



Published in final edited form as:

J Proteome Res. 2013 May 3; 12(5): 2295–2304. doi:10.1021/pr400212z.

Extensive *in vivo* human milk peptidomics reveals specific proteolysis yielding protective antimicrobial peptides

David C. Dallas^{a,b,†}, Andres Guerrero^{c,†}, Nora Khaldi^{d,e}, Patricia A. Castillo^f, William F. Martin^a, Jennifer T. Smilowitz^{a,b}, Charles L. Bevins^f, Daniela Barile^{a,b}, J. Bruce German^{a,b}, and Carlito B. Lebrilla^{b,c,1}

^aDepartment of Food Science, University of California at Davis, One Shields Avenue, Davis, CA, 95616, USA. ^bFoods for Health Institute, University of California at Davis, One Shields Avenue, Davis, CA, 95616, USA. ^cDepartment of Chemistry, University of California at Davis, One Shields Avenue, Davis, CA, 95616, USA. ^dSchool of Medicine and Medical Science, University College Dublin, Belfield, Ireland. ^eComplex and Adaptive Systems Laboratory, University College Dublin, Dublin 4, Republic of Ireland. ^fDepartment of Medical Microbiology and Immunology, School of Medicine, University of California, Davis, One Shields Avenue, Davis, CA, 95616, USA.

Abstract

Milk is traditionally considered an ideal source of the basic elemental nutrients required by infants. More detailed examination is revealing that milk represents a more functional ensemble of components with benefits to both infants and mothers. A comprehensive peptidomics method was developed and used to analyze human milk yielding an extensive array of protein products present in the fluid. Over 300 milk peptides were identified originating from major and many minor protein components of milk. As expected, the majority of peptides derived from β -casein, however no peptide fragments from the major milk proteins lactoferrin, α -lactalbumin and secretory immunoglobulin A were identified. Proteolysis in the mammary gland is selective—released peptides were drawn only from specific proteins and typically from only select parts of the parent sequence. A large number of the peptides showed significant sequence overlap with peptides with known antimicrobial or immunomodulatory functions. Antibacterial assays showed the milk peptide mixtures inhibited the growth of *Escherichia coli* and *Staphylococcus aureus*. The pre-digestion of milk proteins and the consequent release antimicrobial peptides may provide a selective advantage through evolution by protecting both the mother's mammary gland and her nursing offspring from infection.

Keywords

peptidomics; human milk; protease; functional peptide; plasmin

¹**Corresponding Author** To whom correspondence should be addressed: Carlito B. Lebrilla, cblebrilla@ucdavis.edu, phone: (530) 752-6364, fax: (530) 752-4759, Department of Chemistry, University of California at Davis, One Shields Avenue, Davis, CA, 95616, USA.

[†]These authors contributed equally to this work.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Supporting Information. Supporting information includes a three tables: 1) metadata for the mothers providing milk samples; 2) a list of all peptides identified in each sample; 3) a list of peptides found in the heat denatured and untreated sample. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Introduction

Mammalian lactation and its remarkable product, milk, evolved over more than 200 million years to nourish and protect the neonate¹. A large number of milk peptides produced by *in vitro* proteolysis have been found to be functional beyond their simple nutrient provision as amino acids². Activities of milk peptides include immunomodulation^{3,4}, opioid-like activity^{5,6}, antimicrobial action⁷⁻⁹ and probiotic action¹⁰⁻¹². These peptide fragments are not functional when constrained in the context of intact milk proteins¹³. Site-specific proteolysis releases these encrypted fragments. The best described example is the *in vitro* digestion of human lactoferrin by gastric pepsin that produces the peptide fragment lactoferricin that has potent and specific bactericidal properties¹⁴.

Most of these peptides are not naturally occurring—they were produced by *in vitro* digestion, some with the goal of recreating peptides that would be produced in digestion. In this study, we identify the peptides that are naturally occurring in human milk as the first step in understanding where and when milk's peptides can exert specific functions. For this purpose we used a novel, streamlined, high-throughput analytical approach optimized to explicitly capture and identify the complete set of peptides produced by human *in vivo* proteolytic digestion of breast milk. To this point, this study represents the most complete analysis of naturally occurring peptides in human milk.

The antimicrobial functionality of some of the peptides we discovered in our analyses highlights that the pre-digestion of human milk proteins we observe in the mammary gland is likely not of a random nature. The benefit of protecting both the infant and mother are likely major drivers behind the release of some of these peptides.

Materials and Methods

Chemicals and Sample set

Acetonitrile (ACN), formic acid (FA) and trifluoroacetic acid (TFA) were obtained from Thermo Fisher Scientific (Waltham, MA) and trichloroacetic acid (TCA) from EMD Millipore (Darmstadt, Germany). Insulin chain A from bovine pancreas was obtained from Sigma-Aldrich (St. Louis, MO).

Milk samples from five mothers who delivered at term were collected for this study. All milk samples were from day 28 of lactation. All donors were healthy and gave birth to healthy infants. None of the 5 mothers had clinical signs of mastitis on the sampling day. Metadata for the 5 mother-infant dyads are presented in Supplemental Table 1. Milk samples were taken from milk expressed by breast milk pumps, transferred into sterile plastic containers and immediately stored in home freezers. Manual expression typically takes 10–15 min during which milk samples were exposed to room temperature. Milk samples were transported on dry ice to the laboratory where they were stored at -80°C until the moment of the sample preparation. Milk samples were collected with IRB approval.

Sample Preparation

Milk fat fractionation of the sample was performed according to method described by Dallas et al.¹⁵. Briefly, 100 μL of each sample was centrifuged at $16,000 \times g$ for 10 min at 4°C and the skim milk infranate was removed from beneath the fat layer by pipette. The procedure was repeated until no fat was observed.

Proteins were removed by TCA precipitation according to the method of Ferranti et al.¹⁶. Briefly, 100 μL of 200 g/L TCA in nanopure water were added to 100 μL of skim milk. The samples were mixed using a vortex mixer, centrifuged at $3,000 \times g$ at 4°C for 10 min and the

supernatant was collected. The peptide-enriched supernatant was cleaned of contaminants, mainly oligosaccharides, through solid phase extraction (SPE) with 500 mg bed C18 columns (Supelco). The peptides were eluted from the column using 80% ACN, 0.1%TFA solution. Samples were dried and rehydrated for MS injection.

To ensure that peptides identified were not the result of continued proteolytic digestion during sample preparation, fresh milk was exposed to two treatments and analyzed. Fresh milk was provided by a healthy mother with no clinical signs of mastitis. Pumping took 10 min and was delivered to the lab on ice within 2 min thereafter. The first treatment was exactly as the above protocol—samples were immediately frozen and later thawed for sample preparation. The second sample was first boiled for 5 min to deactivate any protease activity. After boiling, this sample was frozen like the other sample and then subjected to the same sample preparation method.

Mass Spectrometry Analysis

Samples were rehydrated with 20 μ L of nanopure water prior to mass spectrometry analysis. Samples (2 μ L/injection) were analyzed on an Agilent (Santa Clara, CA) nano-LC-chip-Q-TOF MS/MS (Chip-Q-TOF) with an Agilent chip C18 column at a flow rate of 0.3 μ L/min. The gradient elution solvents were (A) 3% ACN/0.1% FA and (B) 90% ACN/0.1% FA. The gradient employed was ramped from 0–8% B from 0–5 min, 8–26.5% B from 5–24 min, 26.5–100% B from 24–48 min, followed by 100% B for 2 min and 100% A for 10 min (to re-equilibrate the column). The capillary pump was set to 3.5 μ L/min and 0% B throughout the analysis. Ion polarity was set to positive. The peak collection thresholds were set at 200 ion counts or 0.01% relative intensity for MS spectra and 5 ion counts or 0.01% relative intensity for MS/MS. Data were collected in centroid mode. The drying gas was 325 °C and flow rate was 5 L/min. The required chip voltage for consistent spray varied from 1850 to 1920 V. Automated precursor selection based on abundance was employed to select peaks for tandem fragmentation with an exclusion list consisting of all peptides identified in previous analyses in this study. The acquisition rate employed was .63 spectra/s for both MS and MS/MS modes. The isolation width for tandem analysis was 1.3 m/z. The collision energy was set by the formula (Slope)*(m/z)/100+Offset, with slope = 3.6 and offset = -4.8. Five tandem spectra were collected after each MS spectrum, with active exclusion after 2 MS/MS for 2 min. Precursor ions were only selected if they had at least 1000 ion counts or 0.01% of the relative intensity of the spectra. Mass calibration was performed during data acquisition based on an infused calibrant ion with a mass of 922.009789 Da.

Data Analysis

Agilent Mass Hunter Qualitative Analysis Software (Santa Clara, CA) was used to analyze the data obtained. Molecules identified in the spectral analysis were grouped into compounds by the Find by Molecular Feature algorithm, which groups together molecules across charge state and charge carrier. All tandem-MS from each data file were exported as Mascot Generic Files (.mgf) with a peptide isotope model and a maximum charge state of +9.

Peptide identification was accomplished using both the Mass Spectral-Generating Function Database (MS-GFDB) ¹⁷ (via a command-line interface) and X!Tandem ¹⁸ (using the downloadable graphical user interface). The human milk library used in both searches was constructed based on results from proteomic studies of human milk ¹⁹⁻²¹. A human milk protein library was employed because of the smaller number of proteins involved: because the search required non-specific cleavage, using the whole human protein library took too long. These were exported to FASTA file format. For MS-GFDB, peptides were accepted if p-values were less than or equal to 0.01 corresponding to a 99% confidence level. No p-

values exist in X!Tandem, so a closely related statistic, e-value, was used for the X!Tandem search. The e-value threshold selected was 0.01. In both programs, masses were allowed 20 ppm error. No complete (required) modifications were included but up to four potential modifications were allowed on each peptide. Potential modifications allowed included serine, threonine and tyrosine phosphorylation; methionine and tryptophan oxidation; asparagine and glutamine deamidation; and glutamine dehydration. A non-specific cleavage ($[X][X]$) (where 'X' is any amino acid) was used to search against the protein sequences. For MS-GFDB, the fragmentation method selected in the search was CID and the instrument selected was TOF. For X!Tandem, there was no option for fragmentation type and instrument selection. 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.

Creation of Exclusion List

All masses fragmented in the first round of MS/MS were incorporated into an exclusion list for the following round of mass spectrometry. This exclusion list was composed of mass-to-charge signals, charge state and retention times of the previously fragmented compounds. Ions on the exclusion list were ignored by the instrument and hence were not fragmented again. A +/-20 ppm error window was employed. The retention time window was set at +/-0.5 min. Each sample was run with MS/MS four times. Placing these peaks on the exclusion list allowed the instrument to fragment peaks of lesser abundance that co-eluted with these unidentified compounds. This approach allowed deeper exploration of the data, namely, identification of peaks at low abundance.

Library Search

A complete library of the masses and retention times of all unique peptides found from all MS/MS runs in all samples was compiled. This list was applied back to the MS-only data file for each sample with the Find by Molecular Feature compound-search function of Agilent Qualitative Analysis. This approach allowed extraction of peak volumes for each identified peptide. Find by Molecular Feature was performed with the following parameters: target data type: small molecules (chromatographic); peak height minimum: 1000 counts; retention time restriction: 0–35 min; protonated ion species; peptide isotope model; maximum charge: +7. Extracted compounds were matched to library peptides if they had masses within 40 ppm. Matches were removed if the difference in retention time between the compound and the library was outside ± 3 min. For the boiling versus not boiling experiment, a separate library was created only from the X!Tandem results from a single MS/MS run on each sample. This library was applied to the MS only runs of the boiling versus not boiling experiment.

Search for Known Functional Peptides

To uncover breast milk peptides that overlap with existing functional peptides in the literature, identified peptides were compared to sequences from four functional peptide databases: BIOPEP,²² PeptideDB,²³ CAMP,²⁴ and APD2.²⁵ We merged all four databases and parsed this dataset to remove duplicates. Because hormone peptides in these databases could be very large, the new database was restricted to hormonal peptides less than 60 amino acids in length.

Each breast milk peptide was searched against the database using protein-protein BLAST (BLASTP). For each query, a known functional peptide was retained if E-values were less than 0.5 and at least 50% of the query sequence was covered by the library sequence. This high E-value was chosen to counter-balance the effect of the small size of the milk peptides, which as an effect will have higher E-values. The high E-value threshold allowed for

discovery of overlapping sequences that would be missed with a smaller E-value threshold. The BLASTP output was parsed to remove false positives.

Antimicrobial assays

For the antimicrobial assays, peptides were obtained from the TCA precipitation method for peptide isolation. These peptides were tested for antimicrobial activity against *Escherichia coli* (*E. coli*), strain D31 and *Staphylococcus aureus* (*S. aureus*) as previously described²⁶. In summary, the bacteria were grown aerobically, pelleted by brief centrifugation, washed and resuspended with 10 mM sodium phosphate buffer, pH 7.4. The experiments were performed in triplicate, using either 5×10^4 , 5×10^5 , or 5×10^6 CFU of bacteria. The bacteria were inoculated into the underlay medium, which contained 0.03% Trypticase soy broth (TSB), 1% wt/vol low electroendosmosis (EEO) agarose (Sigma), 0.02% vol/vol Tween-20 dissolved in 10 mM phosphate buffer, pH 7.4. The media was poured into a polystyrene culture dish and allowed to solidify. Once the agarose solidified, 3 mm diameter wells were punched in the plate. On the *E. coli* plate, holes B2, B4, C3 and C5 were loaded with 4 μ L of the peptide mixture at different concentrations (6 μ g/ μ L, 0.6 μ g/ μ L, 0.06 μ g/ μ L and 0.006 μ g/ μ L, respectively), well F1 was loaded with 1 μ g/ μ L maganin—an antimicrobial peptide—as the positive control, well F4 was loaded with 1 μ g/ μ L human defensin-6 as the negative control, and well D6 was loaded with nanopure water as another negative control. For the *S. aureus* assay, wells C5 and D2 were loaded with 4 μ L of the human milk peptide mixture at different concentrations (8 μ g/ μ L and 4 μ g/ μ L, respectively), F2 was loaded with 1 μ g/ μ L maganin, well F4 was loaded with 1 μ g/ μ L human defensin-6, and well E3 was loaded with nanopure water. Then, the plates were incubated for 3 h at 37°C. The plate was then overlaid with 6% TSB, 1% (w/v) low EEO agarose dissolved in 10 mM phosphate buffer, pH 7.4. After solidifying, the plates were incubated overnight at 37°C. Expansion of areas with no bacterial growth around the well demonstrates inhibition of bacterial growth from the compound in that well.

Results

Peptide Identification and Protein Source

Extensive peptidomic analysis of milk yielded a large number of naturally occurring milk peptides. Figure 1 shows an image of several overlapping extracted compound chromatograms (ECC) of peptides enriched from the pooled breast milk. After sequence identification, peptides were matched to their extracted compound chromatograms and labeled in Figure 1. Peptides were identified with the software X!Tandem¹⁸ and MS-GFDB¹⁷ with a database search of the MS/MS spectra using tandem MS spectra such as that shown in Figure 2. The analysis is performed using an accurate determination of the intact mass of the peptides and the sequence as determined by MS/MS. This novel method allowed the identification of 328 unique peptide sequences from human milk. The sequences for all peptides identified are shown in Supplemental Table 2.

Three hundred twenty-eight peptides were identified, with the majority (59%) derived from a single protein, β -casein (Figure 3), which is an abundant protein in milk. However, the representation of other proteins does not appear to follow the natural abundances strictly. For example, other highly abundant proteins such as α -lactalbumin (the most abundant) and lactoferrin are not represented. Other contributors to peptide fragments included polymeric immunoglobulin receptor, butyrophilin, α_{s1} -casein, osteopontin and mucin-1. Overall, 37 proteins with at least one unique peptide (meaning unique sequence, protein of origin and number of phosphorylations) at p-value (MS-GFDB) or e-value (X!Tandem) $\leq .01$ were identified. Peptides ranged in length from 10 to 44 amino acids with the average peptide

length around 22 amino acids. The length corresponded to peptide masses between 1171 and 4486 Daltons, with an average of approximately 2383 Daltons.

Twenty-three percent of peptides identified were phosphorylated at serine, threonine or tyrosine (76 unique peptides). We provide the number of phosphorylations within the protein because tandem MS could not distinguish many of the sites.

Interestingly, the unique peptides identified were not randomly distributed across the overall protein sequence (Figure 4). For butyrophilin, osteopontin, mucin-1, perilipin-2 and polymeric immunoglobulin receptor, the peptides originated from a specific point on the protein. β -casein is unique in this regard as almost the entire sequence was represented, suggesting no specific points of enzymatic digestion.

In order to ensure the identified peptides were not the result of continued protease activity during sample preparation, we compared a sample that was denatured by heat (100°C for 5 min) directly after expression to one that was not. Both samples were generated from fresh milk (less than 15 min after expression at the time of sample preparation initiation). Boiling for 5 min should eliminate most or all of the protease activity in milk. After boiling, both samples were immediately frozen and later extracted by the same method as the previous samples and analyzed for peptides. A new peptide library based on the results of only one round of MS/MS for each sample was created and applied to the MS-only runs of each sample. The results indicate that no perceivable effects were observed upon heating, as ninety-six percent of all unique peptides identified in this sample set were found to overlap (Supplemental Table 3). The results suggest that little to no proteolytic action occurs during our normal sample preparation procedure.

Putative Peptide Functions

The peptides were compared against a database of functional peptides from BIOPEP²², PeptideDB²³, CAMP²⁴, and APD2²⁵. Of the 328 peptides found, 59 shared homology with functional peptides. The high sequence overlap between the endogenous milk peptides with those in the database suggests the likely function of these peptides. Forty-one of these peptides matched known antibacterial sequences and eighteen peptides matched immunomodulatory sequences (sequences shown in Table 1). These peptides, representing about 18% of the total, are primarily from β -casein and κ -casein (Table 1) because they have been the subjects of earlier significant interests. The remaining 82% of the peptides have, as of yet, unknown function. We plan to investigate these peptides in the future for their functions.

Antimicrobial Functions

The antimicrobial activity of the milk peptides was tested in a radial diffusion assay against a Gram-negative (*Escherichia coli*, *E. coli*) and a Gram-positive (*Staphylococcus aureus*, *S. aureus*) bacterial species. Triplicate assays demonstrated that the growth of *E. coli* was inhibited by 6 $\mu\text{g}/\mu\text{L}$ of milk peptides as depicted by the zone of clearance in well B2 of Figure 5. Lower concentrations of milk peptides had no detectable inhibitory effect on the growth of *E. coli*. The growth of *S. aureus* was also inhibited by the ensemble of milk peptides (Figure 6) (demonstrated in triplicate). The growth of *S. aureus* was inhibited at 8 $\mu\text{g}/\mu\text{L}$, as seen by the zone of clearance in well C5 of Figure 6. Similarly to the *E. coli* assays, with lower concentrations of the milk peptide ensemble, our assay could not detect inhibition of *S. aureus* growth.

Discussion

This study used a novel, comprehensive and highly sensitive approach to identify nearly all naturally occurring peptides from human milk. By employing an exclusion list on the tandem MS analysis and four injections of each sample, the number of unique peptides identified increased by 3–5-fold for each sample compared to a single tandem identification run. This strategy probed deeper into peptide data, and could be applied to many other molecule types. Similar exclusion list strategies employed for proteomics with offline-LC MALDI MS/MS^{27, 28} and ESI-MS/MS²⁹⁻³¹ increased the number of proteins identified, however this method seems to be most effective in peptidomics analyses.

Over 300 unique naturally occurring peptides were identified at 99% confidence using a single analytical platform from over 37 proteins. Previous studies on this topic have focused on the hydrolytic products of specific milk proteins and lack a complete description of the complete protein-released peptidome. Ferranti et al.¹⁶ used three different mass spectrometers and Edman sequencing to determine the sequence of around 100 naturally occurring peptides in human milk from β -, α_{s1} - and κ -casein. Armaforte et al.³² confirmed the presence of low-molecular weight fragments (but not the exact sequences) of β - and α_{s1} -casein in human milk by 2D-SDS-PAGE followed by trypsin digestion of gel spots and mass spectrometry³². Christensen et al.³³ found seven naturally occurring peptide fragments of osteopontin, another common milk protein, in intact human milk via immunoaffinity extraction and mass spectrometry.

The unique peptides identified in this study derived from specific locations on the parent protein. That these peptides are released from specific foci of each protein suggests that protein cleavages in the mammary gland are not random events. The exception is β -casein, which yielded peptides from throughout the entire length of the protein. Furthermore, the cleavage sites on the protein commonly matched the site-specificity of plasmin, an *N*- and *O*-linked glycosylated^{34, 35} protease known to exist and be active in human milk³⁶⁻³⁹. Plasminogen, the inactive precursor (zymogen) form of plasmin, is known to exist in milk as a free protein and in association with the casein micelles^{40, 41}. Plasminogen must be cleaved by plasmin activators to become active as plasmin. Two types of plasmin activators are present in human milk: urokinase-type plasmin activators and tissue-type plasmin activators⁴¹. Plasmin inhibitors, which block plasmin activators from converting plasminogen to plasmin, are also present in human milk, including type-I plasminogen activator inhibitor⁴¹. The presence of β -casein derived peptides as the major degradation products in milk despite not being the most abundant protein suggests that there is active casein-bound plasmin that degrade proteins that associate with the micelle structure. This notion is supported by the fact that most of the tissue-type plasminogen activator is associated with casein micelles, as opposed to the whey fraction, of human milk⁴¹. Moreover, casein micelles have been shown to effectively increase the activity of plasmin compared to plasminogen and plasmin activator⁴¹⁻⁴³. Casein's ability to increase plasminogen activation in the micelle was hypothesized to be due to either conformational changes in the micelle associated plasminogen or proximity between the plasminogen and the plasmin activator⁴¹. We hypothesize further that the major reason whey proteins such as α -lactalbumin, secretory immunoglobulin A and lactoferrin do not yield digested peptides in milk is because they do not associate with the micelles and because free plasminogen is blocked from activation by the plasmin inhibitors associated with the whey.

Structural differences in the milk proteins may also be responsible for the differences in appearance of degradation products. More tightly packed protein structures likely have more resistance to protease activity than looser tertiary structures. α -lactalbumin and lactoferrin, which were not found as released peptide fragments, have more tightly packed globular

structures^{44, 45}. β -casein molecules, however, have an elongated, flexible and rheomorphic (non-fixed) structure^{46, 47} that may be more amenable to proteolytic degradation. The structural differences between these proteins may be responsible for their abundance as fragments in the milk.

There is a possibility that the peptides found are degradation products due to prolonged exposure of the proteins in milk to milk's enzymes after expression or due to sample treatment. However, denaturation of the milk proteins appears to have no effect on the peptide product. Furthermore, the procedures used are standard to proteomics analysis and have been optimized to maintain intact peptides. There is a possibility that the peptides are produced intact in the mammary gland via by alternative mRNA splicing and differential mRNA editing. An example of production of different protein products from the same primary transcript is apolipoprotein B48 (apoB48), which is produced from the same primary transcript as apoB100, but with a single nucleotide inserted in the middle of the mRNA through mRNA editing creating a stop codon, leading to translation of a truncated protein⁴⁸. mRNA data has been collected for these mothers and we will assess these possibilities with that data in a future paper.

Peptides isolated from human milk inhibited the growth of *E. coli* and *S. aureus* in triplicate assays. Milk proteases may be specifically releasing functional peptides from milk proteins to benefit maternal and infant health by preventing bacterial infection and guiding the immune system. However, the presence of these peptides in the milk as opposed to their release in the gut of the infant suggests they may exert a greater benefit to the mother. The mother may produce these peptides to aid in the prevention of bacterially induced mastitis. Mastitis, a painful inflammation of the breast that can result in blocked milk ducts, abscesses and septicemia if left untreated or undertreated, is a common problem in breast-feeding mothers (estimates range from 3 to 33% of all mothers⁴⁹⁻⁵⁴). The bacterium most commonly associated with mastitis is *S. aureus*^{50, 55-57}, with *E. coli* association being less frequent⁵⁴. The peptide mixtures were inhibitory at 6 and 8 $\mu\text{g}/\mu\text{L}$ for *E. coli* and *S. aureus*, respectively, and this concentration is above that of milk (the total peptide yield was roughly .3 $\mu\text{g}/\mu\text{L}$ of milk). Because the specific peptides exhibiting antimicrobial properties are not known, their efficacy cannot yet be determined. Nonetheless, the large diversity of peptides present may function against antimicrobial resistance.

In this study, the gross antimicrobial effect of the peptide mixture was considered, however current studies involving the synthesis of key peptides will determine those responsible for the antimicrobial properties. In addition, whether these peptides are beneficial to the infant and survive intact in the infant digestive tract will also be the subject of future investigations. The presence of such antimicrobial peptides in the digestive tract could be important in the prevention of infant gastrointestinal infections, including necrotizing enterocolitis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Cora J. Dillard for her help in editing this manuscript.

Funding Sources

United States Department of Agriculture, National Institute for Food and Agriculture post-doctoral fellowship; National Science Foundation Graduate Research Fellowship Program; National Institute of Health Training Program in Biomolecular Technology (2-T3-GM08799).

ABBREVIATIONS

ECC	Extracted compound chromatogram
E. coli	Escherichia coli
S. aureus	Staphylococcus aureus
apoB48	apolipoprotein B48
ACN	Acetonitrile
FA	formic acid
TFA	trifluoroacetic acid
TCA	trichloroacetic acid
SPE	solid phase extraction
Chip-Q-TOF	nano-LC-chip-Q-TOF MS/MS
TSB	trypticase soy broth
EEO	electroendosmosis

REFERENCES

1. Oftedal OT. The mammary gland and its origin during synapsid evolution. *J. Mammary Gland Biol. Neoplasia*. 2002; 7:225–252. [PubMed: 12751889]
2. Dallas D, Underwood M, Zivkovic A, German J. Digestion of protein in premature and term infants. *J Nutr Disorders Ther*. 2012; 2:2161–0509.1000112.
3. Migliore-Samour D, Floch F, Jollès P. Biologically-active casein peptides implicated in immunomodulation. *J. Dairy Res*. 1989; 56:357–362. [PubMed: 2668358]
4. Jørgensen ALW, Juul - Madsen HR, Stagsted J. Colostrum and bioactive, colostrum peptides differentially modulate the innate immune response of intestinal epithelial cells. *J. Pept. Sci*. 2010; 16:21–30. [PubMed: 19904766]
5. Kampa M, Loukas S, Hatzoglou A, Martin P, Martin P, Castanas E. Identification of a novel opioid peptide (Tyr-Val-Pro-Phe-Pro) derived from human alpha S1 casein (alpha S1-casomorphin, and alpha S1-casomorphin amide). *Biochem. J*. 1996; 319:903–908. [PubMed: 8920997]
6. Brantl V. Novel opioid-peptides derived from human beta-casein - human beta-casomorphins. *Eur. J. Pharmacol*. 1984; 106:213–214. [PubMed: 6529969]
7. Liepke C, Zucht HD, Forssmann WG, Standker L. Purification of novel peptide antibiotics from human milk. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci*. 2001; 752:369–377.
8. Aniansson G, Andersson B, Lindstedt R, Svanborg C. Antiadhesive activity of human casein against streptococcus-pneumoniae and haemophilus-influenzae. *Microb. Pathog*. 1990; 8:315–323. [PubMed: 2215182]
9. Stromqvist M, Falk P, Bergstrom S, Hansson L, Lönnerdal B, Normark S, Hernell O. Human-milk k-casein and inhibition of *Helicobacter pylori* adhesion to human gastric mucosa. *J. Pediatr. Gastroenterol. Nutr*. 1995; 21:288–296. [PubMed: 8523212]
10. Liepke C, Adermann K, Raida M, Mägert HJ, Forssmann WG, Zucht HD. Human milk provides peptides highly stimulating the growth of bifidobacteria. *Eur. J. Biochem*. 2002; 269:712–718. [PubMed: 11856332]
11. Bezkorovainy A, Grohlich D, Nichols J. Isolation of a glycopolypeptide fraction with *Lactobacillus bifidus* subspecies *pennsylvanicus* growth-promoting activity from whole human milk casein. *Am. J. Clin. Nutr*. 1979; 32:1428–1432. [PubMed: 572132]
12. Azuma N, Yamauchi K, Mitsuoka T. Bifidus growth-promoting activity of a glycomacropptide derived from human kappa-casein. *Agric. Biol. Chem*. 1984; 48:159–2162.
13. Schanbacher F, Talhouk R, Murray F. Biology and origin of bioactive peptides in milk. *Livestock Production Science*. 1997; 50:105–123.

14. Bellamy W, Takase M, Yamauchi K, Wakabayashi H, Kawase K, Tomita M. Identification of the bactericidal domain of lactoferrin. *Biochim. Biophys. Acta.* 1992; 1121:130–136. [PubMed: 1599934]
15. Dallas DC, Martin WF, Strum JS, Zivkovic AM, Smilowitz JT, Underwood MA, Affolter M, Lebrilla CB, German JB. *N*-linked glycan profiling of mature human milk by high performance microfluidic chip liquid chromatography time-of-flight tandem mass spectrometry. *J. Agric. Food Chem.* 2011; 59:4255–4263. [PubMed: 21384928]
16. Ferranti P, Traisci MV, Picariello G, Nasi A, Boschi V, Siervo M, Falconi C, Chianese L, Addeo F. 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]
17. Kim S, Mischerikow N, Bandeira N, Navarro JD, Wich L, Mohammed S, Heck AJR, Pevzner PA. The generating function of CID, ETD, and CID/ETD pairs of tandem mass spectra: applications to database search. *Mol. Cell. Proteomics.* 2010; 9:2840–2852. [PubMed: 20829449]
18. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics.* 2004; 20:1466–1467. [PubMed: 14976030]
19. Molinari CE, Casadio YS, Hartmann BT, Livk A, Bringans S, Arthur PG, Hartmann PE. Proteome mapping of human skim milk proteins in term and preterm milk. *J. Proteome Res.* 2012
20. Mange A, Bellet V, Tuailon E, Van de Perre P, Solassol J. Comprehensive proteomic analysis of the human milk proteome: Contribution of protein fractionation. *J. Chromatogr. B.* 2008; 876:252–256.
21. Liao Y, Alvarado R, Phinney B, Lönnerdal B. Proteomic characterization of human milk fat globule membrane proteins during a 12 month lactation period. *J. Proteome Res.* 2011
22. Dziuba J, Minkiewicz P, Nalecz D, Iwaniak A. Database of biologically active peptide sequences. *Food/Nahrung.* 1999; 43:190–195.
23. Liu F, Baggerman G, Schoofs L, Wets G. The construction of a bioactive peptide database in Metazoa. *J. Proteome Res.* 2008; 7:4119–4131. [PubMed: 18707150]
24. Thomas S, Karnik S, Barai RS, Jayaraman V, Idicula-Thomas S. CAMP: a useful resource for research on antimicrobial peptides. *Nucleic Acids Res.* 2010; 38:D774–D780. [PubMed: 19923233]
25. Wang G, Li X, Wang Z. APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucleic Acids Res.* 2009; 37:D933–D937. [PubMed: 18957441]
26. Lehrer RI, Rosenman M, Harwig SSSL, Jackson R, Eisenhauer P. Ultrasensitive assays for endogenous antimicrobial polypeptides. *J. Immunol. Methods.* 1991; 137:167–173. [PubMed: 1901580]
27. Chen H, Rejtar T, Andreev V, Moskovets E, Karger BL. Enhanced characterization of complex proteomic samples using LC-MALDI MS/MS: exclusion of redundant peptides from MS/MS analysis in replicate runs. *Anal. Chem.* 2005; 77:7816–7825. [PubMed: 16316193]
28. Zerck A, Nordhoff E, Resemann A, Mirgorodskaya E, Suckau D, Reinert K, Lehrach H, Gobom J. An iterative strategy for precursor ion selection for LC-MS/MS based shotgun proteomics. *J. Proteome Res.* 2009; 8:3239–3251. [PubMed: 19402737]
29. Wang N, Li L. Exploring the precursor ion exclusion feature of liquid chromatography–electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Anal. Chem.* 2008; 80:4696–4710. [PubMed: 18479145]
30. Muntel J, Hecker M, Becher D. An exclusion list based label - free proteome quantification approach using an LTQ Orbitrap. *Rapid Commun. Mass Spectrom.* 2012; 26:701–709. [PubMed: 22328225]
31. Voisin SN, Krakovska O, Matta A, DeSouza LV, Romaschin AD, Colgan TJ, Siu KWM. Identification of novel molecular targets for endometrial cancer using a drill-down LC-MS/MS approach with iTRAQ. *PloS one.* 2011; 6:e16352. [PubMed: 21305022]
32. Armaforte E, Curran E, Huppertz T, Ryan CA, Caboni MF, O'Connor PM, Ross RP, Hirtz C, Sommerer N, Chevalier F. Proteins and proteolysis in pre-term and term human milk and possible implications for infant formulae. *Int. Dairy J.* 2010; 20:715–723.

33. Christensen B, Schack L, Klänning E, Sørensen ES. Osteopontin is cleaved at multiple sites close to its integrin-binding motifs in milk and is a novel substrate for plasmin and cathepsin D. *J. Biol. Chem.* 2010; 285:7929–7937. [PubMed: 20071328]
34. Picariello G, Ferranti P, Mamone G, Roepstorff P, Addeo F. Identification of N-linked glycoproteins in human milk by hydrophilic interaction liquid chromatography and mass spectrometry. *Proteomics.* 2008; 8:3833–3847. [PubMed: 18780401]
35. Pirie-Shepherd SR, Stevens RD, Andon NL, Enghild JJ, Pizzo SV. Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of human plasminogen 2. *J. Biol. Chem.* 1997; 272:7408–7411. [PubMed: 9054441]
36. Warner RC, Polis E. On the presence of a proteolytic enzyme in casein. *J Am Chem Soc.* 1945; 67:529–532.
37. Okamoto U, Horie N, Nagamatsu Y, Yamamoto J. Plasminogen-activator in human early milk: its partial purification and characterization. *Thromb Haemostasis.* 1981; 45:121. [PubMed: 7196095]
38. Korycha-Dahl M, Dumas BR, Chene N, Martal J. Plasmin activity in milk. *J. Dairy Sci.* 1983; 66:704–711.
39. Astrup, T.; Sterndorff, I. *A Fibrinolytic System in Human Milk.* Royal Society of Medicine; 1953. p. 605-608.1953
40. Heegaard CW, Rasmussen LK, Andreasen PA. The plasminogen activation system in bovine milk: Differential localization of tissue-type plasminogen activator and urokinase in milk fractions is caused by binding to casein and urokinase receptor. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research.* 1994; 1222:45–55.
41. Heegaard CW, Larsen LB, Rasmussen LK, Højberg KE, Petersen TE, Andreasen PA. Plasminogen activation system in human milk. *J. Pediatr. Gastroenterol. Nutr.* 1997; 25:159–166. [PubMed: 9252902]
42. Politis I, White JH, Zavizion B, Goldberg JJ, Guo MR, Kindstedt P. Effect of Individual Caseins on Plasminogen Activation by Bovine Urokinase-Type and Tissue-Type Plasminogen Activators. *J. Dairy Sci.* 1995; 78:484–490. [PubMed: 7782505]
43. Markus G, Hitt S, Harvey S, Tritsch G. Casein, a powerful enhancer of the rate of plasminogen activation. *Fibrinolysis.* 1993; 7:229–236.
44. Permyakov EA, Berliner LJ. α -Lactalbumin: structure and function. *FEBS Lett.* 2000; 473:269–274. [PubMed: 10818224]
45. Legrand D, Pierce A, Ellass E, Carpentier M, Mariller C, Mazurier J. Lactoferrin structure and functions. *Adv. Exp. Med. Biol.* 2008; 606:163–94. [PubMed: 18183929]
46. Livney YD, Schwan AL, Dalgleish DG. A Study of β -Casein Tertiary Structure by Intramolecular Crosslinking and Mass Spectrometry. *J. Dairy Sci.* 2004; 87:3638–3647. [PubMed: 15483147]
47. Swaisgood, HE. *Chemistry of the caseins.* In: P.F. Fox, PLHM., editor. *Advanced Dairy Chemistry.* 3rd ed.. Kluwer Academic/ Plenum Publishers; New York, NY: 2003. p. 139-202.
48. Powell LM, Wallis SC, Pease RJ, Edwards YH, Knott TJ, Scott J. A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. *Cell.* 1987; 50:831. [PubMed: 3621347]
49. Riordan JM, Nichols FH. A Descriptive Study of Lactation Mastitis in Long-Term Breastfeeding Women. *J. Hum. Lact.* 1990; 6:53–58. [PubMed: 2346600]
50. Marshall BR, Hepper JK, Zirbel CC. Sporadic puerperal mastitis. *JAMA.* 1975; 233:1377–1379. [PubMed: 1174212]
51. Jonsson S, Pulkkinen M. Mastitis today: incidence, prevention and treatment. 1994; 208:84–87.
52. Kaufmann R, Foxman B. Mastitis among lactating women: occurrence and risk factors. *Soc. Sci. Med.* 1991; 33:701–705. [PubMed: 1957190]
53. Fetherston C. Risk factors for lactation mastitis. *J. Hum. Lact.* 1998; 14:101–109. [PubMed: 9775842]
54. Semba RD, Neville MC. Breast-feeding, Mastitis, and HIV Transmission: Nutritional Implications. *Nutr. Rev.* 1999; 57:146–153. [PubMed: 10391017]

55. Thomsen A, Espersen T, Maigaard S. Course and treatment of milk stasis, noninfectious inflammation of the breast, and infectious mastitis in nursing women. *Am. J. Obstet. Gynecol.* 1984; 149:492–495. [PubMed: 6742017]
56. Thomsen A, Hansen K, Møller B. Leukocyte counts and microbiologic cultivation in the diagnosis of puerperal mastitis. *Am. J. Obstet. Gynecol.* 1983; 146:938–941. [PubMed: 6688325]
57. Matheson I, Aursnes I, Horgen M, Aabø Ø, Melby K. Bacteriological Findings And Clinical Symptoms. *Acta Obstet. Gynecol. Scand.* 1988; 67:723–726. [PubMed: 3250184]
58. Azuma N, Nagaune S, Ishino Y, Mori H, Kaminogawa S, Yamauchi K. DNA-synthesis stimulating peptides from human β -casein. *Agric. Biol. Chem.* 1989; 53:2631–2634.
59. Hayes M, Stanton C, Fitzgerald GF, Ross RP. Putting microbes to work: Dairy fermentation, cell factories and bioactive peptides. Part II: Bioactive peptide functions. *Biotechnology journal.* 2007; 2:435–449. [PubMed: 17407211]

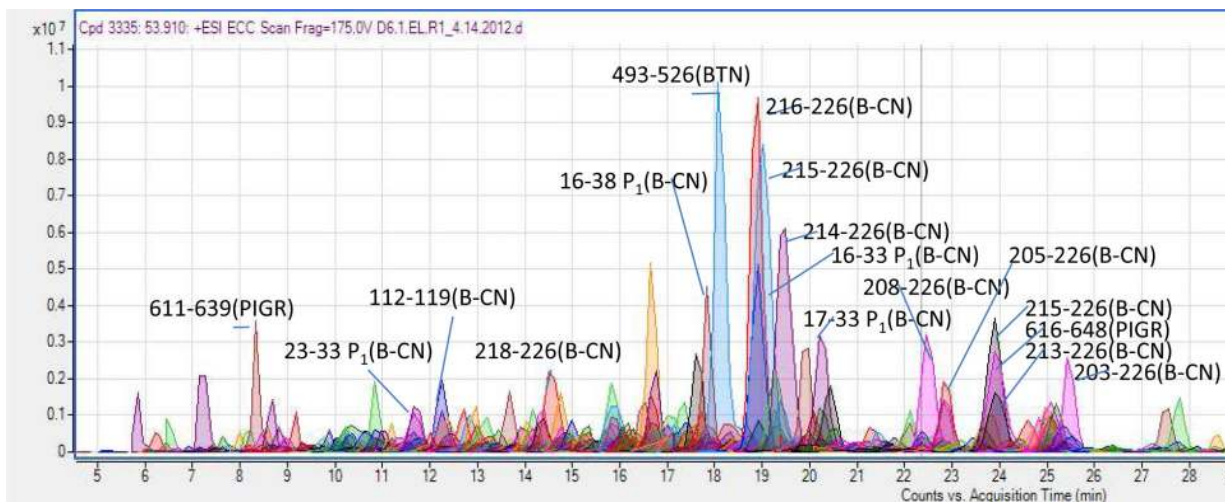


Figure 1.
 Example extracted compound chromatograms from identified peptides. The numbers indicate the position in the protein sequence the peptides were derived from. P₁ indicates the peptide has 1 phosphorylation. Polyimmunoglobulin receptor, PIGR; β -casein, B-CN; butyrophilin, BTN.

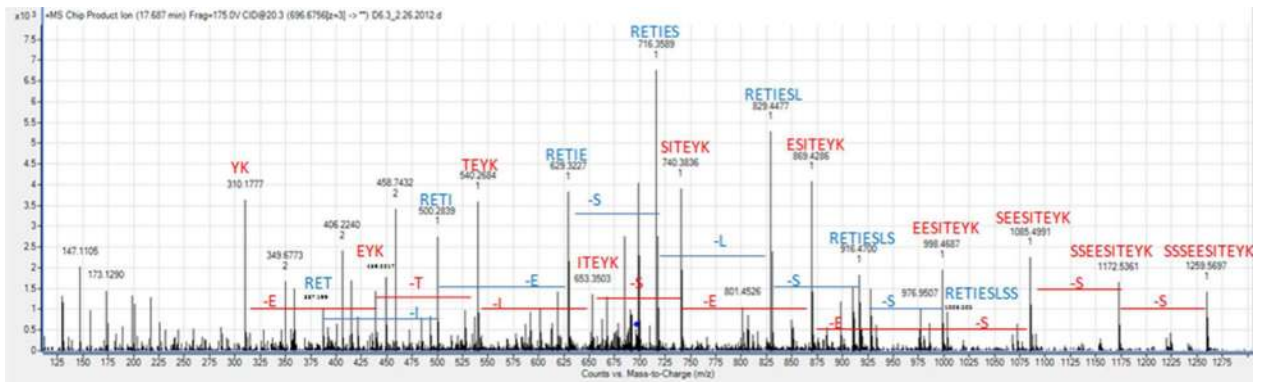


Figure 2.
Example tandem mass spectrum of peptide RETIESLSSEESITEYK from B-CN identified by both X!Tandem and MS-GFDB.

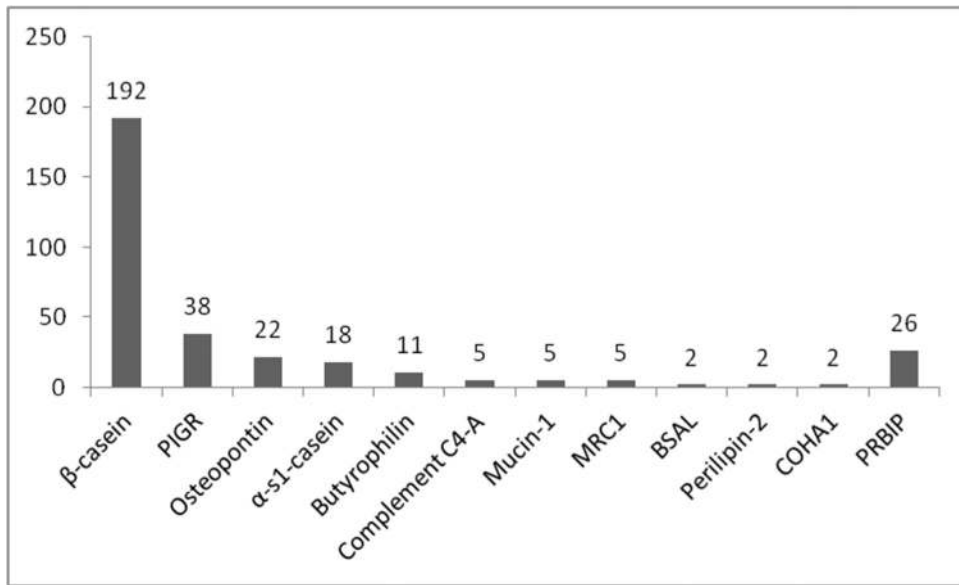


Figure 3. Number of unique peptides identified by protein of origin. PIGR: polymeric immunoglobulin receptor; BSAL: Bile salt-activated lipase; COHA1: collagen α -1(XVII) chain; MRC1: Macrophage mannose receptor 1; PRBIP: Proteins represented by one peptide.

Butyrophilin1.....DG₈₀REQEAEQMPE₉₀YRGRAT..........D₄₉₀LSKEIPLSPM₅₀₀GEDSAPRDAD₅₁₀TLHSKLIPTQ₅₂₀PSQGAP**Osteopontin**1.....RSK₁₇₀SKKFRRPDIQ₁₈₀YDPATDEDIT₁₉₀SHMESEELNG₂₀₀AYKAIPVAQD₂₁₀LNAPSDWDSR₂₂₀GKDSYETSQ₂₃₀DDQSAETHSH₂₄₀KQSRLY.......FRISHELDSAS₃₁₀SEVN**Mucin-1**1.....STDRSPYE₁₂₃₀KVSAGNGGSS₁₂₄₀LSYTNPAVAA₁₂₅₀TSANL**Polymeric Immunoglobulin Receptor**1.....KADA₅₈₀APDEKVLDSG₅₉₀FREIENKAIQ₆₀₀DPRLFAEKA₆₁₀VADTRDQADG₆₂₀SRASVDSGSS₆₃₀EEQGGSSRAL₆₄₀VSTLVPLGLV**Figure 4.**

Unique peptides found from all 5 mothers aligned against parent protein sequence for butyrophilin, osteopontin, mucin-1 and polymeric immunoglobulin receptor.

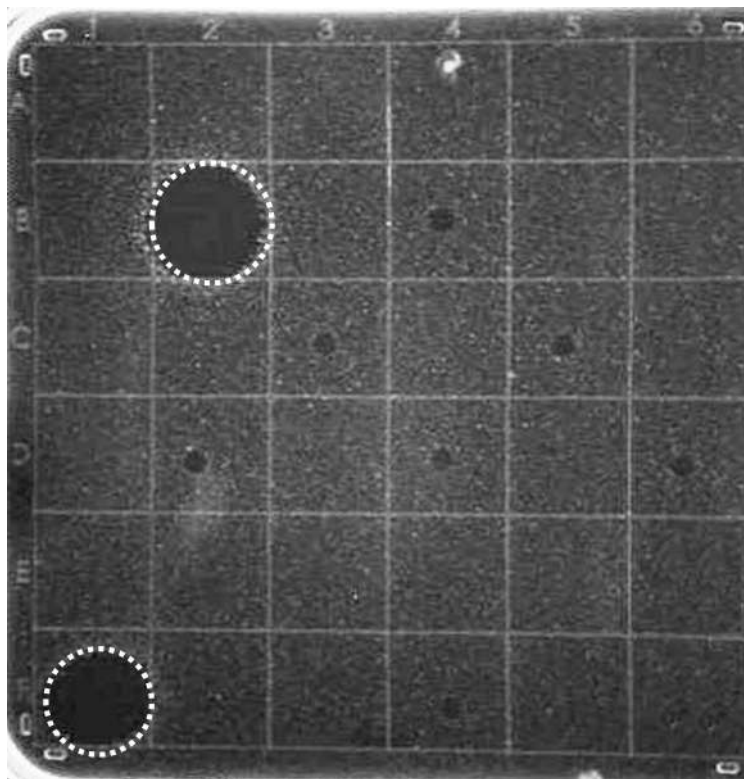


Figure 5. Antibacterial activity of human milk peptides incubated with 10^5 *E. coli*. Wells were loaded as follows: B2, 6 $\mu\text{g}/\mu\text{L}$ milk peptides; B4, 0.6 $\mu\text{g}/\mu\text{L}$ milk peptides; C3, 0.06 $\mu\text{g}/\mu\text{L}$ milk peptides; C5, 0.006 $\mu\text{g}/\mu\text{L}$ milk peptides; F1, 1 $\mu\text{g}/\mu\text{L}$ maganin; F4, 1 $\mu\text{g}/\mu\text{L}$ human defensin-6; D6, nanopure water.

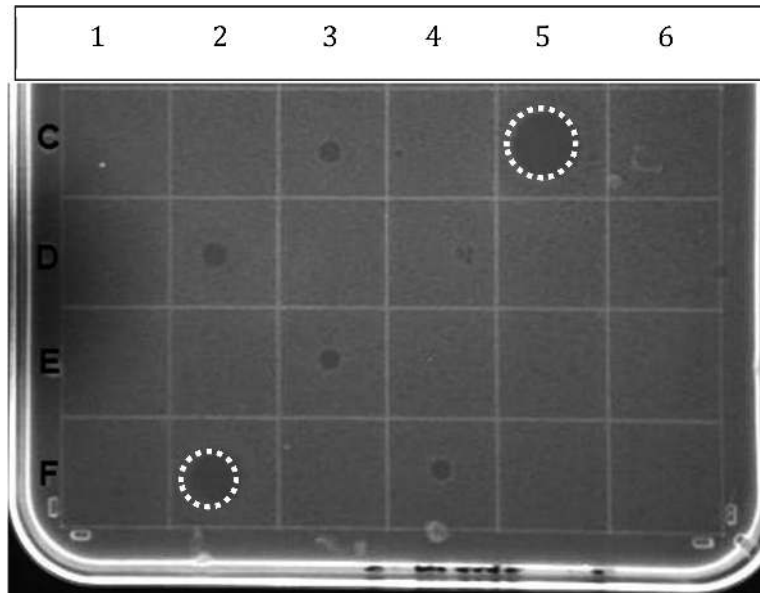


Figure 6. Antibacterial activity of human milk peptides incubated with 10^6 *S. aureus*. Wells were loaded as follows: C5, 8 $\mu\text{g}/\mu\text{L}$ milk peptides; D2, 4 $\mu\text{g}/\mu\text{L}$ milk peptides; F2, 1 $\mu\text{g}/\mu\text{L}$ maganin; F4, 1 $\mu\text{g}/\mu\text{L}$ human defensin-6; E3, nanopure water.

Table 1

Identified peptides containing over 50% of known functional peptide sequences. The overlap between breast milk peptides and the literature peptide is indicated in bolded amino acids. An amino acid mismatch between the literature peptide and our set of peptides is indicated by a red amino acid. Insertion of an amino acid is indicated in green.

Identified peptides	Found peptide protein of origin	Overlapping literature functional peptide	Known activity	Known functional peptide origin	Ref
FDPQIP KLTDLE	human β -CN	QPTIPFFDPQIPK	Immunomodulating	human β -CN (105-117)	58
GRVMPV L KSPTIPFFDP					
GRVMPV L KSPTIPFFDPQIP					
GRVMPV L KSPTIPFFDPQIPK					
GRVMPV L KSPTIPFFDPQIPKLD					
GRVMPV L KSPTIPFFDPQIPKLDLEN					
MPV L KSPTIPFFDPQIP					
MPV L KSPTIPFFDPQIPK					
MPV L KSPTIPFFDPQIPKLDLEN					
PTIPFFDPQIPK					
PTIPFFDPQIPK L					
PTIPFFDPQIPK LD					
PV L KSPTIPFFDPQIPK					
SPTIPFFDPQIPK					
SPTIPFFDPQIPK L					
VMPV L KSPTIPFFDPQIPK					
VMPV L KSPTIPFFDPQIPKLD					
VMPV L KSPTIPFFDPQIPKLDLEN					
A VPVQALLLN Q ELLN P THQIY P VT Q PLAPV H NPISV	human β -CN	QELLN P THQY P VT Q PLAPV H NPISV	Antibacterial	human β -CN (184-210)	59
E LLN P THQIY P VT Q PLAPV H NPIS					
E LLN P THQIY P VT Q PLAPV H NPISV					
H QIY P VT Q PLAPV H NPISV					
L LLN P THQIY P VT Q PLAPV H NPIS					

Identified peptides	Found peptide protein of origin	Overlapping literature functional peptide	Known activity	Known functional peptide origin	Ref
LLLNPTHQIYPTQPLAPVHNPIV					
LLLNQEELLNPTHQIYPTQPLAPVHNPIV					
LLLNQEELLNPTHQIYPTQPLAPVHNPIV					
LLNPTHQIYPTQPLAPVHNPIV					
LLNPTHQIYPTQPLAPVHNPIV					
LLNQELLNPTHQIYPTQPLAPVHNPIV					
LLNQELLNPTHQ					
LLNQELLNPTHQIYPTQPLAPVHNPIV					
LNPTHQIYPTQPLAPVHNPIV					
LNPTHQIYPTQPLAPVHNPIV					
LNQEELLNPTHQIYPTQPLAPVHNPIV					
LNQEELLNPTHQIYPTQPLAPVHNPIV					
NPTHQIYPTQPLAPVHNPIV					
NPTHQIYPTQPLAPVHNPIV					
NQEELLNPTHQIYPT					
NQEELLNPTHQIYPTQPLAPVH					
NQEELLNPTHQIYPTQPLAPVHNPI					
NQEELLNPTHQIYPTQPLAPVHNPI					
NQEELLNPTHQIYPTQPLAPVHNPIV					
NQEELLNPTHQIYPTQPLAPVHNPIV					
PTHQIYPTQPLAPVHNPIV					
PVTQPLAPVHNPIV					
QEELLNPTHQIYPTQPLAPVHNPIV					
QEELLNPTHQIYPTQPLAPVHNPIV					
QIYPTQPLAPVHNPIV					
QPLAPVHNPIV					
THQIYPTQPLAPVHNPIV					
TQPLAPVHNPIV					

Identified peptides	Found peptide protein of origin	Overlapping literature functional peptide	Known activity	Known functional peptide origin	Ref
VLPFQQVVYPQRAVPVQALLLNQELLLNPTHQIYPVTQPLAPVHNPIIS					
VLPFQQVVYPQRAVPVQALLLNQELLLNPTHQIYPVTQPLAPVHNPIISV					
VPVQALLLNQELLLNPTHQIYPVTQPLAPVHNPIIS					
VPVQALLLNQELLLNPTHQIYPVTQPLAPVHNPIISV					
VQALLLNQELLLNPTHQIYPVTQPLAPVHNPIIS					
VQALLLNQELLLNPTHQIYPVTQPLAPVHNPIISV					
VTQPLAPVHNPIISV					
YPVTQPLAPVHNPIISV					