Journal of Chromatography A, 1227 (2012) 219–233

Contents lists available at SciVerse ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Gel-free shotgun proteomic analysis of human milk

Gianluca Picariello[a,*], Pasquale Ferranti[a,b], Gianfranco Mamone[a], Iveta Klouckova[c], Yehia Mechref[c], Milos V. Novotny[c], Francesco Addeo[a,b]

[a] Istituto di Scienze dell'Alimentazione – Consiglio Nazionale delle Ricerche (CNR), Via Roma 64, 83100 Avellino, Italy
[b] Dipartimento di Scienza degli Alimenti, University of Naples "Federico II," Parco Gussone, Portici (NA) 80055, Italy
[c] National Center for Glycomics and Glycoproteomics, Department of Chemistry, Indiana University, Bloomington, IN 47405, USA

ARTICLE INFO

Article history:
Received 1 December 2011
Received in revised form 2 January 2012
Accepted 3 January 2012
Available online 11 January 2012

Keywords:
Human milk
Milk fat globule membrane proteins
Shotgun proteomics
Fourier Transform-Ion Cyclotron Resonance mass spectrometry
Host defense proteins
Lipid droplet proteins

ABSTRACT

The composition of milk has adapted during the evolution of the species to fulfill the specific nutritional needs of the offspring. Currently, it is widely recognized that milk benefits go beyond mere nutrition and serve as a source of a number of functional components to the newborn, particularly host defense effectors. However, the human milk proteome description is still incomplete, primarily because the detection of low-abundance proteins remains challenging. To overcome the limitations of the classical electrophoresis-based approach, previously separated milk fat globule membrane (MFGM) and whey protein fractions were analyzed by nanoflow-high performance liquid chromatography (HPLC)/Fourier Transform-Ion Cyclotron Resonance (FT-ICR) mass spectrometry (MS). This shotgun strategy showed an as yet unmatched potential to profile low-abundance proteins in human milk. Proteins associated with 301 different gene products were identified, some of which could be clustered into subsets of protein isoforms, thus providing one of the largest protein inventories of human milk. The identified proteins, which were derived from multiple metabolic pathways, are involved in different physiological functions, such as membrane trafficking, cell signaling, fat metabolism and transport, metabolite delivery, protein synthesis/proteolysis or folding, and immunity-related actions. Nevertheless, it appears clear from this study that the overall picture of the human milk proteome is still incomplete, although several protein signatures of milk evolution are emerging.

1. Introduction

Milk is a complex biological fluid because it contains a wide range of macro- and micro-nutrients. It has been estimated that the entire mammary gland machinery includes several thousand genes whose evolution can be traced back to more than 300 million years [1]. The astounding preservation of the protein complexity along mammalian phylogenesis signifies the essential role of breast milk for term, pre-term and small-for-date human infants, which goes beyond mere nutritional aspects. Therefore, a basic understanding of the functions accomplished by breast milk, either contingently or in terms of "imprinting", has become a central theme in nutrition research. Many investigations have attempted to comprehensively characterize the proteins in human colostrum and mature milk.

Human milk proteins can be grouped into three major classes: caseins; milk fat globule membrane (MFGM) proteins and predominant whey proteins. Whey contains several low-abundance blood serum-derived proteins in addition to those secreted by neutrophils, by lymphocytes that have infiltrated the mammary gland and by somatic cells. Similar to other biological fluids, the wide range of protein concentrations in the human milk proteome, thought to span eight to ten orders of magnitude, presents a major challenge to systematic cartography. In fact, six or seven polypeptide chains constitute more than 90% of the human milk protein content, and the remaining content is distributed throughout hundreds of cellular proteins [2].

An early proteomic study based on two-dimensional electrophoresis (2DE) and microsequencing identified 22 proteins in human colostrum [3] although as many as 400 spots were observed. Later, through the "classical" 2DE-mass spectrometry (MS) proteomic approach, 107 protein spots, corresponding to 39 gene products, were identified in the human MFGM fraction [4,5]. Indeed, the index of proteins in human milk was increased to 73, including polypeptides identified in the whey fraction [6,7]. Many strategies have been developed for the selective removal of albumin and other highly abundant proteins from human and bovine milk (or colostrum). Nevertheless, only 15 proteins were identified in bovine mature milk and colostrum by 2DE and microsequencing after immunodepletion [8]. Immunoadsorption of the major proteins followed by mono-dimensional (1D) electrophoresis and

* Corresponding author. Tel.: +39 0825 299111; fax: +39 0825 782599.
E-mail address: picariello@isa.cnr.it (G. Picariello).

0021-9673/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2012.01.014

MS identified 151 gene products in human whey colostrum, 83 of which were reported for the first time [9]. Using SDS-PAGE coupled with MS, 120 non-redundant gene products were identified in bovine MFGM protein fractions [10], while 95 and 72 distinct gene products were identified in bovine [11] and water buffalo milk [12], respectively, by liquid chromatography (LC)–MS and 2DE-MS. The non-linear compression of the dynamic range via combinatorial peptide ligand libraries (ProteoMiner), though recently debated [13,14], has been definitively demonstrated to be effective in improving the coverage of complex proteomes in shotgun experiments [15]. Thus, equalization through peptide libraries was successful in cataloguing 149 unique proteins in bovine milk [16] and 115 in human whey [17].

Merging together the results of past studies, a total of 285 non-redundant database entries were catalogued in the inventory of the human milk proteome, including a "core" of 106 proteins that are shared with the bovine milk proteome [18].

The quasi-totality of studies aimed at characterizing proteomes at different depths (deep vs. shallow) relies upon an electrophoretic separation step prior to MS analysis. Conventional 2DE for descriptive proteomics suffers from a number of limitations that are particularly critical for characterizing complex mixtures. For example, very low-abundance proteins, as well as those with extreme isoelectric points or molecular weight values, can escape detection on the gels. In spite of the recent advances in 2DE, the detection and identification of very hydrophobic membrane proteins, such as MFGM proteins, is affected by their solubility [19]. The protein chip array of the SELDI-TOF (surface enhanced laser desorption ionization-time of flight) MS approach was also unsatisfactory to resolve the complexity of the human milk proteome, with particular concern for the MFGM protein fraction [2]. The well-known drawbacks of the 2DE-MS approach have prompted the development of gel-free strategies for high-throughput proteomics, especially when membrane proteins are to be analyzed [20,21]. In particular the use of high resolution capillary chromatographic columns coupled to ultra high mass accuracy (less than 1 ppm error tolerance) LTQ-Fourier Transform-Ion Cyclotron Resonance (FT-ICR) MS was shown to be capable of identifying thousands of peptide components in a single 1D high-resolution chromatographic separation [22]. With this technique, as a general procedure, an entire protein extract is digested with trypsin before nanoflow LC-electrospray (ESI) and MS analysis. In the present work, a gel-free shotgun approach has been exploited to penetrate deeper into the human milk proteome. Very recently, with a UPLC-Orbitrap-based shotgun MS approach, targeted to identify host-defense proteins, 268 and 269 gene products were identified in human and bovine milk, respectively. These results confirm that gel-free techniques are more appropriate for analyzing complex proteomes than the classic electrophoresis-based proteomic approaches [23]. Whey and MFGM proteins were separately trypsinized and analyzed by nanoflow reversed-phase HPLC coupled on-line with high-resolution FT-ICR MS. Validation of data sets was performed with automated bioinformatic tools.

2. Materials and methods

2.1. Sample preparation

Human milk samples were collected from five healthy Italian women by manual expression in a sterile tube, 10–15 days after full-term labor, and pooled (5 mL each). To prevent undesired proteolysis, PMSF was added to a final concentration of 1 mM. Milk freeze–thaw cycles were avoided to prevent breakdown of the fat globule membrane; samples were transferred to the laboratory in a refrigerated box and utilized for protein extraction within 2 h. Milk was gently shaken, and aliquots of 15 mL were centrifuged at 3500 rpm for 20 min at 4 °C to separate the fat, which was manually removed. MFGM proteins were extracted as described by Fortunato et al. [4]. The scarce pellet that resulted from the milk centrifugation, containing somatic cells and large casein micelles, was discarded. The skim milk fraction that here is called "milk whey" still contained a significant amount of casein. No attempt was carried out to completely remove casein by isoelectric precipitation in order to avoid the depletion of possible hydrophobically interacting minor proteins. Milk whey was filtered through Millex sterile 0.22 μm membranes (Millipore, Bedford, MA, USA) and precipitated by adding 4 volumes of −20 °C cold acetone/10% (w/v) trichloroacetic acid (TCA) containing 2% 2-mercaptoethanol. The protein pellet was washed three times with −20 °C cold acetone, separated by centrifugation (4 °C, 20 min, 12,000 rpm) and lyophilized.

2.2. Reduction and alkylation of MFGM and milk whey proteins

Disulphide bridges were reduced by incubating approximately 300 μg of both human MFGM and milk whey proteins in 1 mL of a denaturing/reducing buffer (6 M guanidine HCl, 100 mM Tris–HCl, 1 mM EDTA, 10 mM DTT, pH 8.0) for 1 h at 56 °C. Cysteine residues were alkylated by incubation for 30 min with iodoacetamide at a final concentration of 55 mM at room temperature in the dark. To quench the alkylation reaction, proteins were immediately desalted by PD-10 columns using 50 mM ammonium bicarbonate (AMBIC), pH 8.5, as the eluent. Proteins were quantified by the Bradford assay using BSA as the standard and incubated overnight at 37 °C with modified proteomic grade trypsin (Promega, Madison, WI, USA) using a 50:1 (w/w) protein-to-trypsin ratio. The reaction was stopped by freeze-drying. To prevent the missed detection of peptides due to phosphorylation, digests were re-dissolved in 0.5 mL of AMBIC at pH 8.5 and dephosphorylated overnight at 37 °C with 5 enzymatic units of alkaline phosphatase (Roche, Basel, Switzerland). Finally, to remove AMBIC, peptide solutions were repeatedly lyophilized after the addition of a few drops of aqueous 0.1% TFA. Alkaline phosphatase was discarded from the list of identified proteins in human milk.

To enlarge the proteome coverage, aliquots of the peptides derived from MFGM and milk whey proteins were subjected to N-deglycosylation (18 h, 37 °C, 2 mU/100 μg, in AMBIC) by PNGase F (Roche) and analyzed separately.

2.3. LC–MS/MS analysis

Tryptic digests were separated using an UltiMate 3000 nano-LC system (Dionex, Sunnydale, CA, USA). Peptides derived from 3 μg of the protein fractions were loaded by the autosampler on a Pep-Map300 C18 trapping cartridge (5 μm, 200 Å, Dionex). The nanocolumn used for nano-LC-ESI/MS/MS runs was prepared by packing a slurry of a 90 Å Jupiter C12 bound phase (Phenomenex, Torrance, CA, USA) into a 150 mm × 75 μm i.d. pulled-tip fused silica (Polymicro Technologies, Phoenix, AZ) using pressurized air. The column was washed with 80% acetonitrile mobile phase containing 0.1% formic acid and subsequently equilibrated with 3% acetonitrile with 0.1% formic acid at a 250 nL/min flow rate.

Solvent A was 3% acetonitrile in HPLC-grade water containing 0.1% formic acid, while solvent B was HPLC-grade acetonitrile containing 0.1% formic acid. The separation was carried out at a flow rate of 250 nL/min with a linear gradient from 3% to 55% of solution B over 150 min following 10 min of isocratic elution at 3% of solution B.