

# UC Davis

## UC Davis Previously Published Works

### Title

Comprehensive peptidomic and glycomic evaluation reveals that sweet whey permeate from colostrum is a source of milk protein-derived peptides and oligosaccharides

### Permalink

<https://escholarship.org/uc/item/61r2t9k7>

### Journal

Food Research International, 63(Pt B)

### ISSN

0963-9969

### Authors

Dallas, David C  
Weinborn, Valerie  
de Moura Bell, Juliana MLN  
[et al.](#)

### Publication Date

2014-09-01

### DOI

10.1016/j.foodres.2014.03.021

Peer reviewed

Published in final edited form as:

*Food Res Int.* 2014 September 1; 63(Pt B): 203–209. doi:10.1016/j.foodres.2014.03.021.

## Comprehensive peptidomic and glycomic evaluation reveals that sweet whey permeate from colostrum is a source of milk protein-derived peptides and oligosaccharides

David C. Dallas<sup>a,b,\*</sup>, Valerie Weinborn<sup>a</sup>, Juliana M.L.N. de Moura Bell<sup>a</sup>, Meng Wang<sup>c</sup>, Evan A. Parker<sup>d</sup>, Andres Guerrero<sup>d</sup>, Kasper A. Hettinga<sup>c</sup>, Carlito B. Lebrilla<sup>b,d</sup>, J. Bruce German<sup>a,b</sup>, and Daniela Barile<sup>a,b</sup>

<sup>a</sup>Department of Food Science and Technology, University of California, Davis, One Shields Avenue, Davis, CA 95616, United States <sup>b</sup>Foods for Health Institute, University of California, Davis, One Shields Avenue, Davis, CA 95616, United States <sup>c</sup>Dairy Science and Technology Group, Food Quality and Design Group, Wageningen University, Bomenweg 2, Wageningen 6703HD, The Netherlands <sup>d</sup>Department of Chemistry, University of California, Davis, One Shields Avenue, Davis, CA 95616, United States

### Abstract

Whey permeate is a co-product obtained when cheese whey is passed through an ultrafiltration membrane to concentrate whey proteins. Whey proteins are retained by the membrane, whereas the low-molecular weight compounds such as lactose, salts, oligosaccharides and peptides pass through the membrane yielding whey permeate. Research shows that bovine milk from healthy cows contains hundreds of naturally occurring peptides – many of which are homologous with known antimicrobial and immunomodulatory peptides – and nearly 50 oligosaccharide compositions (not including structural isomers). As these endogenous peptides and oligosaccharides have low-molecular weight and whey permeate is currently an under-utilized product stream of the dairy industry, we hypothesized that whey permeate may serve as an inexpensive source of naturally occurring functional peptides and oligosaccharides. Laboratory fractionation of endogenous peptides and oligosaccharides from bovine colostrum sweet whey was expanded to pilot-scale. The membrane fractionation methodology used was similar to the methods commonly used industrially to produce whey protein concentrate and whey permeate. Pilot-scale fractionation was compared to laboratory-scale fractionation with regard to the identified peptides and oligosaccharide compositions. Results were interpreted on the basis of whether industrial whey permeate could eventually serve as a source of functional peptides and oligosaccharides. The majority (96%) of peptide sequences and the majority (96%) of oligosaccharide compositions found in the laboratory-scale process were mirrored in the pilot-scale process. Moreover, the pilot-scale process recovered an additional 33 peptides and 1

© 2014 Elsevier Ltd. All rights reserved.

\*Corresponding author at: Department of Food Science and Technology, University of California, Davis, CA 95616, United States. Tel.: +1 530 752 4137; fax: +1 530 752 4759. dcdallas@ucdavis.edu (D.C. Dallas).

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.foodres.2014.03.021>.

oligosaccharide not identified from the laboratory-scale extraction. Both laboratory- and pilot-scale processes yielded peptides deriving primarily from the protein  $\beta$ -casein. The similarity of the laboratory- and pilot-scale's resulting peptide and oligosaccharide profiles demonstrates that whey permeate can serve as an industrial-scale source of bovine milk peptides and oligosaccharides.

## Keywords

Whey permeate; Colostrum; Pilot-scale; Peptidomic; Bovine; Oligosaccharide

---

## Introduction

Peptidomics and glycomics are recent technologies that are allowing food scientists to develop a detailed molecular-level understanding of functional compounds in food and food byproducts. Bovine milk contains fragments of intact milk proteins that are visible by gel electrophoresis and referred to as the proteose peptone fraction (Andrews, 1979; Andrews & Alichanidis, 1983; Eigel, 1981; Paquet, Alais, & Aubert, 1989). Mass spectrometry-based analysis has revealed that among this peptide fraction are hundreds of endogenous peptide sequences (Dallas, Guerrero, Khaldi, et al., 2013; Dallas, Guerrero, Parker, et al., 2013). These peptides are released by a select group of proteases that occur naturally in bovine milk, including plasmin (Andrews, 1983; Politis, 1996) and cathepsins B and D (Hinz, Larsen, Wellnitz, Bruckmaier, & Kelly, 2012; Larsen, Boisen, & Petersen, 1993) and elastase (Dallas, Underwood, Zivkovic, & German, 2012). Bovine milk also contains at least 49 oligosaccharides (polymers of several monosaccharides) by composition (shown in Supplementary Table 1), not including isomers; and 62 including isomers (Barile et al., 2010; Barile, Meyrand, Lebrilla, & German, 2011; Barile et al., 2009; Mariño et al., 2011; Parkkinen & Finne, 1987; Saito, Itoh, & Adachi, 1984, 1987; Schneir & Rafelson, 1966; Sundekilde et al., 2012; Tao et al., 2008; Urashima, Saito, Ohmisya, & Shimazaki, 1991; Urashima, Taufik, Fukuda, & Asakuma, 2013; Veh et al., 1981).

In cheese making, the liquid remaining after rennet-based casein precipitation is called “sweet whey.” In the past, whey was treated as a waste product. Now, whey is recognized as a source of potentially valuable compounds. As a result of scientific and technological innovations, whey now serves as an economically important source of functional ingredients for value-added foods. The major products derived from whey are whey protein concentrate and whey protein isolate (Ramchandran & Vasiljevic, 2013). Membrane technology is increasingly deployed as a non-destructive technique for isolation of whey proteins (Marcelo & Rizvi, 2008; Tunick, 2008).

Membrane filtration of whey produces a protein-rich retentate (concentrated whey proteins) and a permeate. Whey permeate is still considered a waste product by the dairy industry and the economic value of this stream remains low (Barile et al., 2009). As the production of whey protein concentrate typically employs 10 kDa membranes (Marcelo & Rizvi, 2008), it was hypothesized that oligosaccharides and protein fragments (peptides) that are smaller than 10 kDa pass through the membrane into the whey permeate phase.

In this research, sweet whey from bovine colostrum was processed by membrane filtration on a pilot-scale as a model for industrial whey permeate products. The permeate was analyzed for peptides and oligosaccharides by nano-liquid chromatography tandem mass spectrometry and database searching. Peptides and oligosaccharides present in this permeate were compared to those present from the laboratory-scale filtration of the same bovine colostrum whey. Furthermore, peptides were assessed for which enzymes were most active in the bovine colostrum sample. This work was conducted to determine whether milk's naturally occurring peptides and oligosaccharides exist within whey permeate, and, thus, whether this whey permeate can serve as an industrial-scale source for these compounds.

## Materials and methods

### Whey from bovine colostrum

Bovine colostrum whey was a gift from Sterling Technology (Brookings, South Dakota, USA). Whey was produced by lipid removal via cream separators and followed by High Temperature Short Time pasteurization and rennet-based casein precipitation. To eliminate lactose contamination of the oligosaccharide fraction, lactose was hydrolyzed by the addition of 0.1% (wt/wt) fungal lactase and incubated for 30 min at 45 °C with constant stirring.

### Sample preparation

An overview of the methodology employed to produce whey permeate from colostrum whey at the laboratory- and pilot-scale is presented in Fig. 1.

### Laboratory-scale production of peptides from whey permeate

For comparison to the pilot-scale whey permeate sample, the colostrum whey (after lactose hydrolysis) was processed in the laboratory with 30 k Damolecular weight cut-off centrifugal filter devices (Amicon Ultra-15 Centrifugal Devices, Millipore, Cork, Ireland) to create a laboratory-scale whey permeate. The membranes in each centrifugal device were made of Ultracel low-binding regenerated cellulose and had an active membrane area of 7.6 cm<sup>2</sup>. One milliliter of the original whey was split into 4 aliquots of 0.25 mL each. Each aliquot was mixed with 8.75 mL water (18-fold dilution) and centrifuged in a swinging bucket rotor at 4000 ×g, 20 min at 4 °C. After centrifugation, the permeate was collected, 9 mL of nanopure water was added to the retentate, and the sample was mixed using a vortex mixer and centrifuged again with the same conditions. This procedure was repeated for a total of 5 wash steps. The combined permeate solution was dried to 10 mL by centrifugal evaporation at 37 °C.

### Pilot-scale production of peptides from whey permeate

The whey was subjected to a two-stage filtration using a tangential-flow membrane system at pilot-plant scale (Model L, GEA Filtration, Hudson, WI, USA). The system was composed of a 2.5-in. diameter spiral membrane housing, a 95 L jacketed stainless steel reactor, a flow-meter, a heat exchanger and a feed pump (7.5 Hp). Ninety five liters of bovine colostrum whey were ultrafiltered as a single batch with a 10 kDa molecular weight cut-off polyethersulfone spiral-wound membrane with an effective area of 1.86m<sup>2</sup>. The

they was concentrated to a concentration factor of 5.3 (concentration factor = volume of feed/volume of retentate). This ultrafiltration and whey concentration produced a protein-rich retentate and a sugar- and peptide-rich permeate. Ultrafiltration was performed at 40–43 °C, with a transmembrane pressure of 3.0 bar and a recirculation flow rate of 10 L/min. The resulting permeate was nanofiltered using a 500–700 Da molecular weight cut-off sulfonated polyethersulfone spiral-wound membrane with an effective area of 1.86 m<sup>2</sup>. The nanofiltration retentate was concentrated to a concentration factor of 8.5 to produce a peptide- and oligosaccharide-rich retentate and a monosaccharide- and salt-rich permeate. Nanofiltration was performed at 45–50 °C, with a transmembrane pressure of 20 bar and a recirculation flow rate of 9.5 L/min. Four diafiltration steps were performed to improve monosaccharide removal from the nanofiltration retentate. All membranes were manufactured by Hydranautics (Oceanside, CA, USA). The peptide- and oligosaccharide-rich retentate from the nanofiltration step was used. After this membrane filtration fractionation, the final whey permeate samples were immediately iced, transferred to the laboratory and stored at –30 °C. Prior to peptide extraction, 100 mL of whey permeate was centrifuged at 4000 ×g for 10 min at 4 °C. Precipitates from remaining salts were removed and the supernatant was collected. The supernatants were centrifuged again with the same parameters and the second supernatant was collected. The sample was then dried to 10 mL by a speed vacuum concentrator at 37 °C.

### **C18 extraction of peptides**

To remove sugars and salts from the laboratory- and pilot-scale whey permeates, C18 solid-phase extraction was performed. The whey permeates were split into five 2-mL aliquots and each aliquot was applied to a 1-g bed C18 solid-phase extraction cartridge (Supelco, Bellefonte, PA, USA). Columns were prepared with 80% acetonitrile (ACN), 0.1% trifluoroacetic acid (TFA) followed by 1% ACN, and 0.1% TFA. The whey permeate was added, and 6 column volumes (1 column volume: 2 mL) of 1% ACN, and 0.1% TFA were added to remove sugars and salts. Three column volumes of 80% ACN and 0.1% TFA were added to elute peptides. Peptides were dried by centrifugal evaporation at 37 °C. The workflow for all sample preparation steps is summarized in Fig. 1.

### **Laboratory-scale production of oligosaccharides from whey permeate**

The isolation of oligosaccharides was performed according to the protocol of Meyrand et al. (2013).

### **Pilot-scale production of oligosaccharides from whey permeate**

Pilot-scale whey permeate was produced according to the same procedure described above for peptide production. Starting with 0.5 mL of whey permeate samples, they were extracted according to the same method described for the above laboratory-scale production of oligosaccharides. The workflow for the oligosaccharide isolation and analysis is shown in Fig. 1.

## Mass spectrometry

Peptides were rehydrated in nanopure water and the peptide concentration was determined by the Bradford spectrophotometric assay. Peptides were prepared for mass spectrometric analysis at 10 ng/μL concentration (for a 10 ng injection) and were analyzed by nanoliquid chromatography mass spectrometry using modifications to a published procedure (Dallas, Guerrero, Khaldi, et al., 2013; Dallas, Guerrero, Parker, et al., 2013). The mass spectrometer employed was the Agilent 6520 (Santa Clara, CA) nano-liquid chromatography-chip-quadrupole time-of-flight mass spectrometer (Chip-Q-TOF) using a C18 chip. Minor modifications to the protocol included using a drying gas flow rate of 3 L/min and a 1 spectra/s spectral acquisition rate.

Oligosaccharides were rehydrated in nanopure water to achieve a 200-fold dilution and injected on the Chip-Q-TOF. Liquid chromatography was performed using a nano-chip with porous graphitized carbon as the solid phase for the enrichment and analytical columns. Mass spectrometric analysis was carried out by modifications of a published procedure (Meyrand et al., 2013); modifications included data acquisition with a 380–2500 mass/charge range and an electrospray capillary voltage of 1850 V.

## Spectral analysis

Spectra were analyzed by database searching according to a published procedure (Dallas, Guerrero, Khaldi, et al., 2013; Dallas, Guerrero, Parker, et al., 2013). Briefly, data were exported from Agilent MassHunter in.mgf format, imported into the offline search engine X! Tandem (Craig & Beavis, 2004) and searched against a bovine milk library compiled from previous bovine milk proteome literature (Reinhardt & Lippolis, 2006, 2008; Wilson et al., 2008). The bovine milk protein library was imported in FASTA file format to X! Tandem. Identified peptides were accepted if e-values were  $\leq 0.01$  corresponding to a confidence level of 99%. Peptide mass tolerance was 20 ppm. No complete (required) modifications were included, but potential modifications allowed were phosphorylation of serine, threonine and tyrosine; oxidation of methionine and tryptophan; deamidation of glutamine and asparagine; and *N*-terminal acetylation. A non-specific cleavage ([X][X]) (where 'X' is any amino acid) was used to search against the protein sequences. Because the instrument did not always select the monoisotopic ion for tandem fragmentation, isotope errors were allowed (allowing up to one C13). No model refinement was employed in X! Tandem.

## Library search

The results from X!Tandem for all samples were compiled. To retain only unique peptide sequences, all duplicates of sequence, protein and modifications combined were eliminated with the “remove duplicates” function in Excel. Peptides representing identical amino acid sequences and modifications but modified in different positions were also removed as duplicates.

## Peptide identification

Sample compounds were identified by the “Find by Molecular Feature” function in Agilent MassHunter Qualitative Analysis version 6. The target data type was small molecule (chromatographic). The retention time was restricted from 3.5 to 36 min (the range of library peptides). Only peaks with at least 1000 ion counts were selected. The selected charge carrier was “protons.” The isotope model was “peptides,” and the maximum assigned charge was 7. After compounds were extracted, they were matched to the peptide library by mass and retention time to identify peptides. The database match employed a mass tolerance of 40 ppm and a retention time tolerance of 3 min. Each exact peptide mass in the library was converted to an  $m/z$  for all charge states from 1 to 7 for the search. The overall workflow for this data processing methodology is shown in Fig. 1.

## Oligosaccharide identification

A library of 49 known bovine milk oligosaccharide compositions was assembled from literature, not including isomers (shown in Supplementary Table 1). This library was used to identify oligosaccharides in the samples based on mass and retention time. Compounds were identified manually using the “Extract Chromatogram” feature in Agilent MassHunter Qualitative Analysis. Extracted ion chromatograms were built for each known oligosaccharide  $m/z$  with a 20 ppm mass tolerance.

## Enzymatic analysis

A custom program written in Python was used to estimate the activity of selected enzymes. This program analyzes the resulting peptides identified by library search for each sample. First, each peptide's start and end positions within the precursor protein were confirmed by sequence matching. Both termini of each peptide were compared to a selected set of proteolytic enzyme rules (in the present work, to the enzyme's plasmin, cathepsin and elastase). Enzymatic cleavage rules were derived from a list published on ExPASy (Gasteiger et al., 2005). As a measure of simplification, rules were assumed to only act on P1 and P1'. P1 is the amino acid position directly before the cleavage site on the *N*-terminal side, whereas P1' is the amino acid after the cleavage on the *C*-terminal side. The enzyme specificity patterns used in the algorithm for evaluating cleavages are shown in Supplementary Table 2. Peptides having termini that pass a comparison to an enzymatic rule have their mass spectral intensity (peak volume) added to the sum of the respective enzyme. In the case that a peptide matches multiple rules simultaneously, the full intensity of that peptide is added to both enzymatic sums, thus the intensity value output represents potential activity rather than uniquely specific activity. Peptides failing all enzymatic comparisons have their intensity added to a list of remainders whose purpose is to assist in the establishment of other enzymes that are active in bovine colostrum.



## Results

### Comparison of peptides and oligosaccharides identified in laboratory- and pilot-scale extractions

An example confirmation of a peptide identification by X!Tandem is shown in Fig. 2. A large number of peaks were detected that matched known fragments of the peptide HQGLPLQEVL, confirming this identification. Overall, 245 peptides were present in all the samples tested (shown in Supplementary Table 3). The pilot-scale samples contained more unique peptides than the laboratory-scale samples: 238 vs. 212 unique peptides. Most peptides (205) were found in samples from both methodologies. Thirty-three peptides were found only in the pilot-scale samples and 7 were unique to the laboratory-scale samples.

Overall, 26 oligosaccharide compositions were present in the samples tested (shown in Table 1) (structural isomers were not determined). The pilot-scale samples contained the same number of oligosaccharide compositions as the laboratory-scale samples (26). The majority of oligosaccharides were identified in samples from both preparation methodologies: 24 oligosaccharides were identified in both samples and only one unique oligosaccharide composition was identified in each sample type.

The peptides found in these experiments are derived from 14 proteins. However, the majority of the peptides identified are derived from  $\beta$ -casein,  $\alpha_{s1}$ -casein, glycosylation-dependent cellular adhesion molecule 1, polymeric immunoglobulin receptor, serum albumin,  $\kappa$ -casein and  $\alpha_{s2}$ -casein. The average number of amino acids was 13.3 and the range was 7–25 amino acids. Peptides with fewer than 7 amino acids in the sample may go unidentified by X!Tandem. Smaller peptides have fewer resulting fragment peaks, which leads to lower X!Tandem e-value scores. The average peptide mass was 1502.728 Da (range 687.346–2925.334). No glycopeptides were searched for with the methodology employed; therefore no glycopeptides were identified.

The total number of unique peptides for each protein according to the method of preparation is shown in Fig. 3. The relative abundances of peptides from each protein were similar between the two extraction methodologies (demonstrated by Fig. 4, which presents the relative intensity of peptides from each protein by sample).

### Enzyme analysis

Analysis of estimated enzymatic activity revealed that some peptides extracted from whey permeate by both laboratory- and pilot-scale protocols were released by plasmin, cathepsin and elastase, however the majority of cleavage sites could not be assigned to these three enzymes (Fig. 5). For both pilot-scale and laboratory-scale samples, the percentage of peptide cleavage sites assigned to cathepsin, plasmin and elastase were similar. The remaining unassigned cleavage sites for the pilot-scale and laboratory-scale samples represented enzymes that were not included in the analysis (Fig. 6).

This enzyme pattern is different from that identified in mature bovine milks, in which plasmin accounts for the majority of cleavage sites (Dallas, Guerrero, Khaldi, et al., 2013; Dallas, Guerrero, Parker, et al., 2013). This difference in enzyme activity may be due to



differences in enzyme activity between colostrum and mature bovine milk. This difference could also be due to enzymatic digestion from microbial enzymes during whey processing. More research is needed to explain the unassigned cleavage sites in colostrum whey.

## Discussion

Cheese production from mature bovine milk starts from a rennet-induced casein precipitation. This precipitation leaves the whey as a separate product stream. Whey is usually processed by membrane filtration to create whey protein concentrate – which is used as a nutritional supplement and an ingredient in a variety of foods – and permeate. Bovine colostrum is also processed to produce whey, which is further processed by membrane filtration to concentrate the whey proteins (mostly immunoglobulins). Isolated colostrum whey proteins are sold worldwide as health-promoting supplements. The production of bovine colostrum-based protein products leaves behind a permeate that, analogous to cheese whey permeate, contains potentially functional low-molecular weight compounds. This study developed a pilot-scale method to mimic the processing performed commonly by the dairy and colostrum industries, and compared the resulting permeate's peptide and oligosaccharide profiles with those from laboratory-scale extraction processes. Hundreds of peptides and 26 oligosaccharides were identified in both pilot-scale and laboratory-scale sweet colostrum whey permeates. The pilot-scale processing produced 96% of all peptides and 96% of all oligosaccharides identified in the laboratory-scale processing. This result was interpreted to indicate that industrial colostrum whey permeate is an otherwise unused and valuable source of endogenous bovine milk peptides and oligosaccharides.

The representation of proteins as peptides does not strictly follow the natural abundances of proteins in colostrum. The presence of a large number of peptides from  $\beta$ -,  $\alpha_{s1}$ -,  $\alpha_{s2}$ - and  $\kappa$ -caseins is interpreted to indicate that these proteins are uniquely susceptible to proteolytic cleavage. Other abundant milk proteins, notably lactoferrin,  $\alpha$ -lactalbumin and immunoglobulins, are not present as milk peptides. The casein proteins' susceptibility to cleavage may be due to their looser, rheomorphic tertiary structure (Livney, Schwan, & Dalgleish, 2004; Swaisgood, 2003) in comparison to the more tightly packed, specific tertiary structures of proteins like lactoferrin,  $\alpha$ -lactalbumin and the immunoglobulins (Legrand et al., 2008; Permyakov & Berliner, 2000; Schroeder & Cavacini, 2010). The lack of these proteins present as peptides in the extracted samples is unlikely to be due to peptide losses in sample preparation. Research in our group has shown that, after *in vitro* digestion of milk with trypsin and chymotrypsin, lactoferrin,  $\alpha$ -lactalbumin and immunoglobulin peptide fragments are successfully eluted from the C18 solid phase extraction column with 60% ACN and 0.1% TFA (Vijayakumar et al., 2012) which is less nonpolar than the elution solution employed in the present research (80% ACN, 0.1% TFA).

The laboratory-scale protocol retained more intact proteins than the pilot-scale protocol. Intact proteins were completely eliminated by the pilot-scale 10-kDa filtration step, whereas the laboratory-scale preparation maintained small intact proteins because it uses a 30-kDa filtration step. This 30-kDa filtration was employed for the laboratory-scale preparation because in previous testing 10-kDa laboratory-scale membranes did not allow for peptide permeation through the membrane. This poor recovery with the 10-kDa membrane at the

laboratory-scale is thought to be caused by membrane fouling due to the use of a dead-ended membrane as opposed to a cross-flow membrane in the pilot-scale setup. The presence of the intact proteins  $\alpha$ -lactalbumin (16.2 kDa) and  $\beta$ -lactoglobulin (19.9 kDa) in the laboratory-scale samples was confirmed by both mass spectrometry and gel electrophoresis. Overall, these results indicate that the pilot-scale extraction retained the diversity of peptides extracted by the laboratory-scale setup, but with less contamination from intact proteins.

Both the laboratory-scale and pilot-scale peptide extractions likely include glycopeptides. Diagnostic ions representing the losses of glucose and *N*-acetylglucosamine were detected in tandem spectra for several precursors that were not identified as peptides. However, collision-induced dissociation, which was employed in this study, fragments the glycan component but releases only few peptide fragments (Mechref, 2001). As the peptide identification approach relies on the extensive fragmentation of the peptide, this fragmentation method did not allow for identification of these endogenous milk peptides. The combination of electron transfer dissociation, which extensively fragments the peptide backbone, and collision-induced dissociation could resolve this problem (Mechref, 2001). Identification of the peptide sequence, site of glycosylation and glycan composition for nonspecifically cleaved endogenous peptides in complex biological mixtures remains a major analytical and computational challenge (Dallas, Martin, Hua, & German, 2012).

This research demonstrates that whey permeate is a source of oligosaccharides and endogenous milk peptides. As these components have demonstrated biological actions, this waste product can serve as a source for functional compounds for novel product development.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors thank Sterling Technology for the donation of the whey used for this research. The authors thank Cora J. Dillard for editing this manuscript.

The authors gratefully acknowledge funding from the Bill and Melinda Gates Foundation, the National Institutes of Health (R01 AT007079, R01 HD059127 and UL1 TR000002), the Peter J. Shields Endowed Chair in Dairy Food Science, the National Science Foundation Graduate Research Fellowship Program, and the USDA National Institute for Food and Agriculture Post-doctoral Fellowship.

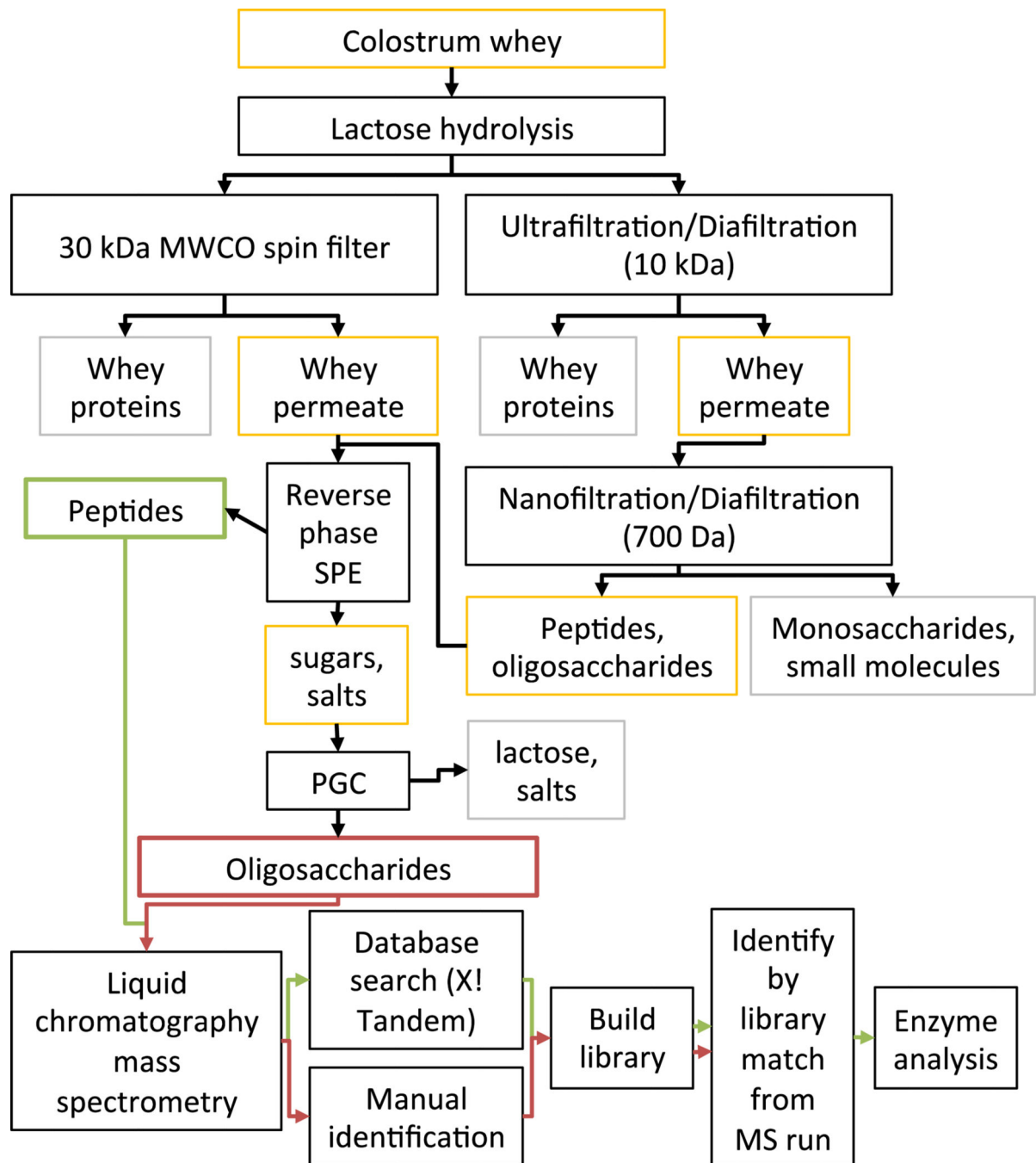
## Abbreviations

<b>ACN</b>	acetonitrile
<b>TFA</b>	trifluoroacetic acid
<b>FA</b>	formic acid
<b>Q-TOF</b>	quadrupole time-of-flight.

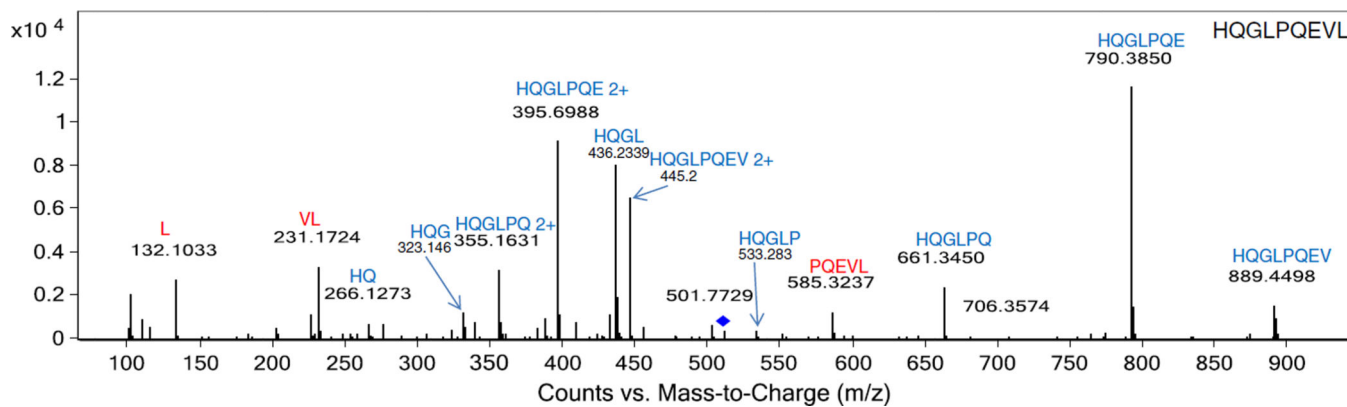
## References

- Andrews AT. The formation and structure of some proteose-peptone components. *Journal of Dairy Research*. 1979; 46(02):215–218. [PubMed: 469045]
- Andrews AT. Proteinases in normal bovine milk and their action on caseins. *Journal of Dairy Research*. 1983; 50:45–55. [PubMed: 6341422]
- Andrews AT, Alichanidis E. Proteolysis of caseins and the proteose-peptone fraction of bovine milk. *Journal of Dairy Research*. 1983; 50(03):275–290.
- Barile D, Marotta M, Chu C, Mehra R, Grimm R, Lebrilla CB, 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. *Journal of Dairy Science*. 2010; 93(9):3940–3949. [PubMed: 20723667]
- Barile D, Meyrand M, Lebrilla C, German J. Examining bioactive components of milk. Sources of complex oligosaccharides (Part 2). *Agro Food Industry Hi Tech*. 2011; 22(4):37–40.
- Barile D, Tao N, Lebrilla CB, Coisson JD, Arlorio M, German JB. Permeate from cheese whey ultrafiltration is a source of milk oligosaccharides. *International Dairy Journal*. 2009; 19(9):524–530. [PubMed: 20161544]
- Craig R, Beavis RC. TANDEM: Matching proteins with tandem mass spectra. *Bioinformatics*. 2004; 20(9):1466–1467. [PubMed: 14976030]
- Dallas DC, Guerrero A, Khaldi N, Castillo PA, Martin WF, Smilowitz JT, et al. Extensive in vivo human milk peptidomics reveals specific proteolysis yielding protective antimicrobial peptides. *Journal of Proteome Research*. 2013; 12(5):2295–2304. [PubMed: 23586814]
- Dallas DC, Guerrero A, Parker EA, Garay LA, Bhandari A, Lebrilla CB, et al. Peptidomic profile of milk of Holstein cows at peak lactation. *Journal of Agricultural and Food Chemistry*. 2013; 62(1): 58–65. [PubMed: 24344900]
- Dallas DC, Martin WF, Hua S, German JB. Automated glycopeptide analysis—Review of current state and future directions. *Briefings in Bioinformatics*. 2012; 14(3):361–374. [PubMed: 22843980]
- Dallas DC, Underwood MA, Zivkovic AM, German JB. Digestion of protein in premature and term infants. *Journal of Nutritional Disorders and Therapy*. 2012; 2(112):1–9.
- Eigel W. Identification of proteose-peptone component 5 as a plasmin-derived fragment of bovine  $\beta$ -casein. *International Journal of Biochemistry*. 1981; 13(10):1081–1086. [PubMed: 6457763]
- Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.; Appel, R., et al. Protein identification and analysis tools on the ExPASy server. In: Walker, JM., editor. *The proteomics protocols handbook*. Totowa, NJ, USA: Humana Press; 2005. p. 571-607.
- Hinz K, Larsen LB, Wellnitz O, Bruckmaier R, Kelly A. Proteolytic and proteomic changes in milk at quarter level following infusion with *Escherichia coli* lipopolysaccharide. *Journal of Dairy Science*. 2012; 95(4):1655–1666. [PubMed: 22459814]
- Larsen LB, Boisen A, Petersen TE. Procathepsin D cannot autoactivate to cathepsin D at acid pH. *FEBS Letters*. 1993; 319(1):54–58. [PubMed: 8454061]
- Legrand D, Pierce A, Ellass E, Carpentier M, Mariller C, Mazurier J. Lactoferrin structure and functions. *Advances in Experimental Medicine and Biology*. 2008; 606:163–194. [PubMed: 18183929]
- Livney YD, Schwan AL, Dalgleish DG. A study of  $\beta$ -casein tertiary structure by intramolecular crosslinking and mass spectrometry. *Journal of Dairy Science*. 2004; 87(11):3638–3647. [PubMed: 15483147]
- Marcelo PA, Rizvi SS. Physicochemical properties of liquid virgin whey protein isolate. *International Dairy Journal*. 2008; 18(3):236–246.
- Mariño K, Lane JA, Abrahams JL, Struwe WB, Harvey DJ, Marotta M, et al. Method for milk oligosaccharide profiling by 2-aminobenzamide labeling and hydrophilic interaction chromatography. *Glycobiology*. 2011; 21(10):1317–1330. [PubMed: 21566017]
- Mechref, Y. *Current protocols in protein science*. John Wiley & Sons, Inc; 2001. Use of CID/ETD mass spectrometry to analyze glycopeptides.

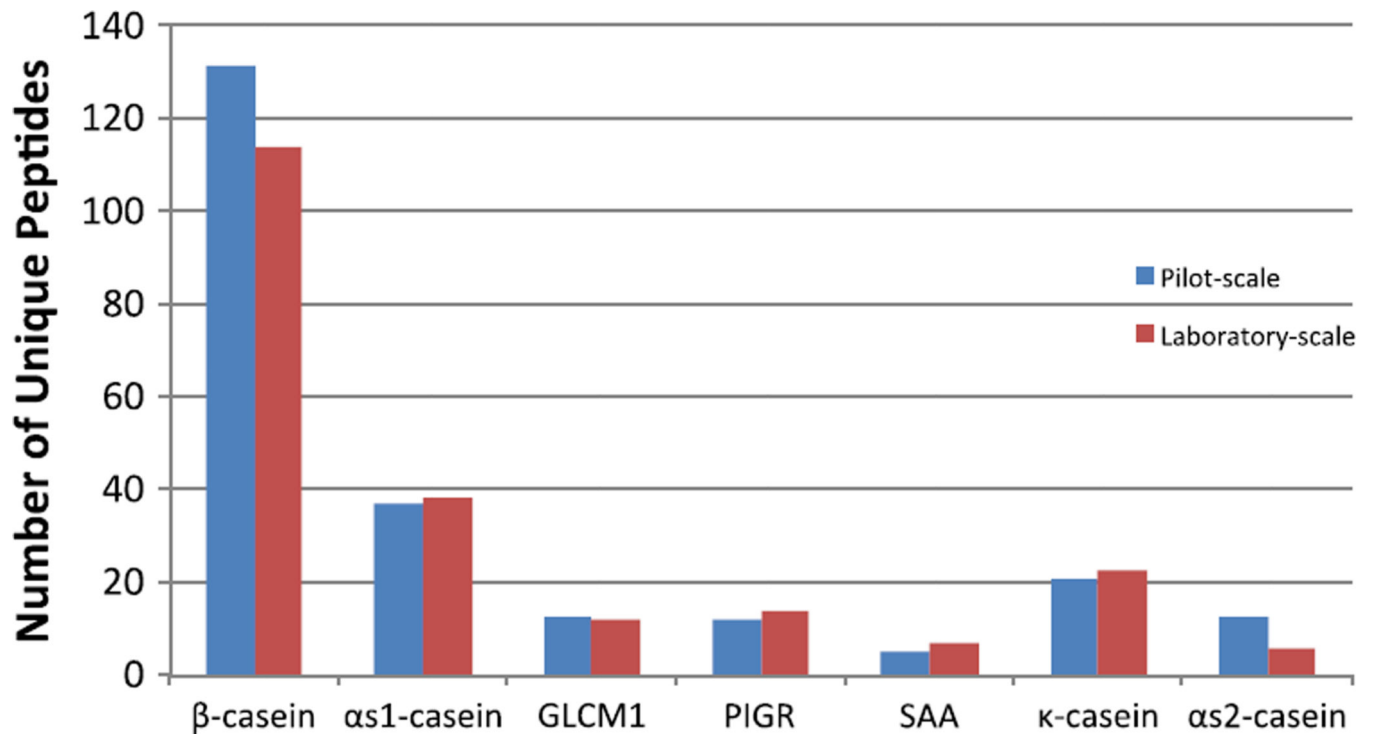
- Meyrand M, Dallas DC, Caillat H, Bouvier F, Martin P, Barile D. Comparison of milk oligosaccharides between goats with and without the genetic ability to synthesize  $\alpha$ 1-casein. *Small Ruminant Research*. 2013; 113(2–3):411–420. [PubMed: 24587592]
- Paquet D, Alais C, Aubert F. Review: The proteose-peptone fraction of milk. *Le Lait*. 1989; 69(1):1–21.
- Parkkinen J, Finne J. Isolation of sialyl oligosaccharides and sialyl oligosaccharide phosphates from bovine colostrum and human urine. *Methods in Enzymology*. 1987; 138:289–300. [PubMed: 3600326]
- Permyakov EA, Berliner LJ.  $\alpha$ -Lactalbumin: Structure and function. *FEBS Letters*. 2000; 473(3):269–274. [PubMed: 10818224]
- Politis I. Plasminogen activator system: Implications for mammary cell growth and involution. *Journal of Dairy Science*. 1996; 79(6):1097–1107. [PubMed: 8827475]
- Ramchandran, L.; Vasiljevic, T. *Membrane processing*. Blackwell Publishing Ltd; 2013. Whey processing; p. 193-207.
- Reinhardt TA, Lippolis JD. Bovine milk fat globule membrane proteome. *Journal of Dairy Research*. 2006; 73(04):406–416. [PubMed: 16834814]
- Reinhardt TA, Lippolis JD. Developmental changes in the milk fat globule membrane proteome during the transition from colostrum to milk. *Journal of Dairy Science*. 2008; 91(6):2307–2318. [PubMed: 18487653]
- Saito T, Itoh T, Adachi S. Presence of two neutral disaccharides containing N-acetylhexosamine in bovine colostrum as free forms. *Biochimica et Biophysica Acta - General Subjects*. 1984; 801(1):147–150.
- Saito T, Itoh T, Adachi S. Chemical structure of three neutral trisaccharides isolated in free form from bovine colostrum. *Carbohydrate Research*. 1987; 165(1):43–51. [PubMed: 3664520]
- Schneir ML, Rafelson ME. Isolation and characterization of two structural isomers of N-acetylneuraminylactose from bovine colostrum. *Biochimica et Biophysica Acta - General Subjects*. 1966; 130(1):1–11.
- Schroeder HW, Cavacini L. Structure and function of immunoglobulins. *Journal of Allergy and Clinical Immunology*. 2010; 125 Supplement 2(2):S41–S52. [PubMed: 20176268]
- Sundekilde UK, Barile D, Meyrand M, Poulsen NA, Larsen LB, Lebrilla CB, et al. Natural variability in bovine milk oligosaccharides from Danish Jersey and Holstein–Friesian breeds. *Journal of Agricultural and Food Chemistry*. 2012; 60(24):6188–6196. [PubMed: 22632419]
- Swaigood, HE. Chemistry of the caseins. In: Fox, PLHMPF., editor. *Advanced dairy chemistry*. 3rd ed.. New York, NY: Kluwer Academic/Plenum Publishers; 2003. p. 139-202.
- Tao N, DePeters EJ, Freeman S, German JB, Grimm R, Lebrilla CB. Bovine milk glycome. *Journal of Dairy Science*. 2008; 91(10):3768–3778. [PubMed: 18832198]
- Tunick MH. Whey protein production and utilization: A brief history. *Whey processing, functionality and health benefits*. 2008:8–9.
- Urashima T, Saito T, Ohmisya K, Shimazaki K. Structural determination of three neutral oligosaccharides in bovine (Holstein–Friesian) colostrum, including the novel trisaccharide; GalNAc $\alpha$ 1–3Gal $\beta$ 1–4Glc. *Biochimica et Biophysica Acta — General Subjects*. 1991; 1073(1):225–229.
- Urashima T, Taufik E, Fukuda K, Asakuma S. Recent advances in studies on milk oligosaccharides of cows and other domestic farm animals. *Bioscience, Biotechnology, and Biochemistry*. 2013; 77(3):455–466.
- Veh RW, Michalski JC, Corfield AP, Sander-Wewer M, Gies D, Schauer R. New chromatographic system for the rapid analysis and preparation of colostrum sialyloligosaccharides. *Journal of Chromatography A*. 1981; 212(3):313–322.
- Vijayakumar V, Guerrero A, Davey N, Lebrilla CB, Shields DC, Khaldi N. Enzyme Predictor: A tool for predicting and visualizing enzymatic cleavages of digested proteins. *Journal of Proteome Research*. 2012; 11(12):6056–6065. [PubMed: 23098558]
- Wilson NL, Robinson LJ, Donnet A, Bovetto L, Packer NH, Karlsson NG. Glycoproteomics of milk: Differences in sugar epitopes on human and bovine milk fat globule membranes. *Journal of Proteome Research*. 2008; 7(9):3687–3696. [PubMed: 18624397]



**Fig. 1.** Schematic representation of laboratory- (left) and pilot-scale (right) production of whey permeate, isolation of peptides and oligosaccharides and compound identification and analysis.

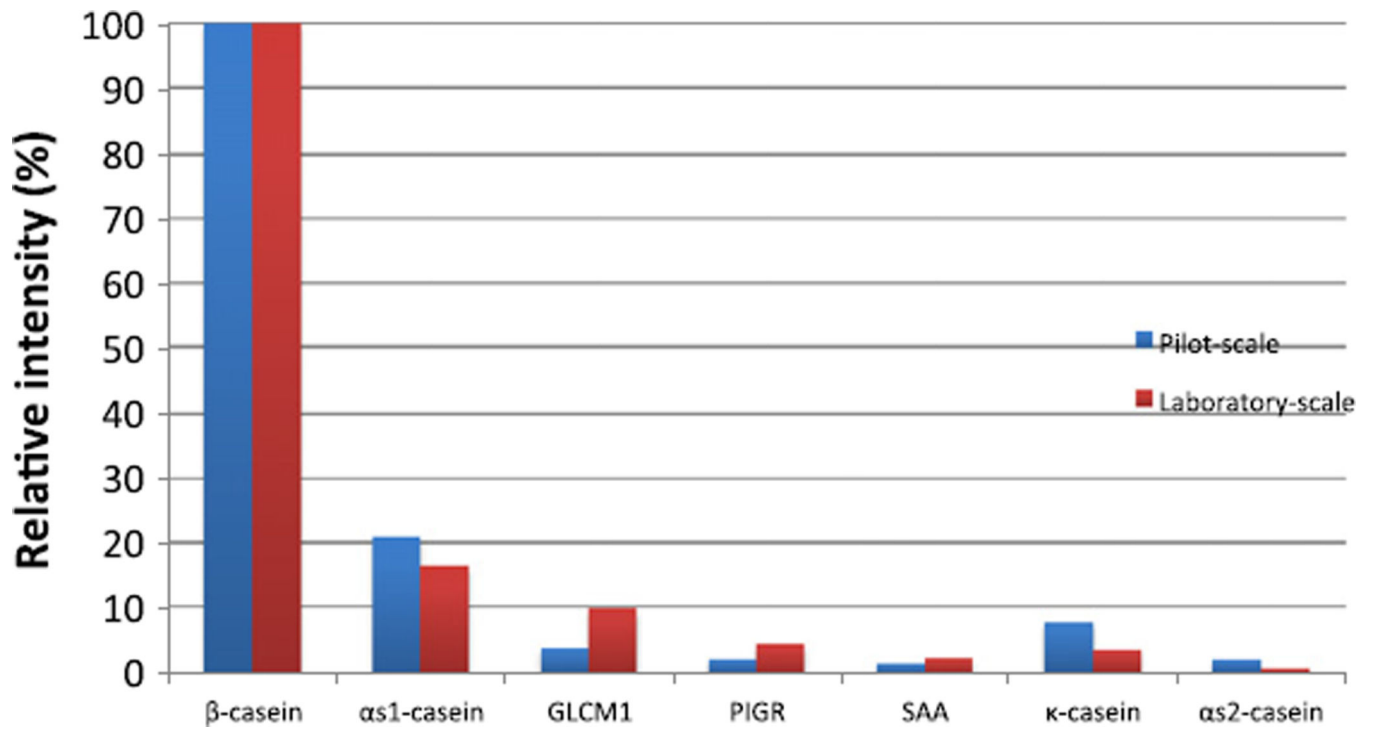


**Fig. 2.** Example of tandem spectral peptide identification performed by X!Tandem. Blue fragments are b-type ions, and red fragments are y-type ions. Fragments are  $z=1$  except those marked as  $2+$ , which are  $z=2$ . This spectrum confirmed the presence of the peptide HQGLPQEV from bovine  $\alpha_{s1}$ -casein in the pilot-scale whey permeate sample. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

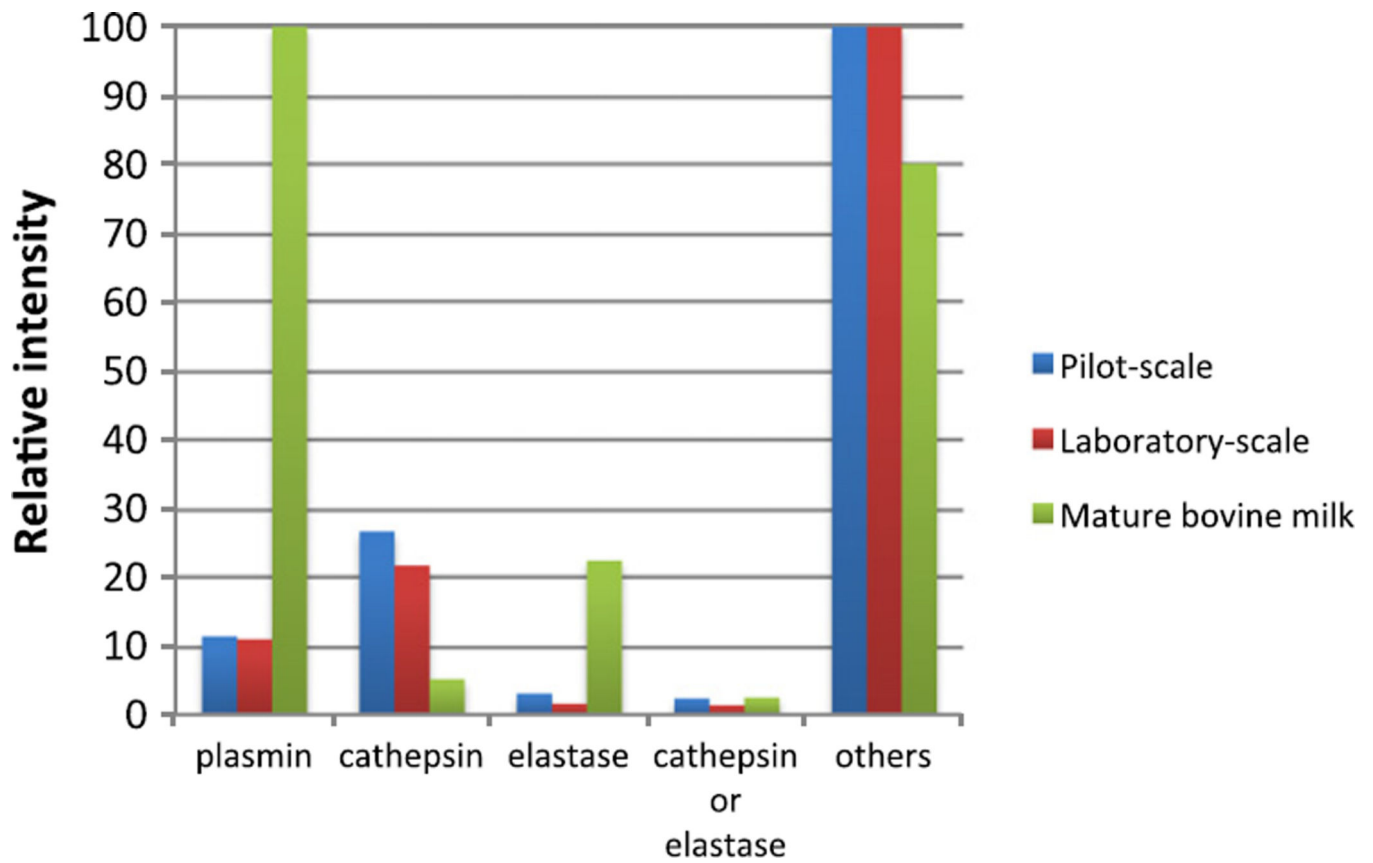


**Fig. 3.** Number of unique peptides for each protein in the laboratory-scale and pilot-scale methodology. GLCM1, glycosylation-dependent cell adhesion molecule 1; PIGR, polymeric immunoglobulin receptor; SAA, serum amyloid A protein; OGP, osteoglycin preprotein; ATPase A, ATP synthase subunit A; XDH, xanthine dehydrogenase.

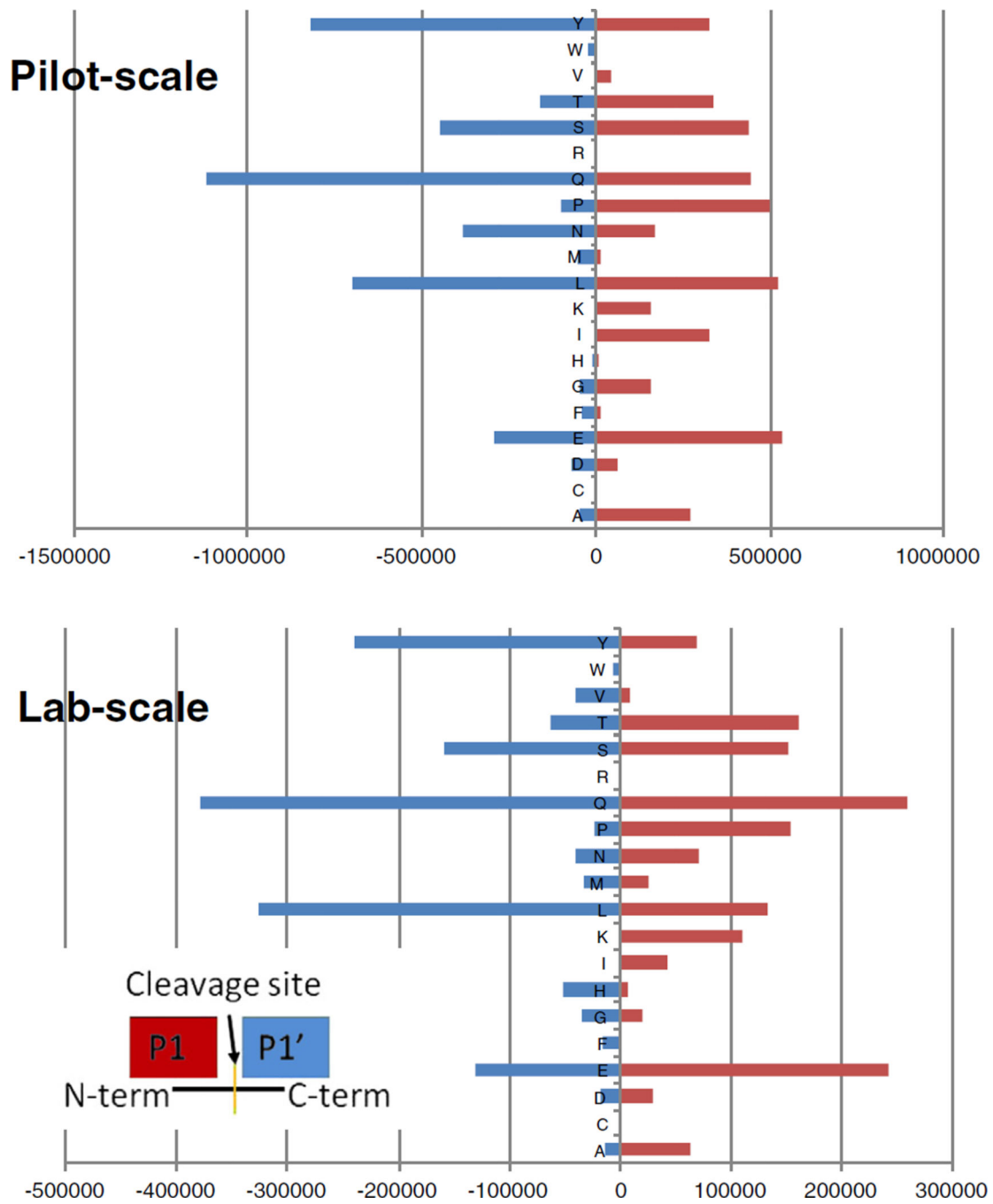




**Fig. 4.** Relative total intensity for all peptides by protein for the laboratory- and pilot-scale methodology. Values are scaled to the highest intensity value for each sample preparation methodology.



**Fig. 5.** Relative intensities (scaled to the highest intensity for each sample type) for all assigned and unassigned cleavage sites for samples prepared by pilot- and laboratory scale methodologies with comparison to average results from 6 bovine milk from previous research.



**Fig. 6.** Cumulative intensities (ion counts) for all unassigned cleavage sites by amino acid position for both methodologies.

**Table 1**

Unique oligosaccharides in each sample. The lab-scale and pilot-scale oligosaccharides are shown separately. Compositions are coded as follows: Hex, hexose; HexNAc, *N*-acetylhexosamine; Fuc, fucose; NeuAc, *N*-acetylneuraminic acid; and NeuGc, *N*-glycolylneuraminic acid.

Present laboratory-scale	Present pilot-scale	Neutral mass	Composition
X	X	633.212	2 Hex 1 NeuAc
X	X	674.238	1 Hex 1 HexNAc 1 NeuAc
X	X	545.196	2 Hex 1 HexNAc
X	X	504.169	3 Hex
X	X	666.222	4 Hex
X	X	690.233	1 Hex 1 HexNAc 1 NeuGc
X	X	924.307	2 Hex 2 NeuAc
X	X	707.248	3 Hex 1 HexNAc
X	X	795.264	3 Hex 1 NeuAc
X	X	383.143	1 Hex 1 HexNAc
X	X	649.207	2 Hex 1 NeuGc
X	X	869.301	4 Hex 1 HexNAc
X	X	998.344	3 Hex 1 HexNAc 1 Fuc
X	X	748.275	2 Hex 2 HexNAc
X	X	828.275	5 Hex
X	X	990.328	6 Hex
X	X	836.291	2 Hex 1 HexNAc 1 NeuAc
X	X	1160.397	4Hex 1HexNAc 1 NeuAc
X	X	910.328	3 Hex 2 HexNAc
X		1072.381	4 Hex 2 HexNAc
X	X	940.302	2 Hex 1 NeuAc 1 NeuGc
X	X	1201.423	3Hex 2HexNAc 1 NeuAc
X	X	1031.354	5Hex 1HexNAc
X	X	1039.370	2Hex 2HexNAc 1 NeuAc
X	X	811.259	3 Hex 1 NeuGc
	X	965.334	1Hex 1HexNAc 2 NeuAc