Human milk glycobiome and its impact on the infant gastrointestinal microbiota

Angela M. Zivkovic a,b,c, J. Bruce German a,b,c, Carlito B. Lebrilla a,c,d,e, and David A. Mills a,c,f,g,1

1Foods for Health Institute, aDepartment of Food Science and Technology, bFunctional Glycobiology Program, cDepartment of Chemistry, dDepartment of Biochemistry and Molecular Medicine, eRobert Mondavi Institute for Wine and Food Science, and fDepartment of Viticulture and Enology, University of California, Davis, CA 95616

Edited by Todd R. Klaenhammer, North Carolina State University, Raleigh, NC, and approved July 2, 2010 (received for review March 10, 2010)

Human milk contains an unexpected abundance and diversity of complex oligosaccharides apparently indigestible by the developing infant and instead targeted to its cognate gastrointestinal microbiota. Recent advances in mass spectrometry-based tools have provided a view of the oligosaccharide structures produced in milk across stages of lactation and among human mothers. One postulated function for these oligosaccharides is to enrich a specific “healthy” microbiota containing bifidobacteria, a genus commonly observed in the feces of breast-fed infants. Isolated culture studies indeed show selective growth of infant-borne bifidobacteria on milk oligosaccharides or core components therein. Parallel glycoprofiling documented that numerous Bifidobacterium longum subsp. infantis strains preferentially consume small mass oligosaccharides that are abundant early in the lactation cycle. Genome sequencing of numerous B. longum subsp. infantis strains shows a bias toward genes required to use mammalian-derived carbohydrates by comparison with adult-borne bifidobacteria. This intriguing strategy of mammalian lactation to selectively nourish genetically compatible bacteria in infants with a complex array of free oligosaccharides serves as a model of how to influence the human supraorganismal system, which includes the gastrointestinal microbiota.

glycoprofiling | human milk oligosaccharides | infant microbiota | Bifidobacterium | diet

The interaction of humans with microorganisms remains one of the most important relationships to both acute survival and long-term health. Humans emerged into a microbial world, and the microbial world continues to shape human evolutionary progress. For example, successes such as the discovery and application of small-molecule antibiotics have not only saved lives, but by intervening in the fundamental relationships between humans and microbes they have imposed selection pressures on the evolution of microorganisms. Understanding how to manage microbial biology in the future will require more sophisticated tools aimed at modifying microbial populations and functions toward human health benefits other than simply preventing pathogenic infection. Insights into how to guide human/microbial interactions to be net favorable for both are needed.

The connection between human breast milk and infants’ growth, development, and health exemplifies this link. Human milk is the culmination of 200 million years of Darwinian pressure on mammalian lactation as the sole source of early infant nourishment. Human milk components not only nourish the infant, they provide myriad bioactive compounds for the offspring that influence the growth, stimulation, and modulation of the immune system, cognitive development, protection from toxins and pathogenic diseases, and perhaps most remarkably, the establishment of the intestinal microbiota (1–3). Considerable efforts made to understand the biology of human milk and its effects on the infant (4) are beginning to elucidate the structure/function properties and benefits that milk provides.

The constant evolutionary pressure on milk as the sole source of nourishment for mammalian infants has resulted in a remarkable model for how diet affects all aspects of development and health. Maternal investment, specifically the composition of breast milk, has been shaped by natural selection acting on both the infant and the mother, maximizing infant survival, growth, and activity while minimizing the costs of lactation for the mother (5). One of the most remarkable apparent functions of breast milk is the selective colonization and support of a protective microbiota. How milk has been able to attain the goal of guiding the evolutionary emergence of specific strains of bacteria in infants for their mutual health benefit is precisely the kind of question that science needs to understand to achieve similar successes for a wide range of human/microbial interactions.

Human milk/colostrum contains between 5 and 23 g/L (6, 7) of oligosaccharides containing a lactose-reducing end elongated with fucosylated and/or sialylated N-acetyllactosamine units (8). This translates to over 200 different human milk oligosaccharide (HMO) structures that differ in their size, charge, and sequence (8).

Despite the fact that oligosaccharides are the third most abundant components in milk (after lactose and lipids; Fig. 1), they were long thought to have no biological significance. It is now known that oligosaccharides in milk are important for the healthy growth of infants (9–11). Certain HMOs derived from the mammary epithelial cells of the mother also share common structural motifs with glycans on the infant’s intestinal epithelia known to be receptors for pathogens. The presence of such structures in milk implies a defensive strategy, with glycans acting as decoys to prevent binding of pathogens to epithelial cells, thereby protecting infants from disease (11). Consistent with these multiple functions, human milk is comprised of a complex mixture of oligosaccharides that differ in size, charge, and abundance (8). Several of the characteristic HMO structures and their isomers found in pooled samples of human milk are shown in Fig. 2. A direct mechanistic link between specific HMO structures and bifidobacterial growth has recently been established and will be discussed in this review (12–19).

Novel Analytical Tools: Structural Determination of HMO

The basic structure of HMOs includes a lactose core at the reducing end, which is elongated by N-acetyllactosamine units, with greater structural diversity provided by extensive fucosylation and/or sialylation wherein fucose and sialic acid residues are added at the terminal positions. Lactose, the most abundant component of milk, is a disaccharide composed of a galactose β-linked to the 4 position of glucose (Galβ1–4Glu) (6). HMOs are composed of both neutral and anionic species with building blocks of five monosaccharides: D-glucose, D-galactose, N-acetylglos-

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, “Microbes and Health,” held November 2–3, 2009, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and audio files of most presentations are available on the NAS Web site at http://www.nasonline.org/SACKLER_Microbes_and_Health. Author contributions: A.M.Z., J.B.G., C.B.L., and D.A.M. wrote the paper. The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence should be addressed. E-mail: damills@ucdavis.edu.
Lactose 70 g/L

HMO 5–15 g/L

Proteins 8 g/L

Lipids 40 g/L

Fig. 1. Human milk composition. (Left) Macronutrient composition of pooled human milk, with lactose being the most abundant component at 70 g/L, followed by lipids at 40 g/L. The third most abundant component is HMO at an estimated 5–15 g/L, followed by protein at 8 g/L. (Right) Pull-out pie chart showing composition of the most abundant HMOs. Masses of individual HMO structures are shown, along with their relative abundances, which were calculated from peak intensities presented in Ninonuevo et al. (21).

Fig. 2. Characteristic HMO structures. The structures of several characteristic HMO structures found across human milk samples, along with their isoforms, are shown. Two isoforms with mass 709.3: (A) LNT, lacto- N-tetraose and (B) LNNnT, lacto- N-neotetraose; (C) mass 1511.6: MSMFLNnH, monofucosylmonosialyllacto- N-hexaose; four isoforms with mass 855.3: (D) LNFP I, lacto- N-fucopentaose, (E) LNFP II, (F) LNFP III, and (G) LNFP V; and five isoforms with mass 1220.4: (H) MFLNH I, monofucosyllacto- N-hexaose, (I) MFLNH II, monofucosylparalacto- N-hexaose, (J) IFNnH I, isomer 1 fucosylparalacto- N-hexaose, and (L) IFNnH III, isomer 3 fucosylparalacto- N-hexaose. Glc, D-glucose; Gal, D-galactose; GlcNAc, N-acetylgalactosamine; Fuc, L-fucose; Neu5Ac, N-acetylneuraminic acid.

Characterize oligosaccharides in human milk include HPLC, high-pH anion exchange chromatography (HPAEC), capillary electrophoresis (CE), NMR, and MS (22–30). MS is evolving as the preferred method for elucidating HMO composition because it provides high sensitivity with structural information. A unique analytical strategy to rapidly profile oligosaccharides in human milk using HPLC-Chip/time-of-flight (TOF) MS technology has recently been developed (8, 31). This analytical technique uses an integrated microfluidic chip coupled with a high mass accuracy TOF mass analyzer. The HPLC-Chip/MS system replaces the traditional column and regular fittings in standard LC/MS systems with an integrated microfluidic chip, providing effective separation with significantly improved sensitivity and reproducibility.

Due to the nature of chemical structures of the oligosaccharides, a single atomic composition (or a single mass) can be made up of several structural isomers (Fig. 2). Variations in structures of isomers can range from linkage isomers (α1–2 vs. β1–2 or vs. α1–4) to variations in branching and positional isomers. The number of isomers can be large—>10 in some cases—with many of the isomers in relatively equal abundances. The extent and roles of this oligosaccharide heterogeneity in milk, or in any biological mixture, are still not well understood (32), primarily because the complexity of the mixture makes it difficult to monitor individual isomers and elucidate their functions.

A systematic method to elucidate the HMO structures has been developed (8). The HMO are first separated by standard HPLC into several fractions, and each fraction is then examined by both matrix-assisted laser desorption/ionization (MALDI) Fourier transform ion cyclotron resonance (FT-ICR) MS and HPLC-Chip/TOF MS to determine the number of structures in each fraction as well as their HPLC-Chip/TOF MS retention times. Tandem mass spectra of each component are obtained with either infrared multiphoton dissociation (IRMPD) or collision-induced dissociation (CID). The fragmentation patterns are used to elucidate the sequences of unknown structures. The exact linkages between each residue and the identity of the monosaccharide residue are then sequentially determined using a suite of biologically relevant exoglycosidases in a stepwise manner followed by MS after each step to evaluate the course of the reaction. Using this process, a complete library of structures found in human milk is being constructed. The combination of retention times, accurate masses,
and tandem MS are being compiled to develop a rapid identification of milk oligosaccharides.

**Structural Diversity and Functional Implications**

An emerging feature of the structural analysis of oligosaccharides in human milk is the diversity among and within women. Approximately 200 molecular species have been identified in pooled human milk samples. These structures consist of neutral and acidic oligosaccharides containing a high degree of fucosylation, and to a lesser extent, sialylation (8). HMOs are terminated by fucose and sialic acid, and even in this aspect vary in the range of 50–70% fucosylated and 5–15% sialylated. Daily profiles of HMOs reflect the fluctuations between and within lactation in individual mothers (8). These results were supported by reports from other laboratories that showed similarly large heterogeneity in milk oligosaccharides using other methods, including HPLC and HPAEC methods but monitoring less than 20 oligosaccharide components (33, 34). This description of the basic composition of HMOs in humans will need to be broadened to an equally detailed understanding of the relationship between the levels of specific milk oligosaccharides and the specific functions these biomolecules contribute to maternal and infant health and development.

Several characteristic oligosaccharides (Fig. 2) illustrate the variation among individual mothers. For some mothers, the most dominant component lacto-N-fucopentaose I (LNFP I; mass 709.3) can be 10× more intense than the next most abundant lacto-N-fucopentaose I/V (LNFP I/V; mass 855.3). For others, the three most abundant components—LNnT, lacto-N-tetraose (LNT; mass 709.3), and LNFP I/V—make up over 50% of the total. Among all samples analyzed to date, a neutral oligosaccharide with neutral mass 709.3 Da (3Hex, 1HexNAc;LNnT) is the most prominent. Importantly, this specific molecule was found to be preferentially consumed by several bifidobacterial strains (13). The next most common structures consist of fucosylated oligosaccharides with masses of 855.3 Da and 1220.4 Da (found in all five donors), and 1511.6 Da, a fucosylated species with a sialic acid residue. These specific HMOs are likely candidates for the important biological roles of milk in maintaining a healthy gut microbiota and in the prevention of pathogenic diseases among infants, as previously reported (6, 11, 18, 35).

The reasons for variation in HMOs among women have not yet been fully elucidated. Some of this variation could be related to diet, lifestyle, ethnicity, and other factors. One factor that appears to be related to the variation among mothers is secretor status. Secretors have oligosaccharides that are more common to other secretors and different from those that are nonsecretors. Secretors produce Fuc1-2 motifs in secreted oligosaccharides. Nonsecretors do not have the gene that produces this motif. Although the secretor status of mothers has not been associated with infant outcomes, the ability of individuals to resist viral and bacterial infection has been correlated with their secretor status (36, 37). Kobata suggested a close relationship between the structures of milk oligosaccharides and the Lewis blood group systems (38). Thurl et al. (39) further observed that HMOs can be used to separate the subjects into four human milk groups corresponding to the presence of Lewis A and B. Lewis group blood types are named because of the fucose-containing glycans found on the surfaces of erythrocytes and in glycoproteins of secreted fluids. Lewis motifs are found on the surfaces of cells and have been implicated in many diseases, from cancer to infection. It has been shown, for example, that Lewis B exhibits preferential binding to pathogens, specifically *Helicobacter pylori* (40). Lewis-related antigens have been shown to have prognostic value for diseases as diverse as cancer and celiac disease (41). The Lewis structures are not synthesized in erythroblasts but are believed to be produced in the gut epithelium where they are shed into the digestive tract, digested, reabsorbed, and transported as glycolipids into the plasma where they are absorbed onto red blood cells (42). Lewis groups are not present on the erythrocytes of newborns, but appear between 3 and 6 mo and stabilize in concentration between 3 and 6 y (41). It is perhaps no coincidence that the mother administers large doses of Lewis structures to the infant during breast-feeding.

Fucosylation and sialylation are involved in structural motifs such as those belonging to the Lewis groups, specifically A, B, X, Y, and sialyl Lewis X. Many of these motifs are found in HMOs. Fucose and sialic acids both have distinct masses that can be readily obtained from the accurate mass. A mass profile of the mixture therefore provides a rapid method for determining the extent of fucosylation in HMO. The possibility that these differences relate to variations in the consequences of different human milks to the colonization, development, and net health properties of individual infants’ microbiota is compelling from these data. With these new analytical techniques in hand it will be possible to address important biological questions related to the fundamental functions of HMO: How do they affect the basic microbial ecology of the human infant intestine?

**Milk Oligosaccharide Interactions with Bacteria**

Oligosaccharides are known to interact directly with the surfaces of bacteria. Because the mechanism of milk oligosaccharide production by mammary cells involves the same enzymes as those for glycoproteins and glycolipids (10), it is believed that HMOs contain the same structural moieties as cell surface glycoconjugates. Bacterial surfaces have oligosaccharide binding proteins, including lectins and other target receptors such as Toll-like receptors (43). Though oligosaccharide binding to the surfaces of commensal bacteria is not well investigated, there is a larger body of work studying glycan binding with pathogenic bacteria. Various oligosaccharides and glycoconjugates in milk are believed to inhibit the binding of pathogenic bacteria and toxins, presumably by acting as decoys and binding to the bacterial surface, thus inhibiting their ability to bind to target oligosaccharides on the surface of epithelial cells (10, 11). Antiadhesive activity of free HMOs has been described for *Streptococcus pneumoniae* (44), enteropathogenic *E. coli* (45, 46), *Listeria monocytogenes* (47), *Vibrio cholerae* (48), *Salmonella typhi* (48), and HIV (49). Milk glycolipids or glycoproteins have also been implicated in protective mechanisms against pathogens such as *Pseudomonas aeruginosa* (50), noroviruses (36), *Cholera* and Shiga toxins (51), and Rotavirus (52).

It is the impressive repertoire of structural diversity present in the aggregate HMO pool that likely contributes to the diverse protective functions against a range of pathogens or toxins. Several studies have shown that discrete fractions of oligosaccharides in human milk can differentially inhibit pathogen adhesion. Newburg and coworkers (35) have shown that fucosylated HMOs inhibit binding of *Campylobacter jejuni*-cultured HEp-2 intestinal cells and to human intestinal mucosa, and also reduce *C. jejuni* colonization of mice. Others have shown that sialylated milk oligosaccharides can block adhesion of enterotoxigenic and uropathogenic *E. coli* to human erythrocytes (53).

In addition to depletion of pathogens, milk oligosaccharides are also believed to enrich a beneficial microbiota containing bifidobacteria, a species commonly found in the feces of breast-fed infants (54–56)—an observation that dates back over 100 y (57). A prominent bifidobacterial population in the feces of breast-fed infants has been reported from numerous culture-based and non–culture-based studies (58). More recent deep-sequencing approaches have challenged the notion of strict bifidobacterial predominance (59, 60). Regardless, bifidobacteria are often overrepresented in the breast-fed infant microbiome by comparison with their appearance in adults.

If milk oligosaccharides act as prebiotic substrates that help shape the microbial content of the infant gastrointestinal tract, one would predict that the commensal bacteria normally enriched in a breast-fed infant would possess the capacity to grow on these complex structures. This concept is not new. Gyorgy and coworkers...
has an inherent and longum and B. longum and infantis or subsp. longum infantis B. longum ATCC15700 was not able to ATCC15697 B. catenulatum HMO-related gene cluster 1 from via an ABC transporter and an as-
subsp. Propionibacterium longum(63). In aggregate, these DJO10A and was only detected and Eubacterium E. coli subsp. subsp. B. animalis and was only detected and was only detected HMO gene cluster 1 from breast milk, the amount consumed by nonsialylated HMO species, LNnT. Although LNnT is an abun-
dant HMO in breast milk, the amount consumed by B. longum subsp. longum and B. breve ATCC15700 consumed only a portion of a single, nonfucosylated/ nonsialylated HMO species, LNnT. Although LNnT is an abun-
dant HMO in breast milk, the amount consumed by B. longum subsp. longum and B. breve represents only a small portion of the overall HMO pool (13). B. breve ATCC15700 was not able to readily consume the bulk of HMO structures, but did grow on all of the monomer constituents of HMO (19), suggesting a possible cross-feeding capacity in the gastrointestinal tract via liberated monosaccharides. The catabolic capacity of these bacteria toward HMO can also be measured by monitoring the sialidase and fucosidase activities required to deconstruct these complex glycan structures. Enzymatic assays showed that B. longum subsp. infants has a higher sialidase activity when grown on lactose as com-
pared with B. longum subsp. longum and B. breve, respectively (13).

Even though we cannot exclude a minimal/nonspecific sialidase activity in B. longum subsp. longum and B. breve, these data suggest that B. longum subsp. infants has an inherent and constitutive ability to process sia
ylated compounds. Furthermore, among the three strains tested, fucosidase activity was only present in B. longum subsp. infants and was only detected upon growth on HMOs (13). A different mode of catalytic activity toward HMO consump-
tion is illustrated by B. bifidum, which exports a 1,2α-fucosidase (AfcA) and a 1–3/4α-fucosidase (AfcB) that defucosylate HMO structures (64, 65). An extracellular lacto-N-biosidase then liber-
ates LNB from HMO species lacking fucosylated and sialylated residues. Wada et al. (66) showed this lacto-N-biosidase activity is minimally conserved in bifidobacteria, being present in select R. bifidum and B. longum subsp. longum strains. Upon release, LNB is transported into B. bifidum via an ABC transporter and an as-
associated LNB-specific solute-binding lipoprotein (67, 68) whereby it is further processed and fed into the central metabolic pathway. The recent genome sequencing of B. longum subsp. infants ATCC15697, a prototypical HMO consumer, has greatly advanced understanding of the genetic underpinnings of this un-
usual phenotype (69). Strain ATCC15697 possesses a host of HMO-related genes clustered into four loci, including one large locus, HMO cluster 1, which contains all of the glycosidases (sialidase, fucosidase, galactosidase, and hexosaminidase) and transporters necessary for importing and metabolizing HMO (Fig. 3). Sequencing of several more isolates confirms this common HMO-related genomic architecture among B. longum subsp. infants isolates and clearly provides a genetic rationale for the vigorous growth of this clade on HMOs. Strain ATCC15697 possesses both sialidase and fucosidase activities when grown on HMOs (13), and these results were corroborated by expression of the fucosidase and sialidase genes, Blon_2336 and Blon_2348, respectively, as observed via proteomics (17).

A particularly interesting aspect of the large HMO cluster (Fig. 3) is the extensive repertoire of extracellular solute-binding pro-
teins (SBP; pfam 01547) predicted to bind oligosaccharides. Six of these cluster I lipoproteins exhibit a pronounced evolutionary divergence relative to other SBP Family 1 proteins in bifidobacteria (17), suggesting a possible relationship with milk oligo-
saccharides. Interestingly, the B. longum subsp. infants genome contains a total of 21 Family 1 SBP, roughly twice as many as observed in the B. longum subsp. longum or B. adolescentis genomes (17). Proteomics of HMO-grown B. longum subsp. infants revealed expression of several Family 1 SBP from the large HMO cluster as well as two additional SBP that are located elsewhere on the genome.

Although the four HMO-related clusters are shared among B. longum subsp. infants isolates, they are notably absent in other sequenced bifidobacteria, such as B. longum subsp. longum DJO10A (70) and B. adolescentis ATCC15703 (GenBank accession no. AP009256), which grow weakly or not all (respectively) on HMOs (12). Interestingly, one possible HMO-related gene set shared between ATCC15697 and DJO10A is the seven-gene op-
eron responsible for LNB metabolism (71). Given that DJO10A is able to weakly grow on HMO, and glycoprofiling indicated a small consummation of LNNt, it is tempting to speculate that this operon is linked to consumption of that particular HMO moiety. Although it is very hard to generalize the mechanisms of HMO catabolism across bifidobacteria because of strain heterogeneity and taxonomic confusion (72) within the genera, several impor-
tant trends have emerged. The most common infant-borne bif-

Fig. 3. HMO-related gene cluster 1 from B. longum subsp. infants ATCC15697. HMO gene cluster 1, shown here, contains all of the necessary glycosidases (sialidase, fucosidase, galactosidase, and hexosaminidase) and carbohydrate transporters necessary for importing and metabolizing HMOs.
**Conclusion**

Microbial ecosystems are guided by a complex interplay of genetics and environment. Even the most narrowly defined ecosystem contains a complex array of microorganisms competing to maintain their respective niches. The human intestine is one such ecosystem in which the host represents an unusual dimension of the environment. The host dictates many of the conditions under which different bacteria compete. So, too, in an interesting sense is the microbiome part of each individual human’s overall genotype. The concept of this mutually interactive community as a supraorganismal system is compelling (73). As scientists gain an increasingly detailed description of this supraorganismal system and its influence on human health, the opportunities to influence it will become increasingly attractive. However, two questions will be central to these opportunities: What direction should it change, and how? The answers to these questions will not be easy. Persistent, even minor, changes in a mature, established microbiota are difficult to achieve even with aggressive antibiotic treatments (74). When achieved, it is not clear whether they are in the long term beneficial to the host’s overall health. Human milk emerged from 200 million years of relentless Darwinian pressure to be nourishing to infants. The pervasive influence of bacteria and other microorganisms on the immunologically naïve infant was clearly a vital element of the evolutionary pressure on lactation. The evidence that is assembling from a wide variety of scientific disciplines points to oligosaccharide structures and their abundance as being a key element in evolution’s strategy to establish and guide the human infant microbiome. Interestingly, weaning results in an immediate change in the infant microbiome (75). Weaning is also associated with an increased risk of a variety of intestine-related diseases (76). This basic strategy of selectively nourishing genetically compatible bacteria with a complex array of free oligosaccharides may well be the most immediate route to designing diets that are capable of influencing the human supraorganismal system for net health benefit.

**ACKNOWLEDGMENTS.** We thank all the students, postdocs, and staff in the University of California–Davis Milk Bioactivities Program (http://mbp.ucdavis.edu) and Functional Glycobiology Program (http://fgp.ucdavis.edu) for their hard work and inspiration. We specifically acknowledge David Sela and Mariana Barboza for their assistance in preparing this manuscript. This publication was made possible in part by support from the University of California Discovery Grant Program, the California Dairy Research Foundation, US Department of Agriculture National Research Initiative Cooperative State Research Education, and Extension Service Award 2008-35200-18776, National Institute on Environmental Health Sciences Superfund P42 ES02710, the Childhood Autism Risks from Genetics and the Environment Study Grant P01 ES11269, and National Institutes of Health–National Institute of Child Health and Development Awards 5R01HD059127 and 1R01HD061923.


