



Published in final edited form as:

J Adv Dairy Res. ; 1(2): 104–. doi:10.4172/2329-888X.1000104.

Coupling Mass Spectrometry-Based “Omic” Sciences with Bioguided Processing to Unravel Milk's Hidden Bioactivities

David C Dallas^{1,2,*}, Hyeyoung Lee^{1,2}, Annabelle Le Parc¹, Juliana Maria Leite Nobrega de Moura Bell¹, and Daniela Barile^{1,2}

¹Department of Food Science & Technology, University of California, One Shields Avenue, Davis, CA 95616, USA

²Foods for Health Institute, University of California, One Shields Avenue, Davis, CA 95616, USA

Abstract

Many of milk's functional molecules could not be discovered until the right concordance of novel separation and analytical technologies were developed and applied. Many health-promoting components still await discovery due to technical challenges in their identification, isolation and testing. As new analytical technologies are assembled, new functional milk molecules will be discovered. Bovine milk is a source of a wide array of known bioactive compounds from a variety of molecular classes, including free glycans, lipids, glycolipids, peptides, proteins, glycoproteins, stem cells and microRNA. Because milk is such a complex mixture, when analyzed without fractionation or purification, many components mask the analytical signal of others, so some components cannot be detected. Modern analytics allow for the discovery and characterization of hundreds of novel milk compounds with high-resolution and high-accuracy. Liquid chromatography paired with electrospray ionization allows the separation of peptides, glycans and glycolipids for improved mass spectrometric detection. Target proteins and glycoproteins can now be purified from intact milk or other dairy streams by chromatography in order to better characterize these proteins for new bioactivities. The combination of advanced analytics with the new engineering capabilities will allow for high molecular resolution and separation techniques that can be scaled-up to semi-industrial and industrial scale for translation of lab-based discoveries. Bioguided analysis and design of dairy processing side streams will result in the transformation of waste into isolated functional ingredients to add value to dietary products.

Keywords

Milk; Bioactive Molecules; Mass Spectrometry; Peptide; Glycan; Oligosaccharide; Glycolipid; Glycoprotein; Purification; Bioguided Processing

Copyright: © 2013 Dallas DC, et al.

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

*Corresponding author: David Dallas, Department of Food Science & Technology, University of California, One Shields Avenue, Davis, CA, 95616, USA, Tel: 530-752-4137; Fax: 530-752-4759; dcdallas@ucdavis.edu.

Introduction

Milk is the product of over 200 million years of mammalian evolution [1]. 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. As all components in milk are derived from the mother's own stores, milk likely contains only molecules that support the health of the infant. For this reason, milk is an incredible source for discovery of functional molecules. As a whole, milk supports growth, immune development, neurological development and microbial development of the infant [2].

Over the years, hundreds of compounds from a variety of molecular classes have been identified in milk. Novel milk molecules continue to be discovered. For each new molecular class, novel detection, separation and analysis methods must be established. The number of functions attributed to milk components is ever increasing. For example, individual classes of molecules in milk are known to limit pathogen growth [3,4], enhance commensal bacterial growth [5], modulate the immune system [6,7], decrease inflammation [8], and to have anti-viral [9] and anti-cancer activities [10]. Surprisingly, many of milk's functional compounds were discovered only recently, including naturally occurring peptides, oligosaccharides, stem cells, exosomes and microRNA. We hypothesize that many more functional molecules remain undiscovered in milk.

Milk's natural structural and compositional complexity precludes the use of a single modern analytical technique for analyzing all milk components. Rather, complete analysis of milk's components requires a combination of old and new tools. Many technological breakthroughs have made discovery of new milk molecules possible. These advances include improved separation science, high-resolution detection instrumentation, bioinformatic toolsets (like proteomic software) and genomics (as a tool to guide proteomics). Currently, toolsets exist for the detection of a wide array of milk lipids (lipidomics), some milk glycolipids, proteins (proteomics), naturally occurring peptides in milk (peptidomics) and free milk oligosaccharides (glycomics).

Now that these technologies exist, assessment of which dairy processing streams yield the highest concentrations of functional compounds is possible. Many processing streams at various steps of production need to be evaluated in detail for all carbohydrate structures: this labor-intensive process has driven the development of new methods and software to aid in automating the analyses. Different processing conditions can be analyzed to determine which produce highest bioactive compound yield and best preserve bioactivity.

After identification, there is a need for scaled-up production of each functional molecule/molecular class in order to test them for function in assays and clinical trials. Molecular classes that could be isolated from dairy streams include peptides, glycoproteins, glycolipids and free oligosaccharides. With successful separation of milk components and verification of their identicalness or similarity to human milk components, these compounds can serve as the basis for improved infant formulas or other therapeutic applications. Application of molecules shown to be functional to commercial products will increase the overall value of the dairy industry.

This review explores how the discoveries of novel components of milk have been made possible through advances in separation science and analytical chemistry, and how these molecules can be isolated at the industrial scale.

Chromatography: an enabling step to reduce milk complexity

In order to clearly identify each molecular type by mass spectrometry, compounds must be purified from all contaminants and chromatographically separated. Without separation, molecules will compete for ionization. Molecules with low ionization efficiency will not be detected as their signals will be suppressed by other molecules that are easily ionized. Liquid chromatography separates molecules over time based on their affinity to a stationary phase. By separating molecules over time with chromatography, fewer molecules are introduced to the mass spectrometer at any one time, decreasing ion suppression and increasing the number of possible compounds detected. Liquid chromatography can be paired in-line with mass spectrometry via electrospray ionization. Nanoflow chromatography systems employ less sample and solvent than normal liquid chromatography and can be packed with a variety of solid phases allowing separation of peptides, glycans and other classes of molecules. Graphitized carbon columns can separate glycans down to the structural isomers level and thus enable structure-specific glycomics [11,12]. Peptides are typically separated with reverse-phase chromatography (non-polar stationary phase). Lipid chromatography can be performed with reverse phase and normal phases (eg. silica).

Modern liquid chromatography allows reproducible retention times that can aid in the identification of known molecules in the chromatogram. The mapping of compounds can be improved with higher reproducibility of column separation and sharper peak elutions (i.e., creation of new columns with higher theoretical plates). Retention times and mass can be used to identify specific known structures.

Compounds can also be fractionated on a preparative basis for further experimentation via flash liquid chromatography. For example, flash liquid chromatography with a porous graphitized carbon stationary phase was used to fractionate milligram quantities of glycan isomers from milk to generate standards that are not available commercially [13].

Fractionation of milligram quantities of sample allows testing of these different fractions in assays to discover specific functionality of each.

Mass spectrometry: enabling molecular identifications in milk

Mass spectrometric platforms are capable of detecting more compounds in a single run than other traditional analytical tools. The high mass accuracy available in modern mass analyzers dramatically reduces the number of possible analyte identities. With small molecules, this high precision measurement can lead to a single compositional possibility. A variety of mass spectrometers can be employed for detection of bioactive milk components (e.g. TOF, Fourier Transform Ion Cyclotron Resonance (FTICR)). These mass analyzers have high-resolution and are highly sensitive. Modern mass spectrometers allow detection of molecules at sub-femtomole concentrations. Figure 1 depicts the generic schema for mass spectrometers.

Advances in ionization technology were crucial to the use of mass spectrometry for milk component detection. Soft ionization techniques (which produce mainly molecular ions) like matrix-assisted laser desorption/ionization (MALDI) [14,15] and electrospray ionization (ESI) [16] are requisite for intact biomolecule analysis for nonvolatile, thermally unstable compounds such as peptides, glycans and glycolipids.

The mass of the intact molecule alone is often not enough to achieve unambiguous molecular assignment. To narrow down the list of potential molecular structures, unknown ions are sequentially isolated and fragmented. The fragments in these spectra allow for confirmation of one compositional identity. In some cases, tandem fragmentation can provide structural information.

Construction and application of molecular libraries allows streamlined, high-throughput identification of known molecules. Application of a molecular library within spectral analysis software (such as Agilent MassHunter) allows fast matching of ions to the library and computational extraction of each compound's abundance. These libraries typically include the mass, retention time and identity of the compound. More sophisticated libraries such as Agilent Personal Compound Database Library also employ representative MS/MS spectra for each compound to increase the certainty of the match. For these libraries, the MS/MS spectra of the unidentified compound must match, within a margin of error, the peaks in the reference spectra. These libraries allow for high-speed evaluation of samples from milk and dairy streams for their bioactive compound content.

Potential uses of the technology to identify milk components

Peptidomics—Most milk proteins are non-functional when intact. Likely, their importance lies in the encrypted functional fragments that are released during digestion. A large number of naturally occurring peptides is known to exist in milk [17]. Milk peptides created by *in vitro* digestion have immunomodulatory, antimicrobial, opioid and nutritional functions [18].

Peptidomic analysis of milk's naturally occurring peptides is only now possible through a confluence of novel preparative isolation, analytical separation, analytical detection and computational techniques. Milk peptides can be isolated by first removing cream by centrifugation, then removing the large proteins via acid precipitation and finally removing the salts, lactose and oligosaccharides by preparative non-polar (C18) solid-phase extraction (SPE). Once isolated, peptides can be separated with a C18 analytical column and detected by a mass analyzer. Even low ppm precision intact mass measurement is not enough to identify the exact peptide sequence. Therefore, compounds can be sequentially isolated and fragmented. The fragment spectra can be analyzed for components of the peptide to provide evidence for which possible sequence is correct. As the mass spectrometer typically produces thousands of tandem spectra in an hour run, data must be processed via automated software. Data can be exported and the fragment spectra analyzed with proteomic software such as X!Tandem. Because milk's naturally occurring peptides are not cleaved *in vitro* by a single enzyme such as trypsin in most proteomic protocols, a search with a nonspecific enzymatic cleavage pattern must be performed. To narrow the possible peptide sequences, a milk protein-specific library can be employed as the basis for the search rather than the usual

complete human proteome. The peptides identified from the automated analysis can be compiled as a library of known peptides with mass, retention time and a unique identity. This library can then be applied back to the mass spectrometer runs of the sample to extract peak volumes. Peak volumes can be compared between identical peptides across different sample sets.

Before the invention of soft ionization methods for mass spectrometry, including ESI and MALDI, peptides could not be analyzed by mass spectrometry as they are not volatile and are thermally labile. Without liquid chromatography (which is paired with ESI), the number of identifiable peptides would be greatly reduced. Without the low ppm mass accuracy measurements of modern mass spectrometers, peptides could not be identified. Without tandem fragmentation, peptides could not be sequenced with mass spectrometry. Advances in programming designed for proteomics were also essential to bringing about milk peptidomics. Simple modifications to the search parameters allowed application of these search algorithms to a new problem, which would have been impossible without this computational advance.

Glycolipidomics—Milk glycolipids exert important health effects, including prebiotic action, intestinal immune response modulation and acting as decoys for toxins [19]. Sialylated glycolipids-gangliosides-influence the microbiota that colonize the gut by supporting the growth of beneficial bacterial species, especially bifidobacteria. Infants fed a ganglioside-supplemented formula show lower count of intestinal pathogenic *Escherichia coli* than infants fed a standard formula [20]. Significant observations from animal studies show that inflamed intestinal mucosa has less sialylated gangliosides than healthy intestinal mucosa [21]. Accelerated catabolism of gangliosides in the intestine increases pro-inflammatory signaling, inflammatory markers and susceptibility to pathogens [22]. Dietary gangliosides can replace the mucosal gangliosides that are progressively degraded in inflammatory conditions [23]. Human and bovine milk gangliosides are similar to human intestinal membrane glycolipids as both contain GM3 (NeuAc α 2-3Gal β 1-4Glc β -Cer) and GD3 (NeuAc α 2-8NeuAc α 2-3Gal β 1-4Glc β -Cer). Therefore, milk consumption might help restore proper ganglioside abundance and function in the intestine. This change could resolve inflammation, increase resistance to infection, and improve gut integrity to induce remission of gut inflammatory diseases like necrotizing enterocolitis, inflammatory bowel diseases and Crohn's disease.

The isolation and analysis of glycolipids are two of the most challenging analytical problems of lipidomics because these unique molecules have characteristics of both lipids (hydrophobic) and glycans (hydrophilic) and are not water-soluble. These physical properties make their separation exceptionally challenging. Because of their amphipathic nature and structural complexity, the classical methods for glycolipid analysis require extensive sample preparation, and provide only partial information on the oligosaccharide and ceramide portions. Recently, a method was described for the rapid and accurate profiling of human or bovine milk gangliosides with both glycan and lipid components identified using high resolution MALDI-FTICR mass spectrometry [24]. GM3 and GD3 are the most abundant gangliosides in the milks, and their ceramide structures were identified simultaneously (Figure 2). Measured accurate mass values combined with the Kendrick

mass defect plot revealed each compound's oligosaccharide headgroup and the carbon distribution in ceramide backbone simply and rapidly.

Each ganglioside species can be quantified by liquid chromatography tandem mass spectrometry with the use of a triple quadrupole mass spectrometer in multiple reactions monitoring (MRM) mode. MRM provides a highly selective and very sensitive determination of gangliosides in a mixture, and it is quantitative when used with pertinent standards [25,26]. This quantitative method can be used to measure changes in gangliosides under various physiological and processing conditions.

Glycomics—Another class of milk components that has challenged researchers for years is free milk oligosaccharides. Recent research suggests that all mammalian milks contain a highly potent mixture of prebiotic glycans (oligosaccharides) with novel bioactivities that appear to protect and promote human health. The two recognized groups of milk oligosaccharides are common to most mammalian milk, including bovine milk, and are composed of neutral oligosaccharides that contain the monomers galactose, N-acetylglucosamine, fucose and a lactose core, and acidic oligosaccharides that contain the same monomers as well as N-acetylneuraminic acid (NeuAc), a type of sialic acid [27,28].

Because of their structural complexity, only bacteria that possess the necessary enzymes to cleave the specific glycosidic bonds that link the individual monosaccharides (fucose, sialic acid, N-acetylhexosamine (HexNAc), hexoses) together are able to utilize such complex oligosaccharides for growth. This genetic capability is thought to be part of the basis by which these beneficial bacteria out-compete pathogenic bacteria in the gut. In particular, the monosaccharide fucose and sialic acid appear to be crucial to the ability of oligosaccharides to selectively enrich beneficial bacteria. Sialylated HMO carrying one or more molecules of NeuAc are also thought to contribute to infant brain development by acting as a source of NeuAc for brain ganglioside assembly [29]. This hypothesis is supported by research showing higher I.Q. for breast-fed infants compared with formula-fed infants [30-32]. Bovine milk oligosaccharides (BMO) are similar in structure to HMO, which enhance the growth of commensal microbiota [33,34]. Bovine colostrum contains high levels of acidic sialylated oligosaccharides [33]. Bovine milk-based formulae contain these oligosaccharides only in trace amounts, whereas human milk contains 500-1000 mg/L of sialyl-oligosaccharides [35].

Many important applications await the isolation and characterization of the non-digestible oligosaccharides in milk. Unfortunately, at present there are no proven commercial sources for HMO-like compounds with high amounts of sialylation and fucosylation, thus the dairy industry is supplementing products with simpler oligosaccharides extracted from plants.

The structural complexity of oligosaccharides makes a single tool and a single method insufficient for their characterization. The major analytical techniques currently include mass spectrometry, chromatography such as high pH anion-exchange chromatography with pulsed amperometric detection and nuclear magnetic resonance spectroscopy. Several groups worldwide are working to complete the annotation of the human and bovine milk glycome [33,34,36-38]. Nano-LC-Q-TOF using graphitized carbon columns is emerging as

the preferred method for oligosaccharide characterization because of the high sensitivity and reproducibility, and the minute amount of sample required, and because of the high-throughput that allowed construction of the first bioinformatic library of HMO containing over 75 fully annotated structures [38,39].

Commercial viability of large-scale extraction of bioactive molecules from agricultural commodities will depend on the richness of these molecules in the sources. Bovine milk contains only trace amounts of these valuable compounds, but many dairy side-streams are potentially rich in bioactive oligosaccharides. The recent discovery that whey permeate contains complex fucosylated and sialylated oligosaccharides analogous to the composition and structures of HMO demonstrates this stream as a promising source for industrial production of human-milk like oligosaccharides [40,41].

This discovery has led to collaborations with the US Dairy Industry to advance the research and development of these specific oligosaccharides as food ingredients. It appears that concentrating certain oligosaccharides from whey permeate can be a cost-effective process for the valorization of whey permeate into high quality and profitable novel dairy ingredients. Purified BMO extracted from cheese whey is currently being used in a human trial (ClinicalTrials.gov Protocol Record 264294).

The use of whey permeates as a source of BMO is especially attractive due to the wide availability and low cost of this by-product compared with those of liquid bovine milk. There is an urgent need, however, to identify which dairy streams (whey permeate, mother liquor, etc.) have the highest concentration of bioactive glycans. The recent creation of a specific bioinformatic library of BMO that includes several novel fucosylated oligosaccharides will enable rapid identification of low abundance oligosaccharides in dairy streams [37]. Figure 3 presents a partial structure library detailing the most important fucosylated structures identified to date in bovine milk. Application of this library will facilitate structural identification in all dairy streams.

Glycoproteomics—Milk proteins are extensively post-translationally modified. Post-translational modifications (PTM) of milk proteins are the result of enzymatic processing of a polypeptide chain in the cell. Hundreds of PTM exist, including include glycosylation, phosphorylation, oxidation and deamidation. PTM influence compartmentalization, function, stability, structure, isoelectric point and proteolytic susceptibility of proteins [42,43]. Many milk proteins are decorated with long carbohydrate chains (glycosylation) during biosynthesis [44,45]. Glycans are linked to the protein through O-glycosidic or N-glycosidic bonds [46]. N-glycans are linked to the asparagine residues of proteins in the specific amino acid sequence Asn-XXX-Ser/Thr. The N-glycan core is composed of 2 N-acetylglucosamines (GlcNAc) and 3 mannoses. The different classes of N-glycan include high mannose (addition of mannose to the core), complex (addition of HexNAc to the core) and hybrid (combination of high mannose and complex). Additional glycan-like sialic acid and fucose can be added to the glycan structure. Protein-linked glycans are involved in multiple cellular mechanisms that contribute to the health and development of the infant [47]. The diverse protein-linked glycans may act as decoys for pathogens by competitively inhibiting their attachment to intestinal cell membrane glycans, thus reducing risk of

infection [48]. The function of glycans varies based on the protein to which they are attached and the precise site of attachment [49]. Determining function of glycosylation on individual milk proteins remains a significant challenge. Whereas the primary structures of linear biomolecules such as nucleic acids and proteins may be elucidated simply by determining their sequence, glycans are branched, with both linkage and anomeric isomers, which significantly complicates their identification.

Complete glycoprotein analysis requires determination of the sites of glycosylation and the glycan structures associated with each site. As complete glycoprotein analysis remains challenging, glycoproteins are typically analyzed either via proteocentric methods where the glycan is removed enzymatically and the protein identified or the glyco-centric method where the protein is removed and the glycans are identified [50]. Protein linked N-glycans can be released from the protein by a commercial enzyme, peptidyl-N-glycosidase F (PNGase F), which cleaves between the asparagine residue and the first GlcNAc. Figure 4 displays the specific of the mechanism of deglycosylation and the workflow for the characterization of released glycans

Released glycans can be purified by graphitized carbon chromatography. N-glycans are challenging to characterize because of their complexity and diversity. Highly sensitive analytical methods are required in order to separate and differentiate among the variety of glycan compositions and structures. Tandem mass spectrometry can help determine glycan structure. In some instances, a series of enzymatic cleavage steps with specific glycosidases is necessary for determining structure [51,52]. Once N-glycan structures are known, they can be compiled into libraries of known mass and retention time. These libraries can then be applied to new samples for fast identification of known molecules.

Glycoprotein isolation

Milk glycoproteins are abundant both in the whey and in the casein micelles. Milk glycoproteins have a variety of health-promoting functions including antimicrobial, antiviral, antioxidant, antiinflammatory, immunomodulatory and anti-cancer actions [53-56]. These functional proteins can be applied as ingredients in functional foods, formulas and other therapeutics. Protein-linked glycans may help protect against microbial and viral attacks and help develop balanced intestinal microbiota [48,57-59]. Glycans impact structural and functional properties of proteins as well as the resistance to proteolysis and thermal stability [60,61].

Modern separation technology also allows the isolation of specific proteins and glycoproteins from milk and whey. For example, lactoferrin, which is one of the most abundant glycoproteins in many mammalian milks, can now be isolated from bovine whey in high purity via readily available chromatographic methods including ion-exchange chromatography [62] and affinity chromatography [48] (Figure 5). In these affinity resins, molecules (ligands) are immobilized on beads and packed in column. Novel chromatographic methods allow separation of glycoproteins based on their glycan composition [63]. For example, immobilized lectins selectively bind specific protein-linked glycans. The lectin concanavalin A, for instance, binds α -linked mannose and terminal glucose. These techniques allow high purity isolation of the protein of interest from a

complex mixture. Optimization is required for each unique glycoprotein of interest. The quality of the purification can be monitored by electrophoretic methods and mass spectrometry. Isolated functional milk proteins can be isolated at industrial scale from milk and dairy streams to provide a functional ingredient for formula production, functional foods or therapeutic applications.

Scale-up

The same techniques used for separation/purification of specific milk molecules at a laboratory scale can be scaled-up to industrial scale. Scaled-up isolation of milk molecules enables their use in functional testing and clinical trials. Some of the processes used to fractionate and purify dairy bioactive components (glycoproteins, peptides, glycolipids, and free glycans) at laboratory-scale include cream separation, ion-exchange and affinity chromatography, gel and capillary electrophoresis, selective precipitation, solid-phase extraction and membrane filtration. Except for cream separation and membrane applications, the commercial adoption of these techniques has been hindered by unacceptable economics at large-scale, operational complexity, low productivity, large effluent production, and product degradation associated with the use of high temperatures, pH, and salts [64,65].

Membrane separation is widely used in the dairy industry for isolation/fractionation of whey proteins [66]. Membranes with a wide array of pore sizes and different configurations are commercially available. Successive filtration steps with different pore size membranes allow the separation of numerous milk components according to their size. The increasing number of successful membrane applications in the dairy industry is a direct result of many advances in membrane technology such as increased membrane reproducibility and chemical resistance, development of modules that can be easily sanitized to comply with food grade requirements, development of commercial inorganic membranes (which allow use with wider temperature and pH ranges, have higher chemical resistance, extend operating lifetime, and have back-flushing capability), development of functionalized ion-exchange membranes (anion- and cation-exchange membranes, a possible alternative to using column chromatography) and improvement of membrane hydrodynamics system hardware [66,67]. Membrane-based separations can be effectively and economically implemented at large-scale, a requirement for most dairy applications [64]. Some of the membrane applications in dairy processing include the use of micro filtration membranes to remove milk fat and bacteria, ultrafiltration membranes to recover and/or fractionate whey proteins and casein micelles from milk, and the combination of ultra-and nanofiltration to recover oligosaccharides [64,66,68].

Currently, chromatography scale-up techniques are expensive because of the cost of their solid phase. However, innovations are helping to make these techniques increasingly economically feasible. For example, research to improve the feasibility of large-scale chromatography through regeneration and reuse of both solid-phase materials and solvents are ongoing [69]. Recent developments in industrial-scale bio-chromatography columns such as the development of new column designs that promote optimal performance with any resin at any scale and that are easy and safe to handle suggest that larger scale isolation of glycoproteins is feasible.

Separation of proteins with similar size but different isoelectric points can be achieved by the use of charged membranes at specific pH and ionic strength [70]. This combination has the potential to overcome a major impediment for protein fractionation—the permeability-selectivity trade-off—and expand the use of membrane filtration in downstream processes where chromatography was previously necessary. One recent development in membrane-based technologies—the coupling of imposed electric fields and pressure-driven filtration [71,72] has potential to improve the separation and purification of milk bioactives. The application of an electric field across a membrane during filtration can enhance the fractionation of negatively or positively charged molecules. Like-charged large molecules are repelled from the membrane surface, which allows for increased transmission of smaller molecules through the membrane [71]. In comparison to normal membrane filtration, electric field-enhanced membrane filtration applied to proteins and peptides has been shown to reduce membrane fouling and, therefore, improve permeate flux and fractionation selectivity [71,72].

The global production of liquid whey amounted to 186 million metric tons in 2008 [73] and it will continue to grow, along with cheese production. For environmental reasons, whey is no longer discharged into rivers because of its high biochemical and chemical oxygen demand (index of pollution). This stream, once considered waste by the dairy industry, is now known to contain a variety of functional components including glycoproteins, oligosaccharides and naturally occurring peptides. Whey proteins are now isolated by membrane filtration, but leave behind whey permeate that still contains the oligosaccharides and naturally occurring peptides. This whey permeate is currently disposed of by the manufacturer. The functional molecules in these dairy streams should be captured for further use. Current technologies allow for the capture of these components. A general schematic for industrial-scale isolation of functional milk components such as glycoproteins, glycolipids, peptides and oligosaccharides from bovine milk is presented in Figure 6.

Glycolipid isolation

Lipids can be isolated by either centrifugation/cream separator or membrane microfiltration (pore size range 0.10-0.5 μm) [66]. Solvent extraction, which is commonly used by the food industry, can be used to isolate glycolipids from neutral lipids (e.g. triglycerides). Then, reverse-phase chromatography can be employed at large-scale to remove other components like oligosaccharides, peptides and lactose.

Whey Protein Isolation

The dairy industry employs ultrafiltration technology extensively for recovery of soluble whey protein concentrates from whey [66]. The development of robust membrane systems, different membrane configurations and continuous operation using multi-stages associated with the use of diafiltration (addition of solvent to wash micro-solutes from the retentate, therefore increasing the purity of the macromolecules retained) have contributed to the success of this technology in the dairy industry [66]. Ultrafiltration of whey proteins generates two streams, a protein-rich retentate and a carbohydrate-rich permeate (oligosaccharides, lactose).

Glycoprotein Isolation

Glycoproteins from whey protein concentrate can be isolated from the whey or casein protein fractions at industrial scale by chromatography. For example, lactoferrin can be isolated at industrial scale by ion-exchange chromatography [74]. Bovine lactoferrin is produced from whey cheese or raw milk by dairy industries for use in bovine milk-based infant formula for health claims.

Oligosaccharide isolation

A major challenge for recovery of oligosaccharides from dairy streams is to simultaneously enrich the oligosaccharide concentration and deplete the content of lactose and other simple sugars that lack bioactive functionality. Lactose removal is important for bioactivity studies because it is the most abundant carbohydrate in whey and can confound studies on the biological activity of oligosaccharides. An inexpensive and scalable approach to recover biologically active oligosaccharides from milk was developed. This approach relies on the combination of enzymatic hydrolysis of lactose, a two-step filtration (ultra- and nano-filtration), followed by diafiltration of the nano-filtration retentate in order to remove the monosaccharides, thus enriching the retentate oligosaccharide content. With this approach, more than 50% of human milk's free oligosaccharides were recovered.

Further processing, enrichment and purification methods (Figure 6) are being investigated at pilot scale to optimize recovery of oligosaccharides from agricultural streams in a way that maximizes the concentration of specific structures present in human milk and depletes structures not present in human milk.

Peptide isolation

Naturally occurring milk peptides will remain in the whey permeate with the oligosaccharides and lactose after ultrafiltration (10-30 kDa). Separation of these peptides from this solution has not been performed yet. This separation would require the use of large-scale chromatography, possibly with reverse- or normal-phases.

The development of feasibly priced, fully scalable fractionation and purification technologies is essential to enable large-scale production of purified bioactive milk compounds for their utilization in clinical trials. These isolated components can be employed as food products and supplements and may have significant health benefits for human nutrition.

Conclusion

Advances in separation and analytical technologies are allowing discovery of hundreds of novel, functional milk molecules. The dairy industry has benefitted and will continue to benefit from the discovery of novel functional molecules in milk.

High-resolution mass spectrometry, molecular libraries and minute sample concentration requirements enable high-throughput investigation of different milk sources, including different dairy processing streams, breeds and species, for highest concentration of these

compounds. The effects of genetic polymorphisms, environmental conditions, feeding strategies and lactation stage on the abundances of functional components will need to be investigated. Some species of lactating farm animals (cows, goats, sheep, etc.) and some breeds within those species are likely to have different levels of certain components more similar to those in human milk. Characterization of such a complexity is crucial for identification of viable sources for purification. For example, a recent study suggested that Danish Jersey milk is richer in bioactive (fucosylated) oligosaccharides than Danish Holstein-Friesian milk, and thus may have greater health-promoting benefits for consumers and may be a richer source for fucosylated oligosaccharide isolation [40]. Each species evolved with different physiological needs, which means each species' milk will have a slightly different repertoire of functional compounds for protecting and nourishing their offspring. Therefore, novel components in dairy animal milk that are lacking in human milk may nonetheless be employed in human therapeutics.

Assessing all of these elements was impossible a decade ago. However, the confluence of modern “omics” sciences and high-throughput mass spectrometry now makes this task is imminently achievable.

The scale-up of laboratory-based processing and purification techniques to industrial level at feasible cost is essential to achieve commercial production of these dairy-derived functional components as health-enhancing ingredients. By identifying functional components in dairy waste streams, this science will allow the dairy industry to improve their profitability. Partnerships between academia and industry will catalyze breakthroughs in isolation and fractionation of these functional ingredients from dairy streams. This will help bridge the gap from laboratory-based research to applications and accelerate the translation of these components to the marketplace.

Acknowledgments

The authors acknowledge CJ Dillard for editorial assistance. This research was partially supported by the Dairy Research Institute and the USDA National Institute of Food and Agriculture post-doctoral fellowship.

References

1. Oftedal OT. The mammary gland and its origin during synapsid evolution. *J Mammary Gland Biol and Neoplasia*. 2002; 7:225–252. [PubMed: 12751889]
2. German JB, Dillard CJ, Ward RE. Bioactive components in milk. *Curr Opin Clin Nutr Metab Care*. 2002; 5:653–658. [PubMed: 12394640]
3. Newburg DS, Ruiz-Palacios GM, Altaye M, Chaturvedi P, Meinzen-Derr J, et al. Innate protection conferred by fucosylated oligosaccharides of human milk against diarrhea in breastfed infants. *Glycobiology*. 2004; 14:253–263. [PubMed: 14638628]
4. Gifford JL, Hunter HN, Vogel HJ. Lactoferricin: a lactoferrin-derived peptide with antimicrobial, antiviral, antitumor and immunological properties. *Cell Mol Life Sci*. 2005; 62:2588–2598. [PubMed: 16261252]
5. LoCascio RG, Ninonuevo MR, Freeman SL, Sela DA, Grimm R, et al. Glycoprofiling of bifidobacterial consumption of human milk oligosaccharides demonstrates strain specific, preferential consumption of small chain glycans secreted in early human lactation. *J Agric Food Chem*. 2007; 55:8914–8919. [PubMed: 17915960]
6. Politis I, Chronopoulou R. Milk peptides and immune response in the neonate. *Adv Exp Med Biol*. 2008; 606:253–269. [PubMed: 18183933]

7. Hanson LA, Ahlstedt S, Andersson B, Carlsson B, Fällström SP, et al. Protective factors in milk and the development of the immune system. *Pediatrics*. 1985; 75:172–176. [PubMed: 3880886]
8. Chichlowski M, De Lartigue G, German JB, Raybould HE, Mills DA. Bifidobacteria isolated from infants and cultured on human milk oligosaccharides affect intestinal epithelial function. *J Pediatr Gastroenterol Nutr*. 2012; 55:321–327. [PubMed: 22383026]
9. Hong P, Ninonuevo MR, Lee B, Lebrilla C, Bode L. Human milk oligosaccharides reduce HIV-1-gp120 binding to dendritic cell-specific ICAM3-grabbing non-integrin (DC-SIGN). *Br J Nutr*. 2009; 101:482–486. [PubMed: 19230080]
10. Svensson M, Håkansson A, Mossberg AK, Linse S, Svanborg C. Conversion of alpha-lactalbumin to a protein inducing apoptosis. *Proc Natl Acad Sci USA*. 2000; 97:4221–4226. [PubMed: 10760289]
11. Chu CS, Niñonuevo MR, Clowers BH, Perkins PD, An HJ, et al. Profile of native N-linked glycan structures from human serum using high performance liquid chromatography on a microfluidic chip and time-of-flight mass spectrometry. *Proteomics*. 2009; 9:1939–1951. [PubMed: 19288519]
12. Niñonuevo M, An H, Yin H, Killeen K, Grimm R, et al. Nanoliquid chromatography-mass spectrometry of oligosaccharides employing graphitized carbon chromatography on microchip with a high-accuracy mass analyzer. *Electrophoresis*. 2005; 26:3641–3649. [PubMed: 16196105]
13. Strum JS, Aldredge D, Barile D, Lebrilla CB. Coupling flash liquid chromatography with mass spectrometry for enrichment and isolation of milk oligosaccharides for functional studies. *Anal Biochem*. 2012; 424:87–96. [PubMed: 22370281]
14. Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, et al. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*. 1988; 2:151–153.
15. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem*. 1988; 60:2299–2301. [PubMed: 3239801]
16. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM, et al. Electrospray ionization—principles and practice. *Mass Spectrometry Reviews*. 1990; 9:37–70.
17. Ferranti P, Traisci MV, Picariello G, Nasi A, Boschi V, et al. Casein proteolysis in human milk: tracing the pattern of casein breakdown and the formation of potential bioactive peptides. *J Dairy Res*. 2004; 71:74–87. [PubMed: 15068070]
18. Schlimme E, Meisel H. Bioactive peptides derived from milk proteins. Structural, physiological and analytical aspects. *Nahrung*. 1995; 39:1–20. [PubMed: 7898574]
19. Rueda R. The role of dietary gangliosides on immunity and the prevention of infection. *Br J Nutr*. 2007; 98(1):S68–S73. [PubMed: 17922964]
20. Rueda R, Sabatel JL, Maldonado J, Molina-Font JA, Gil A. Addition of gangliosides to an adapted milk formula modifies levels of fecal *Escherichia coli* in preterm newborn infants. *J Pediatr*. 1998; 133:90–94. [PubMed: 9672517]
21. Park EJ, Suh M, Thomson B, Ma DW, Ramanujam K, et al. Dietary ganglioside inhibits acute inflammatory signals in intestinal mucosa and blood induced by systemic inflammation of *Escherichia coli* lipopolysaccharide. *Shock*. 2007; 28:112–117. [PubMed: 17510604]
22. Miklavcic JJ, Schnabl KL, Mazurak VC, Thomson AB, Clandinin MT. Dietary ganglioside reduces proinflammatory signaling in the intestine. *J Nutr Metab*. 2012; 2012:280286. [PubMed: 22506104]
23. Park EJ, Suh M, Clandinin MT. Dietary ganglioside and long-chain polyunsaturated fatty acids increase ganglioside GD3 content and alter the phospholipid profile in neonatal rat retina. *Invest Ophthalmol Vis Sci*. 2005; 46:2571–2575. [PubMed: 15980250]
24. Lee H, An HJ, Lerno LA Jr, German JB, Lebrilla CB. Rapid profiling of bovine and human milk gangliosides by matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry. *Int J Mass Spectrom*. 2011; 305:138–150. [PubMed: 21860602]
25. Sørensen LK. A liquid chromatography/tandem mass spectrometric approach for the determination of gangliosides GD3 and GM3 in bovine milk and infant formulae. *Rapid Commun Mass Spectrom*. 2006; 20:3625–3633. [PubMed: 17094170]
26. Zhang J, Ren Y, Huang B, Tao B, Pedersen MR, et al. Determination of disialoganglioside GD3 and monosialoganglioside GM3 in infant formulas and whey protein concentrates by ultra-

- performance liquid chromatography/electrospray ionization tandem mass spectrometry. *Journal of separation science*. 2012; 35:937–946. [PubMed: 22589154]
27. Gopal PK, Gill HS. Oligosaccharides and glycoconjugates in bovine milk and colostrum. *Br J Nutr*. 2000; 84(1):S69–S74. [PubMed: 11242449]
 28. Hakkarainen J, Toivanen M, Leinonen A, Frängsmyr L, Strömberg N, et al. Human and bovine milk oligosaccharides inhibit *Neisseria meningitidis* pili attachment in vitro. *J Nutr*. 2005; 135:2445–2448. [PubMed: 16177210]
 29. Kleene R, Schachner M. Glycans and neural cell interactions. *Nat Rev Neurosci*. 2004; 5:195–208. [PubMed: 14976519]
 30. Wang B, Brand-Miller J. The role and potential of sialic acid in human nutrition. *Eur J Clin Nutr*. 2003; 57:1351–1369. [PubMed: 14576748]
 31. Rodgers B. Feeding in infancy and later ability and attainment: a longitudinal study. *Dev Med Child Neurol*. 1978; 20:421–426. [PubMed: 689307]
 32. Lucas A, Morley R, Cole TJ, Lister G, Leeson-Payne C. Breast milk and subsequent intelligence quotient in children born preterm. *Lancet*. 1992; 339:261–264. [PubMed: 1346280]
 33. Tao N, DePeters EJ, Freeman S, German JB, Grimm R, et al. Bovine milk glycome. *J Dairy Sci*. 2008; 91:3768–3778. [PubMed: 18832198]
 34. Barile D, Marotta M, Chu C, Mehra R, Grimm R, et al. Neutral and acidic oligosaccharides in Holstein-Friesian colostrum during the first 3 days of lactation measured by high performance liquid chromatography on a microfluidic chip and time-of-flight mass spectrometry. *J Dairy Sci*. 2010; 93:3940–3949. [PubMed: 20723667]
 35. Martín-Sosa S, Martín MJ, García-Pardo LA, Hueso P. Sialyloligosaccharides in human and bovine milk and in infant formulas: variations with the progression of lactation. *J Dairy Sci*. 2003; 86:52–59. [PubMed: 12613848]
 36. Ninonuevo MR, Park Y, Yin H, Zhang J, Ward RE, et al. A strategy for annotating the human milk glycome. *J Agric Food Chem*. 2006; 54:7471–7480. [PubMed: 17002410]
 37. Aldredge DL, Geronimo MR, Hua S, Nwosu CC, Lebrilla CB, et al. Annotation and structural elucidation of bovine milk oligosaccharides and determination of novel fucosylated structures. *Glycobiology*. 2013; 23:664–676. [PubMed: 23436288]
 38. Wu S, Tao N, German JB, Grimm R, Lebrilla CB. Development of an annotated library of neutral human milk oligosaccharides. *J Proteome Res*. 2010; 9:4138–4151. [PubMed: 20578730]
 39. Wu S, Grimm R, German JB, Lebrilla CB. Annotation and structural analysis of sialylated human milk oligosaccharides. *J Proteome Res*. 2011; 10:856–868. [PubMed: 21133381]
 40. Sundekilde UK, Barile D, Meyrand M, Poulsen NA, Larsen LB, et al. Natural variability in bovine milk oligosaccharides from Danish Jersey and Holstein-Friesian breeds. *J Agric Food Chem*. 2012; 60:6188–6196. [PubMed: 22632419]
 41. Barile D, Tao N, Lebrilla CB, Coisson JD, Arlorio M, et al. Permeate from cheese whey ultrafiltration is a source of milk oligosaccharides. *Int Dairy J*. 2009; 19:524–530. [PubMed: 20161544]
 42. Solá RJ, Rodríguez-Martínez JA, Griebenow K. Modulation of protein biophysical properties by chemical glycosylation: biochemical insights and biomedical implications. *Cell Mol Life Sci*. 2007; 64:2133–2152. [PubMed: 17558468]
 43. Shental-Bechor D, Levy Y. Folding of glycoproteins: toward understanding the biophysics of the glycosylation code. *Curr Opin Struct Biol*. 2009; 19:524–533. [PubMed: 19647993]
 44. Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*. 2002; 12:43R–56R.
 45. Kornfeld R, Kornfeld S. Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem*. 1985; 54:631–664. [PubMed: 3896128]
 46. Varki, A.; Cummings, RD.; Esko, JD.; Freeze, HH.; Stanley, P., et al. *Essentials of glycobiology*. Cold Spring Harbor Laboratory Pr; 2009.
 47. Froehlich JW, Dodds ED, Barboza M, McJimpsey EL, Seipert RR, et al. Glycoprotein expression in human milk during lactation. *J Agric Food Chem*. 2010; 58:6440–6448. [PubMed: 20415418]

48. Barboza M, Pinzon J, Wickramasinghe S, Froehlich JW, Moeller I, et al. Glycosylation of human milk lactoferrin exhibits dynamic changes during early lactation enhancing its role in pathogenic bacteria-host interactions. *Mol Cell Proteomics*. 2012; 11:M111. [PubMed: 22261723]
49. Petrescu AJ, Milac AL, Petrescu SM, Dwek RA, Wormald MR. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology*. 2004; 14:103–114. [PubMed: 14514716]
50. Dodds ED, Seipert RR, Clowers BH, German JB, Lebrilla CB. Analytical performance of immobilized pronase for glycopeptide footprinting and implications for surpassing reductionist glycoproteomics. *J Proteome Res*. 2009; 8:502–512. [PubMed: 19072223]
51. Aldredge D, An HJ, Tang N, Waddell K, Lebrilla CB. Annotation of a serum N-glycan library for rapid identification of structures. *J Proteome Res*. 2012; 11:1958–1968. [PubMed: 22320385]
52. Zhang J, Lindsay LL, Hedrick JL, Lebrilla CB. Strategy for profiling and structure elucidation of mucin-type oligosaccharides by mass spectrometry. *Anal Chem*. 2004; 76:5990–6001. [PubMed: 15481946]
53. Seganti L, Di Biase AM, Marchetti M, Pietrantonio A, Tinari A, et al. Antiviral activity of lactoferrin towards naked viruses. *Biomaterials*. 2004; 17:295–299. [PubMed: 15222481]
54. Orsi N. The antimicrobial activity of lactoferrin: current status and perspectives. *Biomaterials*. 2004; 17:189–196. [PubMed: 15222464]
55. Lönnnerdal B, Iyer S. Lactoferrin: molecular structure and biological function. *Annu Rev Nutr*. 1995; 15:93–110. [PubMed: 8527233]
56. Legrand D, Ellass E, Carpentier M, Mazurier J. Lactoferrin: a modulator of immune and inflammatory responses. *Cell Mol Life Sci*. 2005; 62:2549–2559. [PubMed: 16261255]
57. Marth JD, Grewal PK. Mammalian glycosylation in immunity. *Nat Rev Immunol*. 2008; 8:874–887. [PubMed: 18846099]
58. Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. *Cell*. 2006; 126:855–867. [PubMed: 16959566]
59. Baker HM, Baker EN. A structural perspective on lactoferrin function. *Biochem Cell Biol*. 2012; 90:320–328. [PubMed: 22292559]
60. van Veen HA, Geerts ME, van Berkel PH, Nuijens JH. The role of N-linked glycosylation in the protection of human and bovine lactoferrin against tryptic proteolysis. *Eur J Biochem*. 2004; 271:678–684. [PubMed: 14764083]
61. Baker EN, Baker HM. A structural framework for understanding the multifunctional character of lactoferrin. *Biochimie*. 2009; 91:3–10. [PubMed: 18541155]
62. Law BA, Reiter B. The isolation and bacteriostatic properties of lactoferrin from bovine milk whey. *J Dairy Res*. 1977; 44:595–599. [PubMed: 338651]
63. Kletter D, Singh S, Bern M, Haab BB. Global comparisons of lectin-glycan interactions using a database of analyzed glycan array data. *Mol Cell Proteomics*. 2013; 12:1026–1035. [PubMed: 23399549]
64. Zydney AL. Protein separations using membrane filtration: new opportunities for whey fractionation. *International Dairy Journal*. 1998; 8:243–250.
65. Martinez-Ferez A, Rudloff S, Guadix A, Henkelb CA, Pohlentz G, et al. Goats' milk as a natural source of lactose-derived oligosaccharides: Isolation by membrane technology. *International Dairy Journal*. 2006; 16:173–181.
66. Munir, C. Ultrafiltration and microfiltration handbook. Technomic Lancaster, CRC Press; 1998.
67. Pouliot Y. Membrane processes in dairy technology—from a simple idea to worldwide panacea. *International Dairy Journal*. 2008; 18:735–740.
68. Goulas AK, Grandison AS, Rastall RA. Fractionation of oligosaccharides by nanofiltration. *Journal of the Science of Food and Agriculture*. 2003; 83:675–680.
69. Scoma A, Pintucci C, Bertin L, Carozzi P, Fava F. Increasing the large scale feasibility of a solid phase extraction procedure for the recovery of natural antioxidants from olive mill wastewaters. *Chemical Engineering Journal*. 2012; 198-199:103–109.

70. Arunkumar A, Etzel MR. Fractionation of α -lactalbumin from β -lactoglobulin using positively charged tangential-flow ultrafiltration membranes. *Separation and Purification Technology*. 2013; 105:121–128.
71. Sarkar B, DasGupta S, De S. Electric field enhanced fractionation of protein mixture using ultrafiltration. *Journal of Membrane Science*. 2009; 341:11–20.
72. Holder A, Weik J, Hinrichs J. A study of fouling during long-term fractionation of functional peptides by means of cross-flow ultrafiltration and cross-flow electro membrane filtration. *Journal of Membrane Science*. 2013
73. Tomita M, Wakabayashi H, Shin K, Yamauchi K, Yaeshima T, et al. Twenty-five years of research on bovine lactoferrin applications. *Biochimie*. 2009; 91:52–57. [PubMed: 18585434]
74. Sarney DB, Hale C, Frankel G, Vulfson EN. A novel approach to the recovery of biologically active oligosaccharides from milk using a combination of enzymatic treatment and nanofiltration. *Biotechnol Bioeng*. 2000; 69:461–467. [PubMed: 10862685]

Abbreviations

BMO	Bovine Milk Oligosaccharides
HMO	Human Milk Oligosaccharides
MALDI	Matrix-assisted Laser Desorption/Ionization
ESI	Electrospray Ionization
Q-TOF	Quadrupole Time-of-flight Mass Spectrometer
Nano-LC-Q-TOF	Nanospray Liquid Chromatography Q-TOF
FTICR	Fourier Transform Ion Cyclotron Resonance
SPE	Solid-phase Extraction
MRM	Multiple Reaction Monitoring
NeuAc	N-acetylneuraminic acid

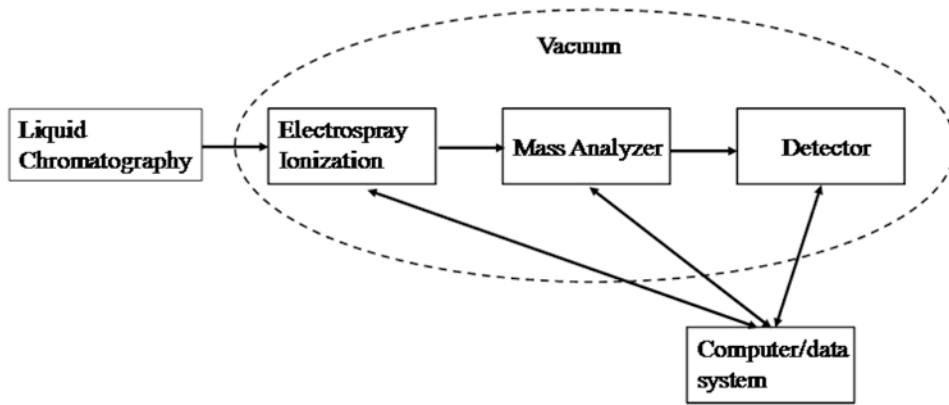


Figure 1.
Basic components of mass spectrometers.

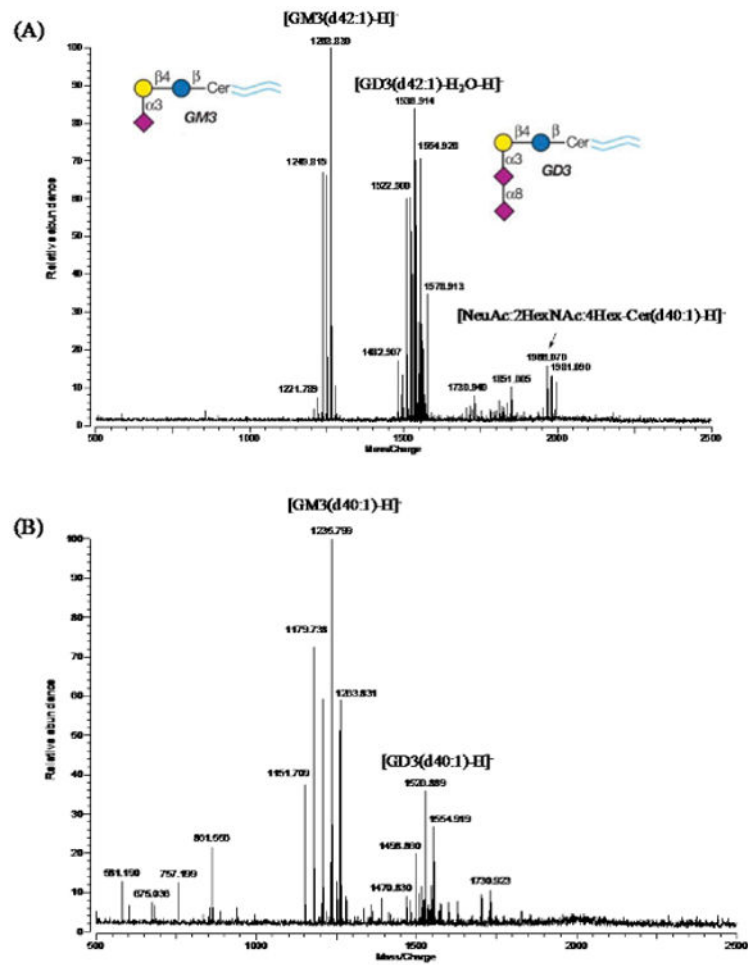


Figure 2.

Representative MALDI-MS spectra of gangliosides from bovine and human milk. Negative mode MALDI MS spectrum of (A) bovine milk gangliosides and (B) human milk gangliosides. The inserts show the structures of GM3 and GD3 gangliosides.



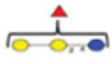



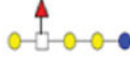

RT	Accurate Mass	Hex	HexNAc	Fuc	NeuAc	NeuGc	Structure
1.78	490.19	2	0	1	0	0	
12.628	490.1897548	2	0	1	0	0	
13.283	652.2425748	3	0	1	0	0	
19.873	652.2425748	3	0	1	0	0	
10.397	693.2691248	2	1	1	0	0	
11.422	855.3219448	3	1	1	0	0	
14.257	1017.374765	4	1	1	0	0	
13.77	1058.401315	3	2	1	0	0	

Figure 3.

Example bioinformatic library for fucosylated BMO. Each entry includes retention time (RT), accurate mass, oligosaccharide composition (Hex (hexose-galactose in yellow and glucose in blue circles), HexNAc (*N*-acetylhexosamine in white squares) Fuc (fucose-red triangles), NeuAc, and NeuGc (*N*-glycolylneuraminic acid) and full/partial structure.

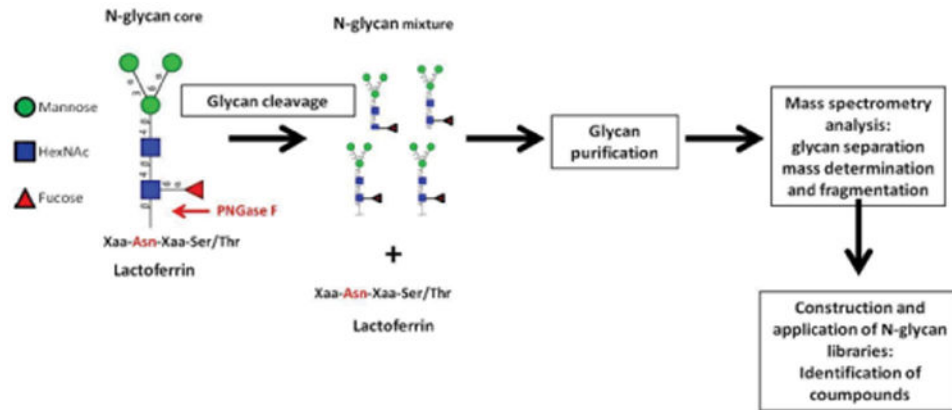


Figure 4.

Schematic of enzymatic N-glycan release from protein and workflow for characterization of the released glycans.

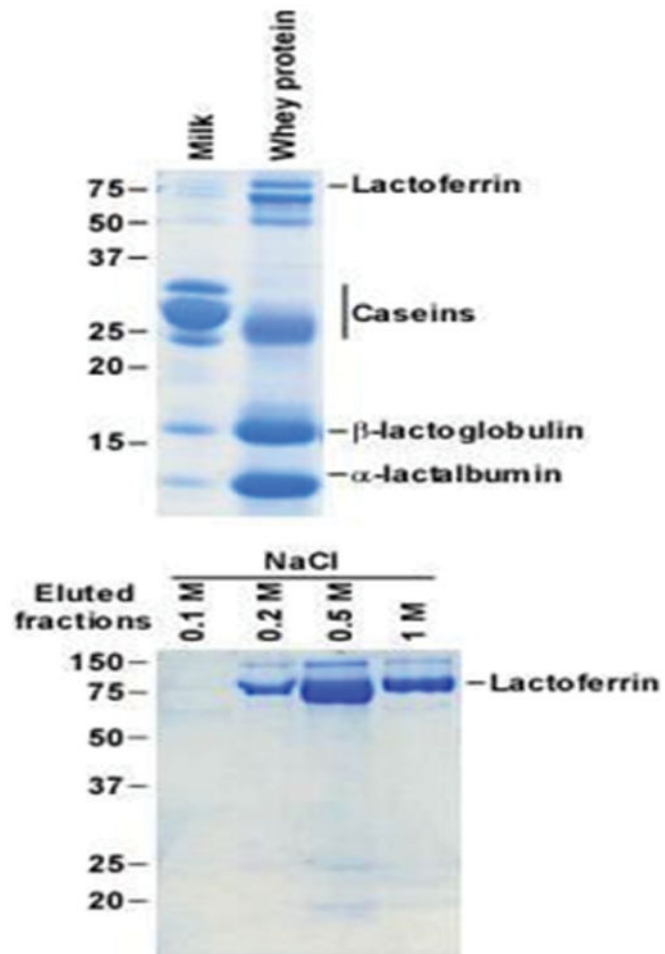


Figure 5.

Electrophoretic gels demonstrate the sequential successful purification of target whey proteins (e.g. lactoferrin from the remaining whey proteins). Isolated fractions were analyzed on 12% SDS-PAGE after affinity chromatography.

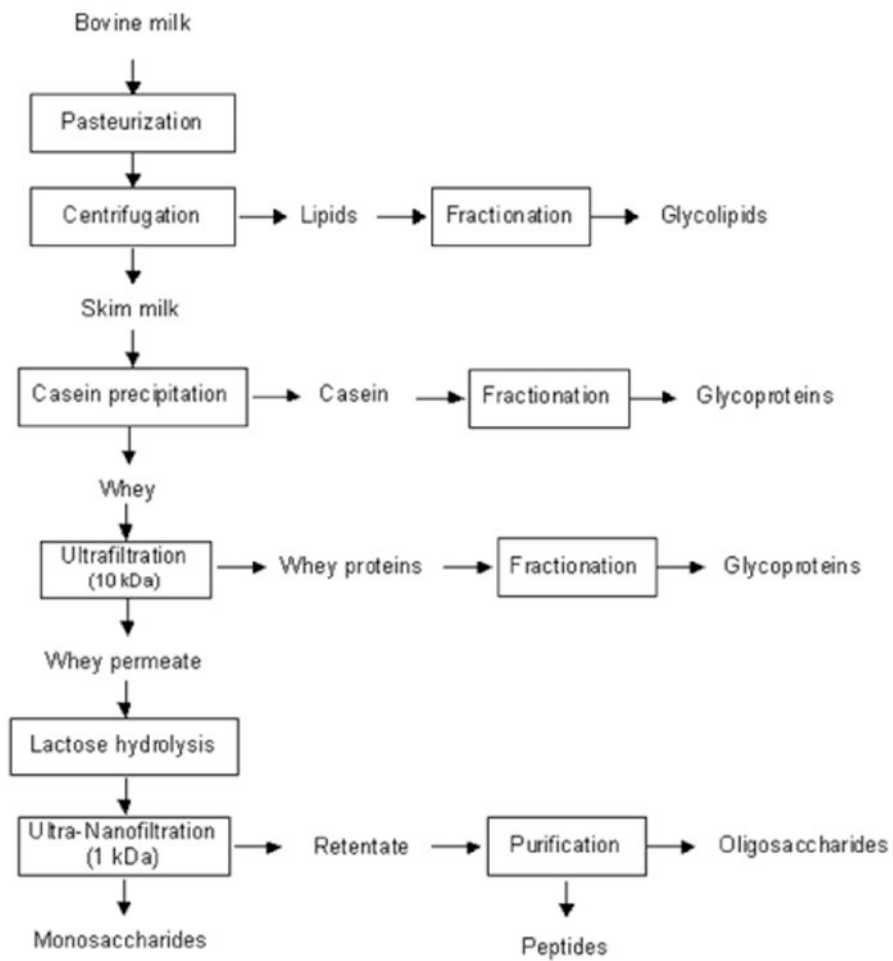


Figure 6. Schematic of potential pathways to maximize the recovery of milk bioactive components.