, R.M., et al., *Higher-affinity oligosaccharide ligands for E-selectin.* J Clin Invest, 1993. **91**(3): p. 1157-66.

49. Bode, L., et al., *Inhibition of monocyte, lymphocyte, and neutrophil adhesion to endothelial cells by human milk oligosaccharides.* Thromb Haemost, 2004. **92**(6): p. 1402-10.

50. Bode, L., et al., *Human milk oligosaccharides reduce platelet-neutrophil complex formation leading to a decrease in neutrophil beta 2 integrin expression.* J Leukoc Biol, 2004. **76**(4): p. 820-6.

51. Schumacher, G., et al., *Human milk oligosaccharides affect P-selectin binding capacities: in vitro investigation.* Nutrition, 2006. **22**(6): p. 620-7.

Chapter 4: Discussion

I. Implications of GBP-HMG Interactions

The major findings of this study are that specific HMG structures interact with specific GBPs expressed in the gastrointestinal (GI) tract. Additionally, at least some of these HMG-GBP interactions are potentially physiologically relevant since the HMG concentrations in human milk are at or above the K_d of the HMG-GBP interactions. Therefore, these studies warrant future studies on the physiological and immunological effects of HMG-GBP interactions on intestinal epithelial cells and dendritic cells.

A key finding in this manuscript was that HMGs specifically interacted with galectins (except galectin-2) and DC-SIGN (refer to **Chapters 2 and 3**). Many galectins are known to be expressed in the GI tract epithelium, including galectin-2, -3, -4, -8, and -9 in the intestine [1-6] and galectin-7 in the esophagus [7]. In order for these galectins to come in contact with HMGs, the galectins need to be cell surface localized and, in the case of the esophageal stratified epithelia, expressed in the outermost epithelial layer to be in contact with the lumen. However, the local concentration of cell surface galectins, affinity of galectins for the cell surface receptors, and degree of galectin exposure to the GI tract lumen *in vivo* is unclear. Moreover, all previous studies on galectin expression in the GI tract have focused on adult human cells and tissues, so the expression profile of galectins in the newborn and infant GI tract is still unknown. On the other hand, DC-SIGN expression has been confirmed to occur on dendritic cells (DCs) in the infant intestine [8]. However, it is unclear whether or not DC-SIGN actually becomes exposed to the GI tract lumen during DC dendrite “sampling” of the intestinal lumen in both

infants and adults. Thus, more studies are needed to determine if galectins and DC-SIGN are in physiologically relevant positions to truly interact with HMGs.

Although specific HMG-GBP interactions were determined to occur at physiologically relevant concentrations, these interactions also serve as a model for understanding GBP interactions with a complex pool of HMGs, as occurs in breast-fed infants. The concentration of HMGs in human milk ranges from about 8-25g/L [9-11], depending on the factors described in **Chapter 1.** GBPs typically recognize specific glycan determinants, as opposed to specific glycan structures [12], and these determinants are typically found on more than just one structure in a given metaglycome. For example, the LacNAc I, Blood Group H Type 1, and Lewis glycan determinants occur on more than just one structure in the HMG metaglycome [13, 14]. Thus, more than one HMG structure likely interacts with HMGs, as was seen in this study for galectin and DC-SIGN binding to the HMG metaglycome presented on the HMG shotgun glycan microarray. This means that one must instead calculate the concentration of *total* HMG binding determinants in human milk rather than the concentration of individual HMG structures. This total HMG determinant concentration, coupled with the K_d of binding to this single determinant or the average adjusted K_d if multiple determinants are bound, will determine the degree of HMG-GBP binding. However, the calculation for the degree of *in vivo* binding is not as straightforward since the *in vivo* interaction may also be influenced by GBP exposure, how strongly the GBP is bound to endogenous or exogenous (ex-microbial) ligands, and changes in HMG concentration during progression through the GI tract. Nonetheless, if HMGs are found to be critical for controlling infant immunity and development, their supplementation into infant formula will be warranted

in cases where human milk feeding is not practical or possible. In these cases, identifying specific HMG structures that exert similar physiological effects as total HMGs would be most beneficial so that the main bioactive HMG(s) can be supplemented at the proper concentrations into infant formula. The use of only a few specific HMGs for supplementation would help create a balance between maximal benefits to the infants and affordable cost, as opposed to using the entire HMG metaglycome (very expensive and currently impractical) or a single HMG (which may not produce all the associated health benefits of HMGs). Based on this study of HMG-GBP interactions and previous functional studies [15-22], LNT, LNnT, 2'-FL, 3'-SL, 6'-SL, and DSLNT have bioactive properties and represent a potential panel of HMGs for supplementation into infant formula.

Some infant formulas are supplemented with glycans such as fructooligosaccharides (FOS) and galactooligosaccharides (GOS). FOS and GOS have been shown to serve as a source of prebiotic fiber to promote establishment of a *Bifidobacteria*-rich gut microflora similarly to human milk and possibly other beneficial effects (reviewed in [23]). While some believe that these glycans may be used as a surrogate for HMGs in infant formulas, previous studies as well as the findings of this study do not support this observation. Recent systematic reviews suggest that FOS/GOS-supplemented infant formulas do not cause any appreciable health benefits vs. unsupplemented infant formulas [24, 25]. The reason for this lack of health benefits is due to the lack of structural similarity between FOS/GOS and HMGs. Thus, FOS and GOS do not have structural features enabling them to act as receptor decoy molecules, replace DSLNT in prevention of necrotizing enterocolitis, or even promote growth of

specific *Bifidobacteria* strains that grow well with HMGs but not FOS/GOS [26]. Additionally, the findings of this current study show that galectins and DC-SIGN bind to specific HMGs but not GOS mixtures or semi-purified GOS structures (refer to **Supplementary File 3** in **Chapter 2** and **Supplementary File 4** in **Chapter 3**). FOS were not tested in this study but, due to a lack of structural similarity between FOS and HMGs, are not expected to bind galectins or DC-SIGN. Therefore, if HMG interactions with galectins and DC-SIGN underlie many of the physiological and immunological effects mediated by HMGs, FOS/GOS will not mimic those functions. These results suggest that infant formula supplementation with specific HMGs may be more appropriate for improving infant health in formula-fed infants compared to FOS/GOS supplementation.

The binding of galectins to HMGs also raises questions about the physiological functions of galectins in the GI tract, especially in breast-fed infants. In breast-fed infants, galectins in the GI tract would constantly be exposed to HMGs, preventing their extracellular activities that are primarily mediated by cell surface glycoconjugate binding. In other words, extracellular galectin activity may be absent, and so the extracellular secretion of galectins would be a futile process. This raises questions as to whether or not galectin secretion in the GI tract may be shut down in breast-fed infants and/or by constant exposure to high concentrations of non-digestible galectin ligands such as HMGs. A major reason for this question is because we have shown in this study that human milk lacks galectins. This was a somewhat surprising finding because galectins are known to have antimicrobial properties [27-29], so galectins would be expected to support the developing newborn and infant immune system. Additionally, most cells and

tissues express and secrete at least one galectin [30], but human milk and hence mammary tissue appear to be examples of fluids and tissues where extracellular galectins are absent. Based on these findings, it is hypothesized that galectin secretion may be inhibited when cells are exposed to high concentrations of HMGs. The underlying mechanism may be the lack of cell surface glycoconjugate cross-linking by galectins upon HMG exposure. In other words, slight basal galectin secretion (which may only be a few molecules and hence undetectable by most detection methods) and specific cell surface glycoconjugate cross-linking may promote further galectin secretion. When HMGs are present, no glycoconjugate cross-linking occurs, so only basal level galectin secretion occurs. This autofeedback mechanism may be important for saving cells energy and galectin proteins by not wastefully secreting galectins into an extracellular environment, such as human milk or the breast-fed infant GI tract, where galectins activity would be inhibited. Since the galectin secretory pathway is a non-classical and unknown process [31], this suggested feedback mechanism of galectin secretion is plausible and should be tested in future studies, which may also unravel the general mechanism of galectin secretion.

DCs may come in contact with microbes in the intestine not only in the lamina propia and Peyer's Patches but also in the intestinal lumen via transepithelium dendrite extension [32] or migration to the lumen [33]. When exposed to the intestinal lumen, DCs may come in contact with intestinal contents, including glycans from mucins and ingested/dietary glycans such as HMGs from human milk. These intestinal glycans may bind to the TLRs, CLRs, and possibly other DC receptors to alter DC cell biology and immune responses. Within the lamina propria and Peyer's Patches, the DCs may still

come in contact with the small percentage of HMGs that are known to cross the intestinal epithelium during breast-feeding [34]. However, direct exposure of DCs to the intestinal lumen would bring DCs and their receptors in contact with a much larger concentration of HMGs. Thus, the higher luminal concentration of target HMG determinants during breast-feeding has a much higher probability of significantly binding to transepithelium-exposed DC-SIGN vs. DC-SIGN in the lamina propria or Peyer's Patches. However, it is unclear if DC-SIGN actually becomes exposed to HMGs during transepithelial dendrite extension, or even if DC-SIGN is expressed by the DC subsets that perform transepithelial dendrite extension or migrate into the lumen. Future studies should be aimed at determining if DC-SIGN and other receptors, including other CLRs and TLRs, become exposed to the intestinal lumen during DC transepithelial dendrite extension or migration into the lumen.

The fact that HMG functions overlap with galectin and DC-SIGN functions, namely in regulation of intestinal epithelial cell physiology and/or immune responses, suggests that there may be a common link between HMG and GBP functions. The results of this study show that galectins and DC-SIGN bind to specific HMGs, which may be the common link between these processes. In particular, it is hypothesized that HMG interactions with galectins and DC-SIGN mediate the physiological and immunological effects reported for HMGs, and a proposed model and method of testing this model are presented in the following two sections.

II. Proposed Model

Based on the findings of this study, three effects of HMG binding to cell surface GBPs involved in signaling can be proposed:

1) HMGs bind to specific GBPs, which leads to activation of the GBP signaling pathway
2) HMGs bind to specific GBPs but, because the GBP receptor requires cross-linking to activate the signaling pathway, no signaling activation occurs upon GBP binding to univalent HMGs. However, this GBP-HMG interaction interferes with GBP cross-linking by endogenous or exogenous multivalent ligands, thereby allowing the HMGs to serve as inhibitors or antagonists.
3) HMGs bind to specific GBPs but, because the HMG is not a multivalent ligand, does not activate the "typical" signaling pathway but instead activates an "alternative signaling pathway" that occurs by ligand binding independently of receptor cross-linking.

In the case of galectins, which are soluble, secreted GBPs as opposed to cell surface GBPs like DC-SIGN, the only expected outcome is that HMG binding to galectins leads to galectin dissociation from cell surface glycans. As a result, galectin-mediated receptor cross-linking is abrogated, and the dissociated galectins cannot rebind and cross-linking cell surface receptors until the HMG concentration is significantly reduced. In other words, galectin-HMG interactions are expected to result in effect #2, the inhibition of GBP-mediated activity. Additionally, effect #2 is the expected model for DC-SIGN-HMG interactions. Previous studies support effect #2 for DC-SIGN-HMG interactions for the following reasons. Dendritic cells have been shown to bind LNFPIII and subsequently promote a more Th2-biased response upon LPS-induced DC

maturation. However, this effect occurs with multivalent LNFPIII ligands but not free, reducing LNFPIII [35]. Moreover, multivalent but not free LNFPIII has functional effects on B cells; in fact, the free LNFPIII actually *inhibits* the multivalent LNFPIII activities in B cells [36]. Thus, free, reducing, univalent glycans such as LNFPIII and other HMGs may actually act as receptor antagonists vs. the multivalent glycoconjugates presented by endogenous and exogenous products such as microbial glycoproteins. The most likely receptor for the LNFPIII is DC-SIGN since LNFPIII contains Lewis x, a major determinant recognized by DC-SIGN (refer to **Figure 2** and **Table 2** in **Chapter 3** of this study). Indirect support for the DC-SIGN-LNFPIII interaction comes from studies using the multivalent glycoprotein *Schistosoma mansoni* soluble egg antigens (SEA), which contains Lewis x determinants similarly to LNFPIII in addition to other glycan determinants such as LDNF. SEA was shown to promote a Th2-biased activation of DCs similarly to LNFPIII, and this process was partially dependent on DC-SIGN in addition to MGL and the mannose receptor [37]. Although direct functional studies on LNFPIII-DC-SIGN interactions in DCs are lacking (ex-the effects of univalent and multivalent LNFPIII on untreated and anti-DC-SIGN-treated DCs), these findings support a functional role of multivalent but not univalent ligands such as HMGs for activation of DC-SIGN-mediated signaling and/or endocytosis. For these reasons, HMGs are likely to act as DC-SIGN antagonists rather than agonists. However, due to the lack of studies and especially direct functional evidence, it cannot be ruled out that HMG binding by DC-SIGN may lead to activation of signaling (effect #1). Indeed, one study suggested that monomeric mannose or mannan could activate the Raf-1 signaling pathway in DCs [38], although no direct proof of DC-SIGN-mannose binding was shown, so alternative CLR

binding (ex-mannose receptor) and signaling pathways may have mediated this effect. Nonetheless, the latter study suggests that HMGs may still lead to effect #1, or perhaps monomeric mannosylated ligands like soluble mannan but not fucosylated ligands like HMGs can activate DC-SIGN signaling.

From these hypothesized outcomes, the following model is proposed (**Figure 1**). In this model, breast-fed infants are exposed to a large amount of HMGs in their GI tract. The HMGs are present at a concentration at or above the K_d of binding to GBPs exposed to the GI tract lumen. These GBPs include galectins expressed by epithelial cells and DC-SIGN expressed by dendritic cells in the lamina propria and Peyer's Patches. The epithelial cells secrete some galectin proteins, which become bound to cell surface glycans found on cell surface receptors. This binding leads to receptor cross-linking and activation of downstream signaling pathways. The epithelial cells are directly exposed to the GI tract lumen and thus HMGs. As a result, HMGs can directly bind to extracellular galectins and cause galectin "elution" from the cell surface in breast-fed infants, thereby inhibiting galectin-mediated receptor cross-linking and signaling. In this way, extracellular galectin activity in the GI tract becomes low, if present at all.

In contrast to constant exposure to the intestinal lumen by intestinal epithelial cells, dendritic cells only become exposed to the intestinal lumen via migration from the lamina propria or dendrite extension from below the intestinal epithelium to "sample" antigens. During these sampling events, DC-SIGN becomes exposed to HMGs in the GI tract lumen. The exposure of DC-SIGN to HMGs leads to constitutive DC-SIGN binding to HMGs containing Lewis antigens, 2'-FL, and 3-FL. Two functional outcomes result from this interaction. First, the interaction prevents DC-SIGN-mediated signaling

initiated by interaction with endogenous or exogenous multivalent glycoconjugates. Second, constitutive HMG binding blocks DC-SIGN-mediated microbial uptake/phagocytosis, although other methods of antigen uptake will still be active. While this model may be especially relevant to the small intestine, the effects may occur all throughout the GI tract including in the mouth, esophagus, and stomach. However, the gut microbiota, which is primarily present in the large intestine but also in the lower portions of the small intestine, can digest HMGs. This digestion may destroy the GBP binding determinants and prevent their interaction with GBPs in these regions. In the other regions of the GI tract including the small intestine, the GBP-HMG interaction would ultimately depend on expression and exposure of galectins and DC-SIGN to the lumen.

Based on the known effects of galectin-receptor interactions, the functional outcomes of the galectin-HMG interactions are proposed to include intestinal epithelial cell differentiation into a more absorptive enterocyte as well as intestinal modeling into a highly absorptive, villi-rich environment. Additionally, altered cytokine secretion by the intestinal epithelium is also expected, namely an increase in Th1 cytokines and decrease in Th2, Th17, and pro-inflammatory cytokines. These cytokine changes will shift intestinal T cells towards a more balanced T cell response as opposed to the highly pro-inflammatory, Th2-biased response in newborns. These phenomena are expected to occur by changes in gene expression and protein expression, which can be tested by RNA sequencing and proteome analysis.

The constitutive activation of DC-SIGN by HMGs may be crucial for proper immunity in infant. First, this targeting is expected to inhibit DC-SIGN-mediated

endocytosis [39], although most other mechanisms of DC phagocytosis will still be active. Second, this interaction may block DC-SIGN activities in DCs that extend their dendrites through the intestinal epithelium as well as intraluminal DC subsets that have migrated from the lamina propria to the intestinal lumen [33]. In these lumen-exposed DCs, blocking DC-SIGN activity may be important during antigen sampling to prevent detrimental immunoregulatory events by DC-SIGN. These potentially detrimental activities of DC-SIGN ligand binding include reduced DC maturation seen for binding to specific mannosylated ligands [40], increased pro-inflammatory effects vs. LPS alone seen for binding to mannosylated ligands [38], and the anti-inflammatory and/or Th2-biasing responses stimulated by binding to fucosylated ligands [37, 41].Therefore, the inhibition of DC-SIGN by HMGs is expected to promote a more balanced immune response by preventing immunosuppression or highly inflammatory conditions as well as preventing pro-allergenic Th2-biased responses, which is especially critical in the immature and highly sensitive newborn and infant intestine.

III. Future Directions

Future studies are aimed at testing the components of this proposed model of GBP-HMG interactions in the GI tract. One important future study is on the effects of loss of extracellular galectin activity in the GI tract, especially the intestine. While galectins are known to regulate some components in epithelial cell biology (reviewed in [42]), the extracellular physiological functions of galectins in the intestinal epithelium are poorly understood. Therefore, studies are needed to determine if extracellular galectins play a role in intestinal epithelial cell biology in the first place and, if so, what roles they do play. To address this question, intestinal epithelial cells will be incubated with thiodigalactoside (TDG), a non-metabolizable analog of lactose and specific inhibitor of galectins. The addition of TDG to intestinal epithelial cells at a concentration well above the K_d of galectin binding (10mM-50mM) is expected to inhibit extracellular galectin activity in much the same way as proposed in the model shown in **Figure 1**. The effects of this treatment will be compared to both a mock-treatment control as well as cells treated with a non-metabolized carbohydrate that does not bind to galectins, such as cellobiose (Glcβ1-4Glc), to confirm that the effects are carbohydrate-specific. Based on previous findings of HMG effects on intestinal epithelial cells as well as galectin activities in other epithelial cells [19, 43], the expected results of this study are increased cellular differentiation (measured by specific enzyme activities such as alkaline phosphatase and sucrase, as well as microscopic examination), increased intestinal epithelium development (including reduced transepithelial resistance), changes in cytokine secretion (especially increased Th1-favoring cytokines such as IL-12 and decreased Th2-favoring cytokines such as IL-4), and changes in gene expression

(measured by RNA sequencing). If such changes are observed by the addition of TDG, the cells will then be treated with the HMGs LNnT, LNT, 2'-FL, or 3-FL. The expected results are the LNT, LNnT, and 2'-FL but not 3-FL, which does not bind galectins, will cause similar changes in intestinal epithelial cells as TDG-treated intestinal epithelial cells. Finally, these TDG and HMG treatment experiments will be repeated on primary small intestinal epithelial cells (available from Lonza) to confirm that the results apply in a more physiologically relevant model. The results of this future study will be instrumental in determining if galectins-HMG interactions have physiological consequences. The findings will also reveal what the actual physiological roles of galectins in intestinal epithelial cells are in the first place.

Galectins in the GI tract have also been shown to exert antimicrobial effects [28]. The antimicrobial effects occur towards microbes carrying specific glycan structures, especially structures resembling human glycan determinants and structures containing epitopes for so-called "natural" antibodies such as the Galα1-3Gal epitope [28, 29, 44, 45]. Thus, future studies should also be aimed at determining if physiologically relevant concentrations of HMGs, including specific HMG structures and/or pooled HMGs, inhibit the antimicrobial effects of galectins both *in vitro* and *in vivo*. This inhibition of galectins may be useful for newborns and infants to help establish an infant-specific microbiome and/or promote development of "natural" antibodies. In other words, inhibition of galectin activity may be needed for the infant immune system to begin sampling microbial antigens otherwise targeted by galectins, such as the microbes expressing the Galα1-3Gal epitope, so that these antibodies are generated. After weaning,

the combination of galectin activity as well as the natural anti-glycan antibodies may help present a double line of defense against specific pathogens carrying these epitopes.

Another future study of interest is testing the hypotheses of the model on HMG interactions in regulation of dendritic cell immune responses. A useful *in vitro* model for this future study would be *Schistosoma mansoni* egg antigen (SEA) interactions with DCs. SEA is a physiologically relevant model since *S. mansoni* eggs are excreted in feces (and thus found in the intestine), are critical to the immunopathogenic mechanism of disease, and modulate DC-mediated immune responses via interactions with CLRs including DC-SIGN (reviewed in [46, 47]). SEA exposure to DCs results in dampened DC maturation upon exposure to specific maturation signals as well as a bias towards a Th2-inducing phenotype upon LPS-mediated DC maturation, a mechanism that is dependent on DC-SIGN, MGL, and the mannose receptor [37]. Treatment of LPS- and SEA-stimulated monocyte-derived DCs in the presence of fucosylated HMGs such as 2'-FL, 3-FL, and LNFPIII are expected to block the effects of SEA on LPS-induced activation and a Th2-biased cytokine response. In other words, treatment with HMGs will partially revert the response to look more like LPS treatment alone, with increased costimulatory molecule expression and a more balanced Th1:Th2 response compared to the anti-inflammatory, Th2-biased effects of LPS and SEA treatment in the absence of HMGs. A non-DC-SIGN-specific HMG such as LNT can be used as a control. A mouse model of *S. mansoni* infection with or without 2'-FL or LNT feeding can then be utilized to determine if the 2'-FL but not LNT feeding promotes a more balanced Th1:Th2 response in the mouse, as expected from the model. Moreover, the effects of 2'-FL feeding may also alter egg shedding into the feces and the outcome of infection

(including symptoms such as weight loss as well as survival), and the 2'-FL but not LNT feeding is expected to reduce egg shedding and symptoms since Th2-biased immune responses is associated with chronic parasite infection [46, 48]. These studies would support a beneficial functional role of HMGs in controlling DC-SIGN activity in intestinal DCs of breast-fed infants.

Many galectins, C-type lectins, and I-type lectins were tested in this study for their ability to interact with HMGs. However, not all human GBPs expressed by dendritic cells and epithelial cells have been tested to date. For example, many other C-type lectins expressed by dendritic cells, including the mannose receptor (MR), DEC-205, and Mincle (reviewed in [49]), have not yet been tested. Future studies are thus aimed at screening these and other GBPs for the ability to bind HMGs as well as their glycan specificity. While some of these GBPs are now available commercially (ex-the mannose receptor and Mincle may now be purchased from R&D), some of these will require recombinant expression in and purification from eukaryotic cell lines.

While the future studies described above may describe potential physiological effects of GBP-HMG interactions, the ultimate goal is to see if the HMGs interact with GBPs in newborns and infants. This means that the experiments measuring the expression of GBPs in the infant GI tract must be performed since previous studies on GBP expression have only utilized adult, not newborn and infant, samples. Fortunately, one study provided good evidence that DC-SIGN is indeed expressed by some dendritic cells in the infant small intestine [8], but evidence for galectin expression in newborn and infant GI tract epithelial cells is still lacking. Once galectin expression in the infant GI tract is confirmed, supplementation of HMGs, particularly the mixture of HMGs

described above (LNT, LNnT, 2'-FL, 3'-SL, 6'-SL, and DSLNT), into infant formula should be performed and tested in clinical trials with non-supplemented control formula as a control. The clinical studies should measure the effects of HMGs on the functions predicted from studies on the physiological effects of GBP-HMG interactions, including a less inflammatory environment in the intestine, a balanced pro- and anti-inflammatory cytokine response, balanced T cell development, reduced incidence of gastrointestinal infections, reduced incidence of necrotizing enterocolitis and food allergies, and proper intestinal epithelium development.

From these future studies, the ultimate goal is to determine whether or not supplementation of specific HMGs into human milk will produce beneficial physiological outcomes vs. non-supplemented infant formula. While exclusive breast-feeding is still the ideal, it is not always practical or, in some cases, not even possible. In these cases, supplementation of infant formula with HMGs may at least help formula reach a state as close as possible to human milk in terms of health benefits. Understanding the mechanisms by which HMGs exert these health benefits will be especially useful in determining the parameters to measure in future clinical studies as well as which HMG structure(s) and concentration(s) should be supplemented to produce the intended outcome. The results of this study have taken us a few steps into this direction by showing that HMG binding to GBPs can occur, which now lead us to studies that should be aimed at determining what the physiological outcomes of these interactions are.

IV. Figures

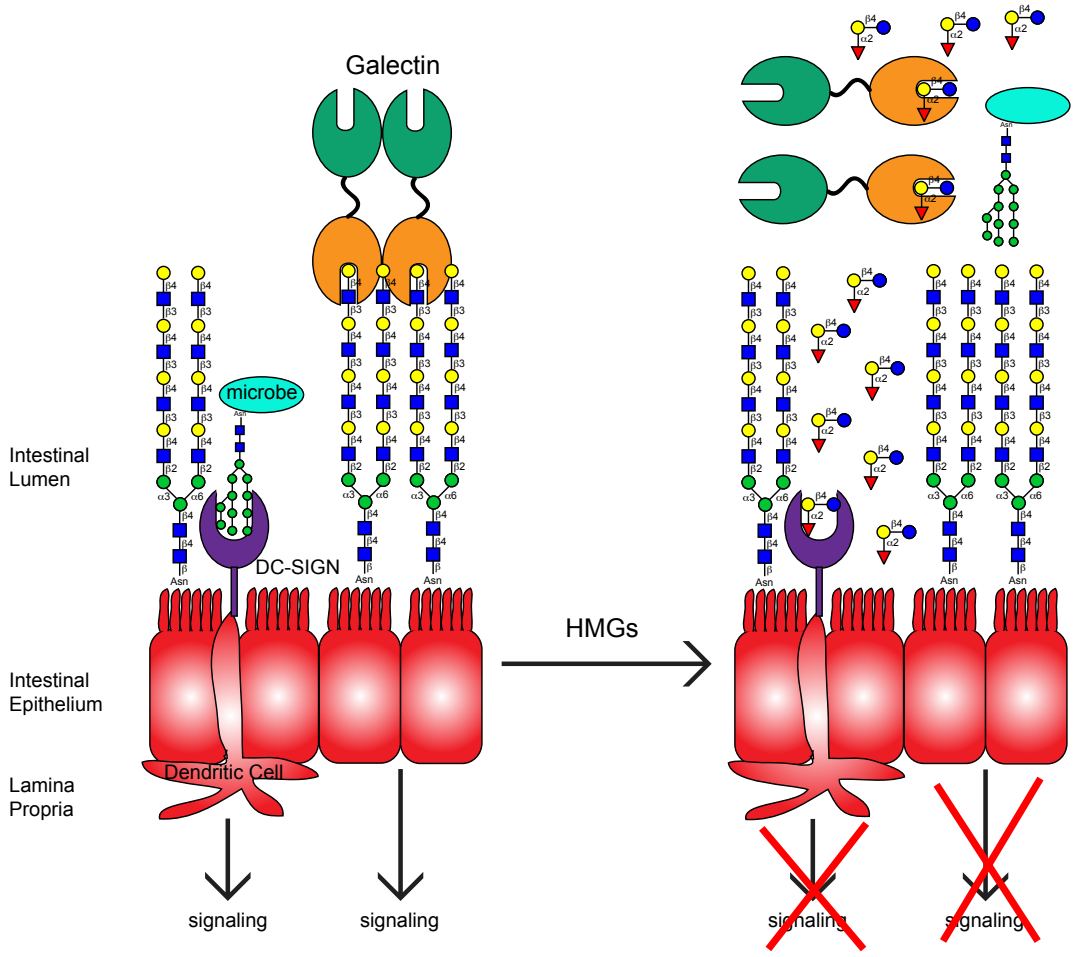


Figure 1. Proposed model of GBP-HMG interactions in the infant small intestine.

In this proposed model, galectins expressed by epithelial cells and bound to the epithelial cell surface as well as DC-SIGN expressed by dendritic cells become exposed to the intestinal lumen. In this state, the galectins bind to and cross-linking cell surface receptors, triggering epithelial cell signaling pathways. Meanwhile, DC-SIGN binds to endogenous or exogenous glycoconjugates, including microbial glycans such as yeast mannans, which leads to initiation of DC signaling pathways and/or microbial endocytosis (not depicted). Breast-fed infants ingest large amounts of HMGs such as 2'-fucosyllactose (2'-FL). The HMGs pass through the most of the GI tract lumen without

digestion or chemical modification and bind to galectins and DC-SIGN. These binding events lead to a loss of galectin binding and cross-linking of cell surface receptors, thereby reducing or even completely inhibiting galectin-mediated cellular signaling. Additionally, the binding of HMGs to exposed DC-SIGN leads to inhibition of endogenous or exogenous ligand binding; as a result, HMGs act as antagonists by binding to DC-SIGN and blocking DC-SIGN signaling as well as DC-SIGN-mediated endocytosis (not depicted). While only the intestine is depicted, this proposed model might be relevant throughout most of the GI tract. Other HMG structures besides 2'-FL (not depicted) are also present during breast-feeding that also contribute to galectin and DC-SIGN binding.

V. References

1. Huflejt, M.E. and H. Leffler, *Galectin-4 in normal tissues and cancer.* Glycoconj J, 2004. **20**(4): p. 247-55.

2. Jiang, K., et al., *Galectin-3 regulates desmoglein-2 and intestinal epithelial intercellular adhesion.* J Biol Chem, 2014. **289**(15): p. 10510-7.

3. Nagy, N., et al., *Galectin-8 expression decreases in cancer compared with normal and dysplastic human colon tissue and acts significantly on human colon cancer cell migration as a suppressor.* Gut, 2002. **50**(3): p. 392-401.

4. Lotz, M.M., et al., *Decreased expression of Mac-2 (carbohydrate binding protein 35) and loss of its nuclear localization are associated with the neoplastic progression of colon carcinoma.* Proc Natl Acad Sci U S A, 1993. **90**(8): p. 3466-70.

5. Saal, I., et al., *Human galectin-2: expression profiling by RT-PCR/immunohistochemistry and its introduction as a histochemical tool for ligand localization.* Histol Histopathol, 2005. **20**(4): p. 1191-208.

6. de Kivit, S., et al., *Intestinal epithelium-derived galectin-9 is involved in the immunomodulating effects of nondigestible oligosaccharides.* J Innate Immun, 2013. **5**(6): p. 625-38.

7. Magnaldo, T., D. Fowlis, and M. Darmon, *Galectin-7, a marker of all types of stratified epithelia.* Differentiation, 1998. **63**(3): p. 159-68.

8. Koning, N., et al., *Human Milk Blocks DC-SIGN-Pathogen Interaction via MUC1.* Front Immunol, 2015. **6**: p. 112.

9. Coppa, G.V., et al., *Changes in carbohydrate composition in human milk over 4 months of lactation.* Pediatrics, 1993. **91**(3): p. 637-41.

10. Urashima, T., et al., *The predominance of type I oligosaccharides is a feature specific to human breast milk.* Adv Nutr, 2012. **3**(3): p. 473S-82S.

11. Kunz, C., et al., *Lactose-derived oligosaccharides in the milk of elephants: comparison with human milk.* Br J Nutr, 1999. **82**(5): p. 391-9.

12. Cummings, R.D., *The repertoire of glycan determinants in the human glycome.* Mol Biosyst, 2009. **5**(10): p. 1087-104.

13. Ashline, D.J., et al., *Structural characterization by MSn of human milk glycans recognized by human rotaviruses.* Mol Cell Proteomics, 2014. **13**(11): p. 2961-2974.

14. Yu, Y., et al., *Human milk contains novel glycans that are potential decoy receptors for neonatal rotaviruses.* Mol Cell Proteomics, 2014. **13**(11): p. 2944-2960.

15. Jantscher-Krenn, E., et al., *The human milk oligosaccharide disialyllacto-N-tetraose prevents necrotising enterocolitis in neonatal rats.* Gut, 2012. **61**(10): p. 1417-25.

16. He, Y., et al., *The human milk oligosaccharide 2'-fucosyllactose modulates CD14 expression in human enterocytes, thereby attenuating LPS-induced inflammation.* Gut, 2014.

17. Vazquez, E., et al., *Effects of a human milk oligosaccharide, 2'-fucosyllactose, on hippocampal long-term potentiation and learning capabilities in rodents.* J Nutr Biochem, 2015. **26**(5): p. 455-65.

18. Castillo-Courtade, L., et al., *Attenuation of food allergy symptoms following treatment with human milk oligosaccharides in a mouse model.* Allergy, 2015. **70**(9): p. 1091-102.

19. Lane, J.A., et al., *Transcriptional response of HT-29 intestinal epithelial cells to human and bovine milk oligosaccharides.* Br J Nutr, 2013. **110**(12): p. 2127-37.

20. Bienenstock, J., et al., *Fucosylated but not sialylated milk oligosaccharides diminish colon motor contractions.* PLoS One, 2013. **8**(10): p. e76236.

21. Angeloni, S., et al., *Glycoprofiling with micro-arrays of glycoconjugates and lectins.* Glycobiology, 2005. **15**(1): p. 31-41.

22. Asakuma, S., et al., *Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria.* J Biol Chem, 2011. **286**(40): p. 34583-92.

23. Oozeer, R., et al., *Intestinal microbiology in early life: specific prebiotics can have similar functionalities as human-milk oligosaccharides.* Am J Clin Nutr, 2013. **98**(2): p. 561S-71S.

24. Mugambi, M.N., et al., *Probiotics, prebiotics infant formula use in preterm or low birth weight infants: a systematic review.* Nutr J, 2012. **11**: p. 58.

25. Mugambi, M.N., et al., *Synbiotics, probiotics or prebiotics in infant formula for full term infants: a systematic review.* Nutr J, 2012. **11**: p. 81.

26. Ward, R.E., et al., *In vitro fermentation of breast milk oligosaccharides by Bifidobacterium infantis and Lactobacillus gasseri.* Appl Environ Microbiol, 2006. **72**(6): p. 4497-9.

27. Kohatsu, L., et al., *Galectin-3 induces death of Candida species expressing specific beta-1,2-linked mannans.* J Immunol, 2006. **177**(7): p. 4718-26.

28. Stowell, S.R., et al., *Innate immune lectins kill bacteria expressing blood group antigen.* Nat Med, 2010. **16**(3): p. 295-301.

29. Stowell, S.R., et al., *Microbial glycan microarrays define key features of host-microbial interactions.* Nat Chem Biol, 2014. **10**(6): p. 470-6.

30. Uhlen, M., et al., *Proteomics. Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.

31. Hughes, R.C., *Secretion of the galectin family of mammalian carbohydrate-binding proteins.* Biochim Biophys Acta, 1999. **1473**(1): p. 172-85.

32. Chieppa, M., et al., *Dynamic imaging of dendritic cell extension into the small bowel lumen in response to epithelial cell TLR engagement.* J Exp Med, 2006. **203**(13): p. 2841-52.

33. Arques, J.L., et al., *Salmonella induces flagellin- and MyD88-dependent migration of bacteria-capturing dendritic cells into the gut lumen.* Gastroenterology, 2009. **137**(2): p. 579-87, 587 e1-2.

34. Gnoth, M.J., et al., *Investigations of the in vitro transport of human milk oligosaccharides by a Caco-2 monolayer using a novel high performance liquid chromatography-mass spectrometry technique.* J Biol Chem, 2001. **276**(37): p. 34363-70.

35. Thomas, P.G., et al., *Maturation of dendritic cell 2 phenotype by a helminth glycan uses a Toll-like receptor 4-dependent mechanism.* J Immunol, 2003. **171**(11): p. 5837-41.

36. Velupillai, P. and D.A. Harn, *Oligosaccharide-specific induction of interleukin 10 production by B220+ cells from schistosome-infected mice: a mechanism for*

regulation of CD4+ T-cell subsets. Proc Natl Acad Sci U S A, 1994. **91**(1): p. 18-22.

37. van Liempt, E., et al., *Schistosoma mansoni soluble egg antigens are internalized by human dendritic cells through multiple C-type lectins and suppress TLR-induced dendritic cell activation.* Mol Immunol, 2007. **44**(10): p. 2605-15.

38. Gringhuis, S.I., et al., *Carbohydrate-specific signaling through the DC-SIGN signalosome tailors immunity to Mycobacterium tuberculosis, HIV-1 and Helicobacter pylori.* Nat Immunol, 2009. **10**(10): p. 1081-8.

39. Appelmelk, B.J., et al., *Cutting edge: carbohydrate profiling identifies new pathogens that interact with dendritic cell-specific ICAM-3-grabbing nonintegrin on dendritic cells.* J Immunol, 2003. **170**(4): p. 1635-9.

40. Geijtenbeek, T.B., et al., *Mycobacteria target DC-SIGN to suppress dendritic cell function.* J Exp Med, 2003. **197**(1): p. 7-17.

41. Gringhuis, S.I., et al., *Fucose-specific DC-SIGN signalling directs T helper cell type-2 responses via IKKepsilon- and CYLD-dependent Bcl3 activation.* Nat Commun, 2014. **5**: p. 3898.

42. Viguier, M., et al., *Galectins in epithelial functions.* Tissue Barriers, 2014. **2**: p. e29103.

43. Holscher, H.D., S.R. Davis, and K.A. Tappenden, *Human milk oligosaccharides influence maturation of human intestinal Caco-2Bbe and HT-29 cell lines.* J Nutr, 2014. **144**(5): p. 586-91.

44. Galili, U., et al., *Interaction between human natural anti-alpha-galactosyl immunoglobulin G and bacteria of the human flora.* Infect Immun, 1988. **56**(7): p. 1730-7.

45. Posekany, K.J., et al., *Induction of cytolytic anti-Gal antibodies in alpha-1,3-galactosyltransferase gene knockout mice by oral inoculation with Escherichia coli O86:B7 bacteria.* Infect Immun, 2002. **70**(11): p. 6215-22.

46. van Die, I. and R.D. Cummings, *Glycan gimmickry by parasitic helminths: a strategy for modulating the host immune response?* Glycobiology, 2010. **20**(1): p. 2-12.

47. Olveda, D.U., et al., *The chronic enteropathogenic disease schistosomiasis.* Int J Infect Dis, 2014. **28**: p. 193-203.

48. Caldas, I.R., et al., *Human schistosomiasis mansoni: immune responses during acute and chronic phases of the infection.* Acta Trop, 2008. **108**(2-3): p. 109-17.

49. Geijtenbeek, T.B. and S.I. Gringhuis, *Signalling through C-type lectin receptors: shaping immune responses.* Nat Rev Immunol, 2009. **9**(7): p. 465-79.