6440 J. Agric. Food Chem. 2010, 58, 6440–6448
DOI:10.1021/jf100112x

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

Glycoprotein Expression in Human Milk during Lactation

John W. Froehlich,[†] Eric D. Dodds,[†] Mariana Barboza,[†] Erica L. McJimpsey,[†] Richard R. Seipert,[†] Jimi Francis,[‖] Hyun Joo An,[†] Samara Freeman,[‡] J. Bruce German,[‡] and Carlito B. Lebrilla*,[†,§]

[†]Department of Chemistry, [‡]Department of Food Science and Technology, and [§]School of Medicine, University of California, Davis, California 95616, and [‖]Department of Biochemistry, University of Nevada, Reno, Nevada 89557

While milk proteins have been studied for decades, strikingly little effort has been applied to determining how the post-translational modifications (PTMs) of these proteins may change during the course of lactation. PTMs, particularly glycosylation, can greatly influence protein structure, function, and stability and can particularly influence the gut where their degradation products are potentially bioactive. In this work, previously undiscovered temporal variations in both expression and glycosylation of the glycoproteome of human milk are observed. Lactoferrin, one of the most abundant glycoproteins in human milk, is shown to be dynamically glycosylated during the first 10 days of lactation. Variations in expression or glycosylation levels are also demonstrated for several other abundant whey proteins, including tenascin, bile salt-stimulated lipase, xanthine dehydrogenase, and mannose receptor.

KEYWORDS: Breast milk; glycoprotein; lactation; milk; post-translational modification

INTRODUCTION

Breast milk is the remarkable product of the mammary gland emerging from evolutionary pressure as the exclusive food for the growing infant (*1*−*11*). Because it serves as the sole source of nutrition in early life, milk has been widely studied to better understand the quantities and bioavailabilities of essential nutrients. Milk not only delivers nutrients but also provides the infant with a number of physiological advantages such as enhanced immunity (*9*, *12*) and the development and maintenance of a complex microflora in the gastrointestinal tract (*11*, *13*−*15*). In light of these varied and extensive benefits, it is critical to study the structure−function relationship of different components in milk and how they confer advantages to the infant.

Proteins make up a major component of human milk. Proteins as a whole are the fourth most abundant component of human milk after lactose, lipids, and oligosaccharides. Total protein concentrations are typically in the 10−20 mg/mL range and have been shown to decrease during the lactation period, reaching a minimum after approximately 1 week. This change is also correlated with a drastic increase in total milk production and a change in overall protein composition. A more detailed study of these dynamic behaviors is especially interesting because milk proteins and their degradation products have been associated with a number of important biological activities. These include immunomodulatory (*16*), antibacterial (*17*), and pathogen-binding (*18*) activities.

The rapidly developing infant requires a dynamic and complex source of nutrients, and the overall health, development, and protection of the infant are enhanced by a variety of bioactivities also present in milk. To understand these activities, research in this group has examined the behavior of important components in human milk during the course of the lactation period. We have recently quantified human milk oligosaccharides (HMOs) during the lactation period in these samples and have found them to remain relatively constant for all components (*19*) during lactation. One proposed role of HMOs is as a prebiotic (*15*, *20*), and the constant production of oligosaccharides is consistent with that role.

Variation of protein expression is similarly important given the many roles of proteins in the infant diet. While thorough conventional proteomic studies have been performed on both the whey (*4*, *21*) and the milk fat globule membrane (*22*, *23*) fractions of human milk, there are very few studies that address any variation in post-translational modifications. Even fewer studies address milk protein glycosylation, although there have been a few notable exceptions. In 2002, Charlwood et al. (*23*) described the N-glycosylation of four abundant milk fat globule membrane proteins. This study interrogated samples from a wide range of lactation stages, yet reported neither any temporal variations in glycan abundance that are apparent in this study nor variations in protein abundances.

While there have been studies establishing the variation in both total protein quantity and protein composition in human milk, little progress has been made in determining any qualitative or quantitative variations in glycosylation of the major milk glycoproteins. Protein glycosylation is one of the most commonly occurring post-translational modifications and directly affects protein structure, function, and recognition. Protein glycosylation is particularly relevant to milk, as it has been shown to have

*To whom correspondence should be addressed. Tel: 1-530-752-6364. Fax: 1-530-754-5609. E-mail: cblebrilla@ucdavis.edu.

 Published on Web 04/23/2010

direct implications in pathogen binding (*8*), infant development, and proteolytic susceptibility (*24*).

In this work, glycosylation of specific milk proteins is shown to vary during lactation. The work presented here both outlines a strategy for rapidly determining gross protein glycosylation profiles over time and demonstrates that there are dynamically glycosylated glycoproteins in human milk. Dynamic protein glycosylation suggests many interesting biological roles for individual proteins. An extensive review has been published on the potential roles of protein-linked glycosylation (*25*). As changes in protein glycosylation can significantly affect protein function, stability, and structure (*26*−*28*), tracking these changes could facilitate the discovery of additional bioactive species in milk.

The degradation products from these compounds may be equally as important as the intact protein, as they have been shown to perform various biological functions both in vivo and in vitro (*8*, *12*, *29*, *30*). Whereas glycosylation will greatly influence the function, identity, and quantities of these degradation products, a more thorough understanding of temporal variations of protein glycosylation is clearly necessary. As the neonate's intestinal microflora, immune system, and digestive systems change the most rapidly during the first few days of life, understanding changes in protein expression and glycosylation are critically important.

MATERIALS AND METHODS

Materials. Organic solvents [all high-performance liquid chromatography (HPLC) grade or higher] were purchased from Burdick and Jackson. Sequencing-grade trypsin (modified by reductive methylation to reduce autolysis) was purchased from Promega. Bisacrylamide gels, Bradford reagent, and Coomassie brilliant blue G-250 were purchased from Bio-Rad. All water used was 18 MΩ deionized water. Pro-Q emerald 300 was purchased from Invitrogen. Glycerol-free Peptide: *N*-glycosidase F (PNGaseF) was purchased from New England Biolabs. Porous graphitized carbon cartridges were obtained from Glygen. All other reagents were purchased from Sigma-Aldrich.

Samples. Human milk samples were obtained from four healthy women. Overall, samples from days 1, 2, 5, 10, 15, 16, 17, and 30, 31, or 32 of lactation were interrogated in this study, although none of the individuals provided milk for all time points. Quantitative comparisons were made with a minimum of three samples for each time point, and days 15, 16, and 17 were grouped for this comparison as were days 30, 31, and 32. All milk samples were manually expressed and immediately frozen. Samples were then transferred to a −80 °C freezer within 3 h and stored there until analysis.

Protein Extraction. One half milliliter of raw milk was centrifuged at 4 °C for 30 min, and the fat and cellular layers were removed. Residual lipids were removed by the method of Wessel and Flugge (*31*). Briefly, three volumes (1.5 mL) of 2:1 chloroform/methanol was added, the solution was agitated, and the supernatant was retained. An ethanol precipitation of protein from the supernatant was performed overnight at 4 °C by adding 5 mL of HPLC-grade ethanol, and following centrifugation, the supernatant was removed. Precipitates were resuspended in 50 mM ammonium bicarbonate buffer (pH 7.5), and protein quantities were determined by the Bradford method. The precipitated protein was stored at −20 °C until analysis.

Soduim Dodecyl Sulfate−Polyacrylamide Gel Electrophoresis (SDS-PAGE). An aliquot containing 10 μg of protein was mixed 1:1 (v/v) with Laemmli sample buffer containing 350 mM dithiothreitol (DTT) (*32*) and was denatured and reduced by 2 min of heating at 95 °C. SDS-PAGE separations were achieved with discontinuous gradient 10−20% bisacrylamide gels and constant 20 mA running conditions. Following separations, gels were washed four times for 10 min with water. Subsequent staining with Coomassie brilliant blue and the glycoprotein-specific Pro-Q Emerald 300 was performed in parallel on duplicate gels according to the manufacturer's instructions. Gels were scanned using an HP ScanJet 4890 scanner (in the case of Coomassie-stained gels) or a Bio-Rad transilluminating scanner (in the case of Pro-Q Emerald-stained gels).

In-Gel Trypsin Digestion. Selected gel bands were thoroughly washed with deionized water and gently agitated for a total of 1 h. Gel pieces were cut into 1 mm wide cubes and dried in a vacuum centrifuge. Washed, dehydrated gel pieces were incubated with 10 mM DTT at 55 °C for 1 h and 55 mM iodoacetamide (IoAA) at room temperature in the dark for 45 min. Gel pieces were then washed with 100 mM ammonium bicarbonate with gentle agitation for 10 min and briefly dehydrated with acetonitrile. This wash step was repeated twice, and gel pieces were dried. Proteins were then digested with 0.5 μg of trypsin in 100 mM ammonium bicarbonate at 37 °C for 16 h. Following digestion, peptides were extracted with acetonitrile, water, and 50% ethanol and dried in a vacuum centrifuge. Samples were reconstituted with deionized water, desalted using a zip-tip C_{18} microtip, and eluted into 5 μL of 50% acetonitrile. One microliter was used for subsequent matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) analysis, as described below.

Peptide Mass Fingerprinting Analysis with MALDI-MS. Tryptic peptides were analyzed on an IonSpec ProMALDI 7.0 T Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) equipped with an external MALDI source and capable of external ion accumulation with vibrational cooling. All spectra were internally calibrated using the InCAS technique (*33*) and a mixture of standard peptides to increase mass accuracy. The peptides used were Bradykinin fragment peptides 1−7, angiotensin II, $P_{14}R$, human adrenocorticotropic hormone fragment peptides 18−39, and oxidized B chain of bovine insulin. Protonated, monoisotopic masses for the calibrant peptides were m/z 757.39915, 1046.54179, 1533.85765, 2465.19833, and 3494.65077, respectively. Spectra were deisotoped using the IonSpec PeakHunter software, and the monoisotopic mass lists were then relieved of known contaminants and nonpeptide masses using the Mass Sieve approach (*34*).

MASCOT Database Searching. Processed mass lists were submitted to MASCOT (*35*) for protein identification. Carbamidomethylation of cysteine residues was selected as a fixed modification. No variable modifications were included in the searches. Peptide mass tolerances were set to 10 ppm. All proteins were identified above the 95% confidence interval (CI), and mass errors for all assigned peptides were less than 5 ppm. Each band identified only one protein (discounting splice variants) at that confidence level.

Liquid Chromatography/Tandem Mass Spectrometry (LC-MS/MS) Analyses. A total of 11 samples were examined via LC-MS/MS. One hundred micrograms of extracted milk proteins was dried and solubilized with 60 μL of 8 M urea. Samples were then reduced with DTT and alkylated with IoAA. After dilution in 180 μL of water, an overnight digestion with trypsin was performed. Peptides were then concentrated and desalted using C_{18} zip-tip before LC separation and online MS/MS.

A nanoLC-2D system (Eksigent, Dublin, CA) coupled with an LTQ ion trap mass spectrometer (Thermo Finnigan) was used with a homemade fritless reverse phase microcapillary column (75 μm × 180 mm; packed with Magic C18AQ, 3 μm 100 Å: Michrom Bio Resources) and vented column configuration. Digested samples were transferred from the autosampler to the online trap column (0.15 mm × 20 mm; packed with Magic C18AQ, 3 μm 100 Å) and desalted. Peptides were eluted from the trap and separated on the capillary column using a reverse-phase gradient at a flow rate of 300 nL/min and directly electrosprayed into the mass spectrometer. A cycle of one MS survey scan followed by 10 MS/MS scans was repeatedly acquired over the LC gradient. Dynamic exclusion for 1 min duration was utilized. Buffers were 0.1% formic acid in water (buffer A) and 0.1% formic acid in acetonitrile (buffer B). A 107 min gradient (2−40% B for 95 min, followed by 40−80% B for 12 min) was used.

Protein identification based on LC-MS/MS was performed using X! Tandem with a fragment ion mass tolerance of 0.40 Da and a parent ion tolerance of 1.8 Da. IoAA derivatization of cysteine was specified as a fixed modification. Deamidation of asparagine and glutamine, oxidation of methionine and tryptophan, sulfonation of methionine, tryptophan oxidation to formylkynurenin of tryptophan, acetylation of lysine and the N-terminus, and phosphorylation of serine, threonine, and tyrosine were specified as variable modifications.

Scaffold software was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide

Prophet algorithm (*36*). Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet (*37*).

In-Gel Deglycosylation and MALDI FTICR-MS Analysis. To verify changes in lactoferrin glycosylation, 10 μg of milk protein corresponding to lactation days 2 and 5 was reduced, alkylated, and loaded on to 10% polyacrylamide gels (86 mm × 68 mm × 1 mm). After separation, protein visualization was performed using Bio-Safe Coomassie blue staining. Lactoferrin gel bands were excised and destained with 5% methanol−7% acetic acid in water and frozen overnight. After extensive washing with acetonitrile and 50 mM NH_4HCO_3, glycans were released using 750 units of PNGaseF at 37 C for 16 h. Released glycans were extracted from gel pieces by washing them with 200 μL of water, followed by 200 μL of acetonitrile, with a 10 min sonication step following each addition. Supernatants obtained in each step were recovered and combined. This wash cycle was repeated three times, and supernatants were dried down in a vacuum centrifuge.

Samples were reconstituted in deionized water, and oligosaccharides were purified by solid-phase extraction using porous graphitized carbon cartridges. Hypercarb carbon cartridges were conditioned with 3 column volumes of deionized water, followed by 3 volumes of 80% acetonitrile in 0.10% aqueous trifluoroacetic acid (TFA) (v/v) and with another 3 volumes of deionized water. The samples were loaded onto the cartridge, incubated for 10 min at room temperature, washed with 3 volumes of deionized water, eluted with 3 volumes of 40% acetonitrile and 0.1% TFA in water (v/v), and dried in vacuo. Glycans were reconstituted in 5 μL of water. MALDI Fourier transform ion cyclotron resonance (FTICR)-MS was then performed as described previously, using a mixture of maltooligosaccharides as an external calibrant (*38*). All oligosaccharide compositions were identified by accurate mass, utilizing a 10 ppm tolerance, and five replicate spectra were acquired for each sample.

RESULTS

Gel Electrophoretic Analyses of Milk Proteins. Gel electrophoresis was used to determine semiquantitative changes in expression of the major milk proteins. Samples from days 1, 15, and 30 of lactation were examined in these studies. The total protein amount was determined, and 10 μg aliquots were investigated via SDS-PAGE. Shown in **Figure 1** are representative SDS-PAGE analyses. The 10 abundant proteins indicated were identified using in-gel digestion and peptide mass fingerprinting. The assignments are summarized in **Figure 1** and **Table 1**, along with pertinent identification parameters such as the number of peptides matched (n_P), the probability of a stochastic assignment, and the score of the next best protein identification (ΔS_M). For reference, a significant match at the 95% CI corresponds to a score of 65. All proteins were identified with scores of 69 or higher under the Mascot protocol. In addition, these results are characterized by large values of ΔS_M, which is indicative of a high level of discrimination and a low likelihood of false positive identification afforded by accurate mass analysis (*39*).

To explore the glycosylation status of the abundant milk proteins, gels were run in duplicate and stained with colloidal Coomassie brilliant blue G-250 (for total protein) and Pro-Q Emerald 300 (for glycan content) stains separately. To obtain quantitative information, each gel lane was examined by an in-house image-processing program. Several proteins were shown to change in total Coomassie signal intensity during lactation. These trends are demonstrated in **Figure 2**, which shows an overlay of three SDS-PAGE electropherograms generated from a single individual at differing lactation times. Upon examination, it is clear that the expression level of the heavy chain of IgA (immunoglobulin α-1 chain C, IGHA1) is decreasing for this mother over the first lactation month, as expected. In addition to IGHA1, there are four other proteins detailed below whose abundances are found to vary during the first month of lactation.

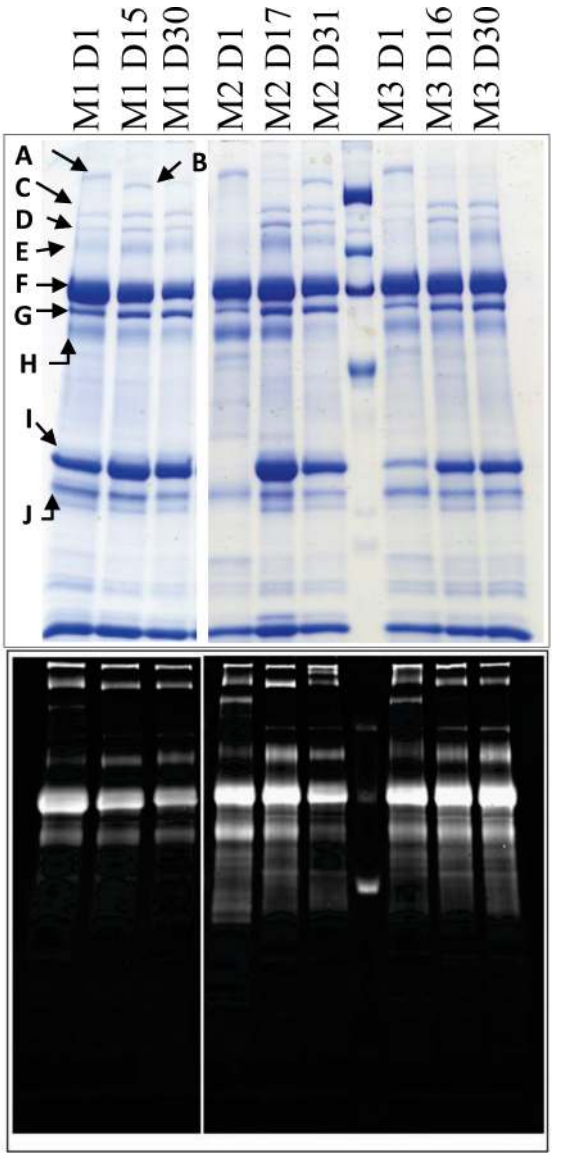

Figure 1. Coomassie (**top**) or glycoprotein-specific (**bottom**) stained SDS-PAGE separations for three mothers at varying time points. Indicated protein bands identified by peptide mass fingerprinting are annotated in **Table 1**, along with several measures of the quality of the identification. Abbreviations: M, mother code; D, lactation day; and MWM, molecular weight markers. Markers for Coomassie stain are, from top, 203, 120, 95, and 49 kDa. For pro-Q stain, markers are 180, 82, and 42 kDa. Protein K is α-lactalbumin.

All quantitative comparisons of protein and glycan-level variations were made using background-subtracted integrals of Coomassie- or pro-Q-stained electrophoretic peaks, respectively. The dynamic behaviors of five glycoproteins including tenascin, xanthine dehydrogenase (XD), bile salt-stimulated lipase (BSSL), lactoferrin, and IGHA1 protein are shown in **Figure 3**. These proteins were selected for this comparison because they each signaled under both the Coomassie and the glycoprotein-specific stains. The upper bar graph shows the averaged glycoprotein-specific stain intensity, and the lower graph shows the averaged Coomassie response at days 1, 15, and 30 for each protein. Proteins were identified as differentially glycosylated or expressed if the response to the appropriate stain was significantly different between two time points at greater than the 90% CI. Remarkably, tenascin, IGHA1 protein, XD, and lactoferrin expression levels all decrease significantly from colostrum to mature milk in these

samples. Interestingly, XD expression peaks at 2 weeks lactation time and drops to approximately half of the peak concentration by 1 month lactation time, while lactoferrin, IGHA1, and tenascin decrease from colostrum to mature milk. Mannose receptor and Ig κ light chain both increase from day 1 to day 30 of lactation (data not shown). Mannose receptor and XD are each potential N-glycoproteins by virtue of the presence of the consensus sequence for N-glycosylation. XD responds to the glycoprotein-specific stain used in this study. There is no corresponding response from the mannose receptor.

Interestingly, there are also glycoproteins that exhibit disparate behavior under protein- or glycan-specific staining. A small section of the SDS-PAGE gel band is shown for illustrative purposes in **Figure 4**, with samples from the first 10 days of lactation shown. This detailed examination was intended to determine if any interesting dynamic behavior was missed between days 1 and 15 as examined previously. The CBB stain intensities suggest that the gross quantity of lactoferrin remained relatively constant over the first 10 days of lactation. Surprisingly, the total glycan stain intensity decreases by approximately 60% over that time. By day 15, this decrease in glycosylation is restored to near its previous levels. As one of the most abundant glycoproteins in milk, this trend represents a significant change in the total glycoconjugate content present in the infant's digestive system. Additionally, while the expression level of XD peaks at 2 weeks and decreases afterward (days 15 and 30 are different at the 95% CI), the glycosylation level is largely unchanged from 2 to 4 weeks, suggesting an increase in the degree of glycosylation of this glycoprotein during this time.

Although the CBB and glycan-specific staining methods have previously both been shown to be valid quantitative platforms (*40*,*41*), the changes in lactoferrin glycosylation were also verified

Table 1. Summary of All Proteins Identified by Peptide Mass Fingerprinting[a]

gel band	protein identity	MW (kDa)	S_M	ΔS_M	n_P	SC (%)	e_{rms}(ppm)
A	tenascin	240.6	163	135	30	24	3
B	fatty acid synthase	275.5	143	111	30	20	2
C	mannose receptor	168.9	93	61	19	19	4
D	xanthine dehydrogenase	148.4	80	42	10	12	4
E	bile salt-stimulated lipase	66.6	94	71	12	25	2
F	lactoferrin	79.8	97	69	15	29	3
G	serum albumin	67.7	178	147	22	48	1
H	Ig α-1 chain C	54.3	73	32	6	20	2
I	κ-casein	20.4	75	37	5	45	1
J	Ig κ light chain	23.5	71	N/A	5	41	2

[a] Several parameters relating to the quality of the identification are also tabulated as follows: S_M, Mowse score; ΔS_M, difference between the top two highest scoring protein identifications; n_p, number of peptides identified; SC (%); percent sequence coverage of the protein identified; and e_{RMS} (ppm), root mean square mass error of identified peptides, in parts per million. The identification for Ig κ chain C has no ΔS_M listed because all other proteins identified at a nonstochastic level were splice variants.

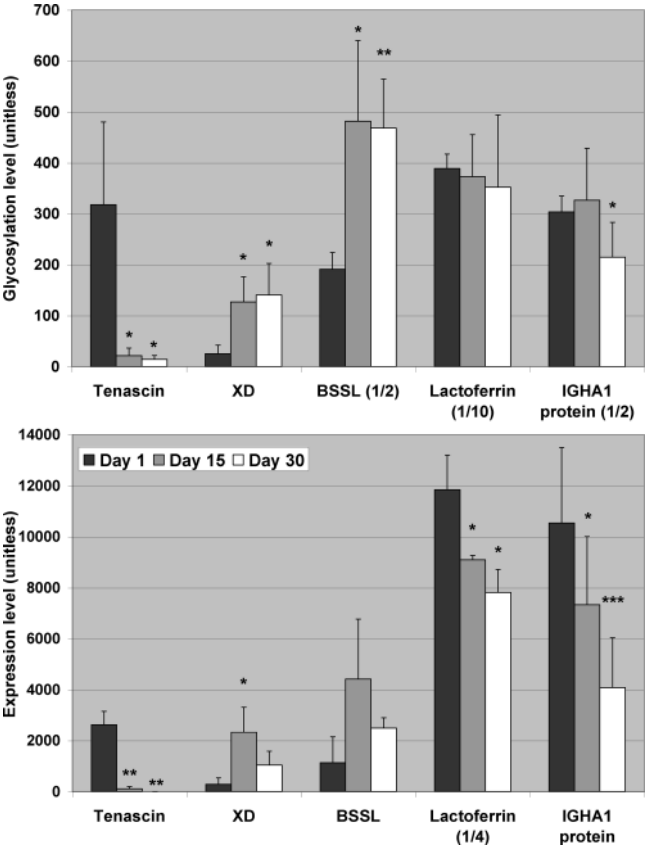

Figure 3. Glycosylation (**upper**) or expression (**lower**) levels of selected human milk glycoproteins. Singly, doubly, or triply marked entries are statistically different at the 90, 95, or 99% CI from the corresponding day 1 entry. Standard deviations for each measurement are shown as error bars. When necessary, signals were scaled by the factor in parentheses to enable direct comparison between highly and moderately abundant glycoproteins.

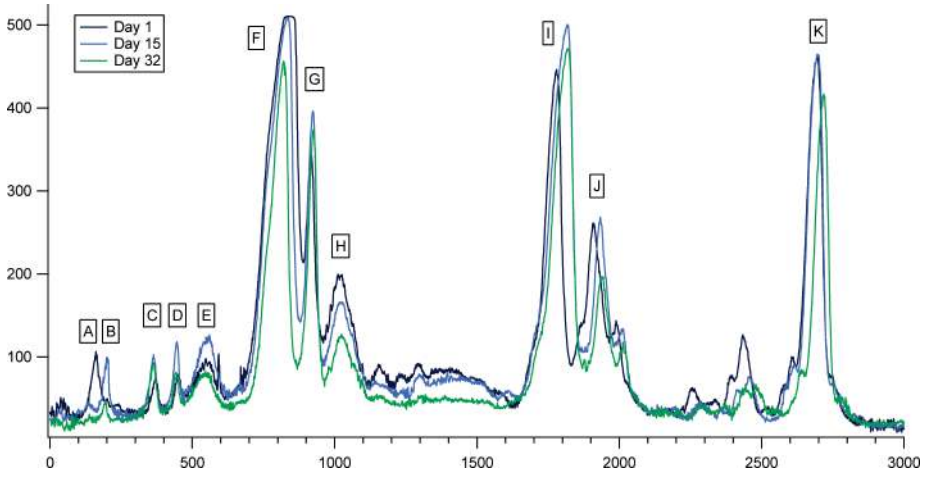

Figure 2. Electropherograms of SDS-PAGE-separated, Coomassie-stained proteins with annotated identifications as detailed in **Table 1**. Protein K is α-lactalbumin.

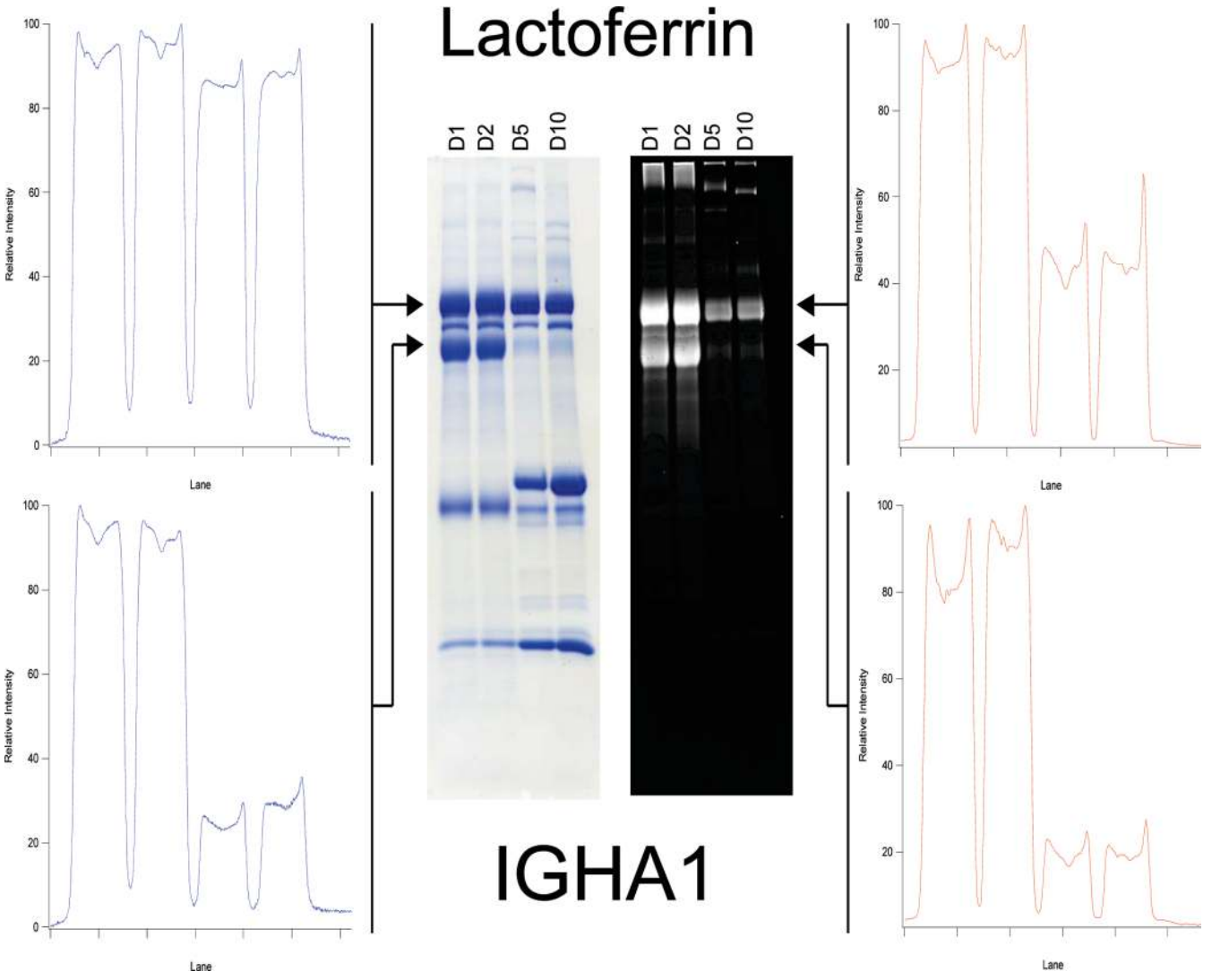

Figure 4. Image line profiles of isolated milk proteins stained with either Coomassie (left) or Pro-Q (right) stains. Proteins were isolated from a single mother on lactation days 1, 2, 5, and 10. While lactoferrin is shown to have a relatively constant total protein level (**top left**), the degree of glycosylation at day 10 is approximately 40% of that present at day 1 (**top right**). The lower two traces showing Ig α-1 chain C behavior show little variation.

by another method. Enzymatic release of *N*-glycans followed by analysis by MS was used both to verify the gross glycosylation trends that were observed via the glycan-specific visualization and to identify the compositions of the *N*-glycans. Shown in **Figure 5** are the mass spectra obtained from enzymatic release of electrophoretically separated lactoferrin from different lactation days. The glycosylation trends obtained from the Pro-Q Emerald 300 staining approximate the total glycan intensity changes observed via high-resolution, high mass accuracy FTICR-MS. The glycans observed were primarily biantennary, complex type, and fucosylated, and compositions for each peak are labeled on the spectra. A decrease in both the total level of glycosylation and the degree of fucosylation was observed between days 2 and 5.

Shotgun Proteomics Analysis of Milk Proteins. LC-MS/MS analyses were performed to determine the gross protein and glycoprotein content in milk. As expected, the shotgun analyses identified significantly more proteins than the gel-based analyses. The LC-MS/MS analyses identified a total of 55 proteins present in 11 samples analyzed. Identified proteins were searched in Swiss-Prot to determine their annotated glycosylation status. Shown in **Table 2** are the proteins identified along with their potential for glycosylation. Of the 55 proteins identified, 27 are glycosylated, and an additional 12 are potentially N-glycosylated based on the presence of the consensus sequence NXS or NXT, where X is any amino acid except proline. O-Glycosylation can occur at essentially any serine or threonine residue, so there is no analogous classification for potential O-glycosylated proteins, as essentially no filter exists based on amino acid sequence.

DISCUSSION

While there are a small number of published studies (< 10) that detail the proteomics of the milk globule membrane, there have been surprisingly few contemporary reports on the proteomic analysis of the whey protein in human milk. However, there are a few notable exceptions. Using a combination of casein removal, immunodepletion, PMF analysis, and MudPit (*42*), Palmer et al. identified 151 proteins in colostrum, 83 of which were identified for the first time. Interestingly, the authors identified 27 proteins that are implicated in defense, 10 that have growth-modulating activities, and several that are involved in vitamin or mineral transport, supporting the hypothesis proposed by many researchers that milk is a multifunctional food. Regrettably, because of the difficulty in attaining a sufficient quantity of high-quality colostrum samples, they were limited to a single pooled sample from 100 donors and could not examine any temporal or

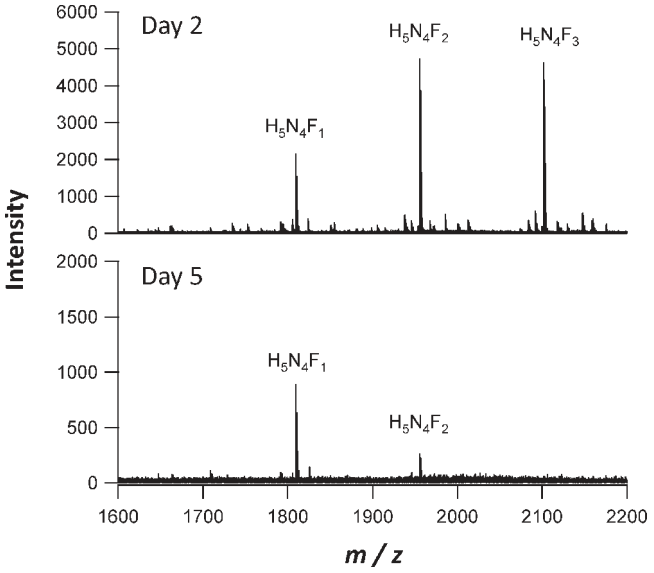

Figure 5. Variations of lactoferrin glycosylation during lactation. Representative mass spectra of enzymatically released lactoferrin *N*-glycans from lactation days 2 (**top**) and 5 (**bottom**) from a single individual are shown. Monosaccharide compositions of the major peaks are annotated as H, hexose; N, *N*-acetyl hexosamine; and F, deoxyhexose. A dramatic decrease in both quantity and diversity of glycosylation is observed between lactation days 2 and 5, as shown by the glycosylation-specific Pro-Q staining.

individual variations of milk protein. However, their study was fundamental to the emerging view of milk as a vanguard of human health and development.

The work presented here shifts the focus from what has historically been done in MS-based analyses of milk (identifying a large number of proteins) to identifying the dynamic behavior of the most abundant proteins and glycoproteins during lactation. Glycosylation has been shown to influence protein function, stability, susceptibility to proteolysis, and cell−cell interactions; therefore, changes in glycosylation could greatly influence both the function and the degradation state of glycoproteins in the infant's gut. As the degradation of the most abundant glycoproteins will have the greatest impact on the identity of potentially bioactive glycopeptides, tracking glycosylation changes in these species is essential.

Of the 10 abundant proteins identified via PMF in this study, eight are glycosylated or potentially glycosylated. Five proteins (tenascin, BSSL, Lactoferrin, Ig A, and κ-casein) are well-characterized glycoproteins. In addition, α-lactalbumin is highly abundant but seldom (∼1%) glycosylated (*43*), and XD was only recently identified as a glycoprotein (*44*). Most of these abundant glycoproteins are also found to change in concentration during the first month of lactation. Only one of these has previously been shown to vary in glycosylation during the course of lactation, namely, BSSL (*45*). Its glycosylation was shown to vary greatly between the first and the sixth month of lactation, both in absolute quantity of monosaccharide residues and in identity of glycans. BSSL was found to eventually be replaced by a completely nonglycosylated analogue during late lactation.

In this study, we examine glycosylation changes of BSSL during the first month and show that there is dynamic glycosylation during this time as well. While the protein amount does not show significantly different changes in the samples examined here, the glycosylation amount does, showing a 2-fold increase from day 1 to 15 of lactation. This increase is maintained until day 30, despite an apparent decrease in the total amount of BSSL expression at day 30 of lactation. Whereas BSSL glycosylation will facilitate transport of the active form of the enzyme through the infant's stomach, these changes may influence the efficiency of lipid digestion and absorption, a critical energy source for the neonate. BSSL glycans contain Lewis epitopes and have been shown to contribute to host defense both through the presence of these epitopes and by virtue of the enzyme's lipolytic activity (*46*). Therefore, BSSL glycosylation affects host defense both directly (via acting as a receptor analogue) and indirectly by preserving enzymatic activity of BSSL in the infant's digestive tract.

Ig A is one of the most abundant glycoproteins in colostrum but decreases to much lower levels a few days after the onset of lactation. Ig A is commonly believed to augment the infant's own nascent immune system. Gross glycosylation profiles of Ig A vary in the same manner as the total protein expression, meaning that the degree of Ig A glycosylation shows little variation. Despite the apparent lack of dynamic Ig A glycosylation, the decrease in protein amount leads to a dramatic reduction of the total glycoconjugate during early lactation.

Lactoferrin presents a compelling glycosylation transformation. During the first 10 days of lactation, it has been well-established that both the concentration of lactoferrin and the total protein concentration decrease in similar fashion. In the present study, a constant mass of protein was analyzed for each sample, and the amount of lacroferrin was unchanged. Despite a constant amount of lactoferrin analyzed, we have found a significant change in both the identity and the extent of glycosylation present. These changes in glycosylation provide another dimension to this well-studied glycoprotein that has not been examined in any detail, and at this point, biological implications of this change can only be hypothesized. Compellingly, lactoferrin's digestion products include peptides that have been shown to be antimicrobial (*30*), and as glycosylation has been shown to influence lactoferrin's susceptibility to proteolysis (*47*), the modification could directly influence the generation of these antimicrobial species. It has also been shown that the identity of oligosaccharide present on lactoferrin can affect the efficiency of brush border binding and iron transport in vitro (*48*), so changes in glycosylation during lactation will directly affect both the identity of the microbiota and the availability of essential minerals in the infant's digestive system via regulation of the degree of digestion of lactoferrin during early life. In addition to demonstrating an overall drop in degree of glycosylation, FTICR-MS-based analyses show a decrease in *N*-glycan fucosylation between days 2 and 5.

On the basis of global LC-MS/MS analyses, we estimate that 70% of the abundant proteins identified here are likely glycosylated, which is larger than estimates of overall glycosylation in human serum proteins (*49*). What makes human milk especially unique however is the large amount of total glycosylation in the proteins. With 8 of the 10 most abundant proteins being glycosylated, the absolute quantity of protein-linked oligosaccharide is extremely high. Combined with the highly abundant HMOs (present at 10−20 g/L), there is a predominance of indigestible carbohydrate present in early lactation. Evolutionary pressures suggest that these compounds are present to increase the survival of the infant, as the mother is synthesizing and excreting them at significant personal cost. While models have been established for the presence of HMO, more research is needed to elucidate all of the roles of protein-linked glycosylation. However, the dynamic nature of this glycosylation suggests that it adjusts to meet the needs of the infant and may suggest what roles these glycans may play biologically.

Table 2. All Proteins Identified in the LC-MS/MS Experiments, Listed with Corresponding Glycosylation Status as Annotated in the SwissProt Database[a]

protein name	MW (kDa)	Swiss-Prot no.	# UP	seq coverage (%)	glycosylation status
actin, cytoplasmic 1	42	P63261	4	12	X
adipose differentiation-related protein	48	Q99541	2	11	P
α-1 antichymotrypsin	48	Q2TU88	4	17	N
α-1 antitrypsin	47	P01009	6	23	N
α-lactalbumin	16	P00709	21	68	N
α-s1 casein	22	P47710	10	51	X
antistreptococcal Ig λ chain variable region	10	Q96SB0	2	31	P
apolipoprotein A-1	31	P02647	3	16	X
β-1,4-galactosyltransferase 1	44	P15291	6	28	P
β-2-microglobulin	14	P61769	2	27	X
β-casein	25	P05814	31	92	X
bile salt-stimulated lipase	79	P19835	17	29	N, O
butyrophilin, subfamily 1 member A1	59	Q4VAN2	5	14	P
CD59 glycoprotein	14	P13987	2	20	N, O
chordin-related protein 2 variant III	50	Q6WN33	5	20	P
clusterin	58	Q5HYC1	10	25	N
complement C3	187	P01024	11	11	N
complement C4	193	P0C0L4	2	1	N
fatty acid synthase	273	P49327	26	20	P
fatty acid-binding protein	15	P05413	4	28	X
gollistatin-related protein-1	35	Q12841	3	11	N
fructose-1,6-biphosphate adolase A	39	Q6FI10	2	13	X
galectin-3-binding protein	65	Q08380	5	18	X
γ-enolase	47	P09104	2	11	P
haptoglobin	38	P00738	2	8	N
Ig κ chain V−II	13	P06310	2	35	X
Ig κ chain V−III	13	P18135	3	34	X
Ig λ constant 2	22	Q8N355	8	40	X
IgG heavy chain locus; IGHG1	54	Q569J1	19	38	N
IGHG1 immunoglobulin heavy constant γ 1	41	Q6N030	9	37	N
immunoglobulin heavy chain variant (fragment)	37	Q9NPP6	4	16	P
immunoglobulin J chain	18	P01591	4	25	N
immunoglobulin κ variable 1−5	23	Q6GMX8	11	44	X
insulin-like growth factor-binding protein precursor 2	35	P18065	7	19	X
κ-casein	20	P07498	13	55	O
lactadherin	43	Q08431	9	33	P
lactoferrin	78	P02788	66	81	N
leucine-rich α-2-glycoprotein	38	P02750	2	14	N, O
lipoprotein lipase	53	P06858	4	17	P
lysozyme C	17	P61626	8	40	X
mannose receptor, C-type	166	Q5VSK2	6	6	P
monocyte differentiation antigen CD14	40	P08571	4	20	N
oroscomucoid	24	P02763	4	20	N, O
osteopontin	35	P10451	11	25	N, O
platelet glycoprotein 4	53	P16671	4	29	N
polymeric Ig receptor	83	P01833	45	13	N
prolactin-inducible protein	17	P12273	2	47	N
selenium binding protein 1	52	Q13228	2	19	X
serotransferrin	77	P02787	8	21	N
serum albumin	69	P02768	35	65	X
similar to Ig γ-2 chain C region	22	Q49AS2	4	46	N
tenascin C	231	P24821	40	30	N
trefoil factor 3	14	Q07654	2	14	X
XD/oxidase	146	P47989	8	1	N
zinc α-2-glycoprotein 1	34	P25311	7	30	N

[a] Abbreviations: MW, molecular weight in kDa; #UP, number of unique peptides; and seq coverage, % sequence coverage of identified protein. Abbreviations for glycosylation status are as follows: N, known N-linked glycoprotein; O, known O-linked glycoprotein, N, O, known N-and O- linked glycoprotein; P, potential N-linked glycoprotein; and X, no potential N-glycosylation and no known O-linked glycans. Proteins are identified as potentially N-glycosylated if they possess the consensus amino acid sequence required for N-glycosylation, NX(S/T), where X is any amino acid except proline.

In summary, milk glycoproteins are both prevalent and abundant. Contrary to the view of milk proteins functioning solely as a source of amino acids for digestion and mineral absorption, the predominance of protein glycosylation implies a more structure-specific role. Research to date has provided evidence that these glycoproteins and their digestive products provide an important source of bioactive compounds with diverse beneficial properties. We have demonstrated in this work that milk presents a dynamic glycoprotein component to the infant. Compellingly, several abundant glycoproteins are shown to vary in expression and glycosylation in an independent fashion.

ABBREVIATIONS USED

BSSL, bile salt-stimulated lipase; CI, confidence interval; HMOs, human milk oligosaccharides; IGHA1, immunoglobulin alpha-1 chain C; XD, xanthine dehydrogenase; PNGase F, peptide: *N*-glycosidase F.

ACKNOWLEDGMENT

We thank the UC Davis proteomics staff for assistance with the shotgun experiments. Also, thanks go to Dr. Milady Niñonuevo, Caroline Chu, and Larry Lerno for helpful advice. Thanks to Scott Kronewitter for assistance with the image analysis software.

LITERATURE CITED

(1) Binns, C.; Lee, A.; Harding, H.; Gracey, M.; Barclay, D. The CUPDAY Study: Prebiotic-probiotic milk product in 1−3-year-old children attending childcare centres. *Acta Paediatr*. **2007**, *96* (11), 1646–1650.

(2) Colome, G.; Sierra, C.; Blasco, J.; Garcia, M. V.; Valverde, E.; Sanchez, E. Intestinal permeability in different feedings in infancy. *Acta Paediatr*. **2007**, *96* (1), 69–72.

(3) Kokinopoulos, D.; Photopoulos, S.; Varvarigou, N.; Kafegidakis, L.; Xanthou, M. The effect of human milk, protein-fortified human milk and formula on immunologic factors of newborn infants. *Adv. Exp. Med. Biol*. **1991**, *310*, 77–85.

(4) Kverka, M.; Burianova, J.; Lodinova-Zadnikova, R.; Kocourkova, I.; Cinova, J.; Tuckova, L.; Tlaskalova-Hogenova, H. Cytokine profiling in human colostrum and milk by protein array. *Clin. Chem*. **2007**, *53* (5), 955–962.

(5) Lonnerdal, B. Nutritional and physiologic significance of human milk proteins. *Am. J. Clin. Nutr*. **2003**, *77* (6), 1537S–1543S.

(6) Lonnerdal, B.; Lien, E. L. Nutritional and physiologic significance of alpha-lactalbumin in infants. *Nutr. Rev*. **2003**, *61* (9), 295–305.

(7) Moro, G. E.; Arslanoglu, S. Reproducing the bifidogenic effect of human milk in formula-fed infants: Why and how? *Acta Paediatr. Suppl*. **2005**, *94* (449), 14–17.

(8) Newburg, D. S. Human milk glycoconjugates that inhibit pathogens. *Curr. Med. Chem*. **1999**, *6* (2), 117–127.

(9) Newburg, D. S.; Ruiz-Palacios, G. M.; Morrow, A. L. Human milk glycans protect infants against enteric pathogens. *Annu. Rev. Nutr*. **2005**, *25*, 37–58.

(10) Packard, V. S. *Human Milk and Infant Formula*; Academic Press: New York, 1982; p x, 269 pp.

(11) Parracho, H.; McCartney, A. L.; Gibson, G. R. Probiotics and prebiotics in infant nutrition. *Proc. Nutr. Soc*. **2007**, *66* (3), 405–411.

(12) Hakansson, A.; Zhivotovsky, B.; Orrenius, S.; Sabharwal, H.; Svanborg, C. Apoptosis induced by a human milk protein. *Proc. Natl. Acad. Sci. U.S.A*. **1995**, *92* (17), 8064–8068.

(13) Ziegler, E.; Vanderhoof, J. A.; Petschow, B.; Mitmesser, S. H.; Stolz, S. I.; Harris, C. L.; Berseth, C. L. Term infants fed formula supplemented with selected blends of prebiotics grow normally and have soft stools similar to those reported for breast-fed infants. *J. Pediatr. Gastroenterol. Nutr*. **2007**, *44* (3), 359–364.

(14) Ward, R. E.; Ninonuevo, M.; Mills, D. A.; Lebrilla, C. B.; German, J. B. In vitro fermentation of breast milk oligosaccharides by Bifidobacterium infantis and Lactobacillus gasseri. *Appl. Environ. Microbiol*. **2006**, *72* (6), 4497–4499.

(15) Boehm, G.; Stahl, B.; Jelinek, J.; Knol, J.; Miniello, V.; Moro, G. E. Prebiotic carbohydrates in human milk and formulas. *Acta Paediatr. Suppl*. **2005**, *94* (449), 18–21.

(16) Moriuchi, M.; Moriuchi, H. A milk protein lactoferrin enhances human T cell leukemia virus type I and suppresses HIV-1 infection. *J. Immunol*. **2001**, *166* (6), 4231–4236.

(17) Li, Y. M.; Tan, A. X.; Vlassara, H. Antibacterial activity of lysozyme and lactoferrin is inhibited by binding of advanced glycation-modified proteins to a conserved motif. *Nat. Med*. **1995**, *1* (10), 1057–1061.

(18) Naarding, M. A.; Dirac, A. M.; Ludwig, I. S.; Speijer, D.; Lindquist, S.; Vestman, E. L.; Stax, M. J.; Geijtenbeek, T. B.; Pollakis, G.; Hernell, O.; Paxton, W. A. 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), 3367–3374.

(19) Ninonuevo, M. R.; Perkins, P. D.; Francis, J.; Lamotte, L. M.; LoCascio, R. G.; Freeman, S. L.; Mills, D. A.; German, J. B.; Grimm, R.; Lebrilla, C. B. Daily variations in oligosaccharides of human milk determined by microfluidic chips and mass spectrometry. *J. Agric. Food Chem*. **2008**, *56* (2), 618–626.

(20) Coppa, G. V.; Bruni, S.; Morelli, L.; Soldi, S.; Gabrielli, O. The first prebiotics in humans: Human milk oligosaccharides. *J. Clin. Gastroenterol*. **2004**, *38* (6 Suppl.), S80–S83.

(21) Palmer, D. J.; Kelly, V. C.; Smit, A. M.; Kuy, S.; Knight, C. G.; Cooper, G. J. Human colostrum: Identification of minor proteins in the aqueous phase by proteomics. *Proteomics* **2006**, *6* (7), 2208–2216.

(22) Cavaletto, M.; Giuffrida, M. G.; Conti, A. The proteomic approach to analysis of human milk fat globule membrane. *Clin. Chim. Acta* **2004**, *347* (1−2), 41–48.

(23) Charlwood, J.; Hanrahan, S.; Tyldesley, R.; Langridge, J.; Dwek, M.; Camilleri, P. Use of proteomic methodology for the characterization of human milk fat globular membrane proteins. *Anal. Biochem*. **2002**, *301* (2), 314–324.

(24) Brines, R. D.; Brock, J. H. The effect of trypsin and chymotrypsin on the in vitro antimicrobial and iron-binding properties of lactoferrin in human milk and bovine colostrum. Unusual resistance of human apolactoferrin to proteolytic digestion. *Biochim. Biophys. Acta* **1983**, *759* (3), 229–235.

(25) Varki, A. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **1993**, *3* (2), 97–130.

(26) Wormald, M. R.; Dwek, R. A. Glycoproteins: Glycan presentation and protein-fold stability. *Structure* **1999**, *7* (7), R155–R160.

(27) O'Connor, S. E.; Imperiali, B. Modulation of protein structure and function by asparagine-linked glycosylation. *Chem. Biol*. **1996**, *3* (10), 803–812.

(28) Imperiali, B.; Rickert, K. W. Conformational implications of asparagine-linked glycosylation. *Proc. Natl. Acad. Sci. U.S.A*. **1995**, *92* (1), 97–101.

(29) Lopez-Exposito, I.; Recio, I. Protective effect of milk peptides: antibacterial and antitumor properties. *Adv. Exp. Med. Biol*. **2008**, *606*, 271–293.

(30) Jones, E. M.; Smart, A.; Bloomberg, G.; Burgess, L.; Millar, M. R. Lactoferricin, a new antimicrobial peptide. *J. Appl. Bacteriol*. **1994**, *77* (2), 208–214.

(31) Wessel, D.; Flugge, U. I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem*. **1984**, *138* (1), 141–143.

(32) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227* (5259), 680–685.

(33) O'Connor, P. B.; Costello, C. E. Internal calibration on adjacent samples (InCAS) with Fourier transform mass spectrometry. *Anal. Chem*. **2000**, *72* (24), 5881–5885.

(34) Dodds, E. D.; An, H. J.; Hagerman, P. J.; Lebrilla, C. B. Enhanced peptide mass fingerprinting through high mass accuracy: Exclusion of non-peptide signals based on residual mass. *J. Proteome Res*. **2006**, *5* (5), 1195–1203.

(35) Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20* (18), 3551–3567.

(36) Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem*. **2002**, *74* (20), 5383–5392.

(37) Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem*. **2003**, *75* (17), 4646–4658.

(38) Clowers, B. H.; Dodds, E. D.; Seipert, R. R.; Lebrilla, C. B. Dual polarity accurate mass calibration for electrospray ionization and matrix-assisted laser desorption/ionization mass spectrometry using maltooligosaccharides. *Anal. Biochem*. **2008**, *381* (2), 205–213.

(39) Dodds, E. D.; Clowers, B. H.; Hagerman, P. J.; Lebrilla, C. B. Systematic characterization of high mass accuracy influence on false

discovery and probability scoring in peptide mass fingerprinting. *Anal. Biochem.* **2008**, *372* (2), 156–166.
(40) Steinberg, T. H.; Pretty On Top, K.; Berggren, K. N.; Kemper, C.; Jones, L.; Diwu, Z.; Haugland, R. P.; Patton, W. F. Rapid and simple single nanogram detection of glycoproteins in polyacrylamide gels and on electroblots. *Proteomics* **2001**, *1* (7), 841–855.
(41) Vincent, S. G.; Cunningham, P. R.; Stephens, N. L.; Halayko, A. J.; Fisher, J. T. Quantitative densitometry of proteins stained with coomassie blue using a Hewlett Packard scanjet scanner and Scanplot software. *Electrophoresis* **1997**, *18* (1), 67–71.
(42) Washburn, M. P.; Wolters, D.; Yates, J. R., 3rd Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001**, *19* (3), 242–247.
(43) Giuffrida, M. G.; Cavaletto, M.; Giunta, C.; Neuteboom, B.; Cantisani, A.; Napolitano, L.; Calderone, V.; Godovac-Zimmermann, J.; Conti, A. The unusual amino acid triplet Asn-Ile-Cys is a glycosylation consensus site in human alpha-lactalbumin. *J. Protein Chem.* **1997**, *16* (8), 747–753.
(44) Picariello, G.; Ferranti, P.; Mamone, G.; Roepstorff, P.; Addeo, F. Identification of N-linked glycoproteins in human milk by hydrophilic interaction liquid chromatography and mass spectrometry. *Proteomics* **2008**, *8* (18), 3833–3847.
(45) Landberg, E.; Huang, Y.; Stromqvist, M.; Mechref, Y.; Hansson, L.; Lundblad, A.; Novotny, M. V.; Pahlsson, P. Changes in glycosylation of human bile-salt-stimulated lipase during lactation. *Arch. Biochem. Biophys.* **2000**, *377* (2), 246–254.
(46) Hernell, O.; Ward, H.; Blackberg, L.; Pereira, M. E. A. Killing of Giardia-Lamblia by human-milk lipases—An effect mediated by lipolysis of milk lipids. *J. Infect. Dis.* **1986**, *153* (4), 715–720.
(47) van Veen, H. A.; Geerts, M. E.; van Berkel, P. H.; Nuijens, J. H. The role of N-linked glycosylation in the protection of human and bovine lactoferrin against tryptic proteolysis. *Eur. J. Biochem.* **2004**, *271* (4), 678–684.
(48) Davidson, L. A.; Lonnerdal, B. Specific binding of lactoferrin to brush-border membrane: ontogeny and effect of glycan chain. *Am. J. Physiol.* **1988**, *254* (4 Part 1), G580–G585.
(49) Apweiler, R.; Hermjakob, H.; Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* **1999**, *1473* (1), 4–8.

Received for review January 12, 2010. Revised manuscript received March 18, 2010. Accepted April 7, 2010. Funding provided by UC Discovery, California Dairy Research Foundation, and Dairy Management Incorporated is gratefully acknowledged. This research was also supported in part by NIEHS Superfund (Grant P42 ES04699) and by the CHARGE study (Grant P01 ES11269).