Journal of Chromatography B, 876 (2008) 252–256

Contents lists available at ScienceDirect

Journal of Chromatography B

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

Comprehensive proteomic analysis of the human milk proteome: Contribution of protein fractionation

A. Mangé[a,b,c], V. Bellet[a,b,c], E. Tuaillon[a,d], P. Van de Perre[a,d], J. Solassol[a,b,c,*]

[a] University of Montpellier I, Montpellier, France
[b] CHU Montpellier, Hôpital Arnaud de Villeneuve, Department of Cellular Biology, 191 Avenue du Doyen Giraud, 34295 Montpellier Cedex 5, France
[c] Val d'Aurelle Cancer Institute, Department of Clinical Oncoproteomic, Parc Euromédecine, 208 rue des Apothicaires, 34298 Montpellier Cedex 5, France
[d] CHU Montpellier, Hôpital Arnaud de Villeneuve, Department of Bacteriology/Virology, 191 Avenue du Doyen Giraud, 34295 Montpellier Cedex 5, France

ARTICLE INFO

Article history:
Received 2 May 2008
Accepted 2 November 2008
Available online 7 November 2008

Keywords:
Fractionation
Milk
Proteomics
Low-abundance proteins

ABSTRACT

In-depth analysis of the milk proteome by mass spectrometry is challenged by the presence of few high-abundance proteins that interfere with the detection of lower-abundance proteins. Here, we evaluated the proteomic analysis of milk samples following a strong anion exchange fractionation procedure using denaturating conditions ensuring the disruption of protein–protein interactions. Crude whey or skim milk and their different resulting fractions were analyzed by protein chip array mass spectrometry. Using protein chip array mass spectrometry, several high-abundance proteins were localized in distinct fractions increasing the total number of unique peptides and proteins detected. This total number increased by about 20–30% by combining different chromatographic surface arrays used for capture. Reproducible results were obtained in human skim milk and whey; however this approach was not successful with milk fat globule membrane and required refinement. Hence, milk profiling by anion exchange fractionation combined to protein chip array mass spectrometry represents a promising tool to detect unknown low-abundance milk proteins that may ultimately prove useful as biomarkers of diseases transmitted by breastfeeding.

1. Introduction

Measuring changes in the expression of multiple proteins in biological fluids provides a powerful means of identifying new biomarkers of potential medical interest. Recent studies have emphasized the need for large-scale proteomic approaches that, rather than focus on single protein markers which may not be sufficiently specific as a diagnostic tool, focus on patterns of protein co-expression through proteomic profiling. Two-dimensional polyacrylamide gel electrophoresis (2-DE) allows high throughput protein separation and identification. However, this process remains both labor-intensive and difficult to standardize between laboratories limiting its clinical applications for large scale clinical proteomic investigations. Surface enhanced laser desorption ionization time-of-flight mass spectrometry (SELDI-TOF MS) is an interesting analytical tool to profile the protein content of human fluids. The SELDI-TOF MS system utilizes affinity surfaces to capture proteins according to their biochemical properties and measures the mass of retained proteins by mass spectrometry. Since SELDI-TOF MS offers high-throughput protein profiling of native biological specimens, this technology has been used successfully to detect several disease-related proteins in complex biological specimens, such as cell or tissue lyzates or serum [1].

Surprisingly, milk protein investigations using 2-DE have generally given poor results and none have used SELDI-TOF MS. This can probably be explained by the large dynamic range of protein concentrations in milk, rendering its proteome difficult to characterize. Casein, α-lactalbumin, lactotransferrin, IgA, secretory component and human serum albumin (HAS) represent the six most abundant proteins (on a mg mL^{-1} basis) and constitute more than 90% of the mass of total milk proteins in human milk. The remaining 10%, believed to represent the population of biological interest, is composed of thousands of very low-abundance proteins that are very difficult to detect without specific sample preparation. As a result, methods to reduce the complexity of samples by removing highly abundant proteins are rapidly becoming an essential first step in many schemes for proteomics analysis. Several different strategies

Abbreviations: SELDI-TOF MS, Surface-enhanced laser desorption/ionization Time-of-flight mass spectrometry; MALDI, matrix-assisted laser desorption/ionization; HAS, human serum albumin; MW, molecular weight; p*I*, isoelectric point.

* Corresponding author at: Bâtiment de Recherche CRLC Val d'Aurelle, Parc Euromédecine, 208 rue des Apothicaires, 34298 MONTPELLIER CEDEX 5, France. Tel.: +33 4 67 61 24 12; fax: +33 4 67 33 95 90.
E-mail address: jerome.solassol@univ-montp1.fr (J. Solassol).

1570-0232/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2008.11.003

to achieve this goal have been described for human milk samples including isoelectric focusing electrophoresis, ultracentrifugation and depletion [2,3]. Chromatographic separation has been shown to improve protein or peptide detection in other complex samples such as serum [4] or bovine milk [5], however this has not as yet been used on human milk.

The aim of our investigation was therefore to validate a novel technique using SELDI-TOF MS associated with an anion-phase solid chromatographic fractionation assay and denaturating conditions to study milk protein expression. In the first part of the study, we assessed the accuracy of the fractionation procedure and we subsequently evaluated its reproducibility and ability to increase the detection of low-abundance proteins by SELDI-TOF mass spectrometry.

2. Materials and methods

2.1. Human milk and protein assay

Human milk samples were collected from consenting volunteers by bimanual expression into polypropylene containers in the Arnaud de Villeneuve Hospital, in Montpellier, France. Milk was stored on ice during transport to the laboratory (up to 2 h), and then centrifuged (1000 × *g*, 20 min, 4 °C) to remove the lipid layer and cell debris pellet. The aqueous phase (skim milk) was then aliquoted with protease inhibitor (Complete mini EDTA-free, Roche Applied Science) and stored at −80 °C. An aliquot of the skim milk was centrifuged at 100,000 × *g* for 60 min to pellet the casein micelles, and the clear supernatant (whey) was removed and stored at −80 °C. Milk fat globule membrane (MFGM) fractions were produced by sonication and subsequent high-speed centrifugation (100,000 × *g* for 90 min at 30 °C) of the washed and resuspended fat layer produced by the initial centrifugation, as previously described [6]. The pellet from the high-speed centrifugation was resuspended in water and stored at −80 °C. Protein concentrations of all the samples were determined by the bicinchonic acid method (BCA) (Pierce), using BSA as a standard.

2.2. Milk fractionation

For the anion exchange fractionation, 20 μl of the milk samples were denaturated in 9 M Urea, 2% 3-[(2-cholamidopropyl)-dimethylammonio]-1-propane-sulfonate hydrate (CHAPS), and 50 mM Tris–HCl, pH 9. They were separated on a 96-well filter plate containing Qhyper DF resin (BioSepra Corp.) into six different fractions according the manufacturer's instructions. The wash buffers for the different fractions were 50 mM Tris–HCl, 0.1% *n*-octyl b-D-glucopyranoside (OGP), pH 9 (F1 + flow through), 50 mM Tris–HCl, 0.1% OGP, pH 7 (F2), 100 mM Na-acetate, 0.1% OGP, pH 5 (F3), 100 mM Na-acetate, 0.1% OGP, pH 4 (F4), 50 mM Na-citrate, 0.1% OGP, pH 3 (F5) and 33.3% isopropanol/16.7% acetonitrile/0.1% trifluoroacetic acid (F6).

2.3. SELDI-TOF MS analysis

Each sample was analyzed on four different array surfaces: hydrophobic (H50), cation-exchange (CM10), anion-exchange (Q10) and metal binding (IMAC30-Cu) arrays. Samples were randomly applied to a ProteinChip array surface (Ciphergen Biosystems, Inc.) in a 96-well format. IMAC arrays were pre-treated with 50 μL of 100 mM $CuSO_4$ with vigorous shaking for 5 min, washed twice with 100 μL of water and neutralized with 100 mM sodium acetate pH 4 for 5 min. ProteinChip arrays were equilibrated twice with 100 μL of binding buffer on a shaker for 5 min. Five micrograms of each sample was added to the binding buffer, and the arrays were incubated for 30 min with vigorous shaking. Arrays were then washed three times with 100 μL of binding buffer, followed by a final wash with water. Binding buffers used for the different arrays were 10% CH_3CN, 0.1% trifluoroacetic acid (TFA) for H50; 100 mM NaAC pH 4.0 for CM10; 100 mM Tris–HCl pH 9.0 for Q10; and 100 mM phosphate buffer pH 7.0, 500 mM NaCl for IMAC30-Cu. Arrays were removed from the bioprocessor and allowed to air dry. One microlitre of 50% sinapinic acid was applied twice to the spots. Finally, arrays were analyzed on a PBSIIc ProteinChip Reader (Ciphergen Biosystems, Inc.). The data were averaged over 200 laser shots for each spot. Mass detection accuracy of PBSIIc was calibrated externally by using the All-in-1 protein II molecular mass standards (Ciphergen Biosystems). To minimize experimental variation, samples were analyzed concurrently, and no sample was subjected to more than two freeze–thaw cycles. In all experiments, SELDI-TOF MS profiles were obtained in triplicate for each sample.

2.4. Peak detection

Peak detection was performed using the ProteinChip Biomarker software (Version 3.2.0, Ciphergen Biosystem Inc.). The background was subtracted from the spectra, and the peak intensities were normalized to the total ion current of m/z between 2.5 and 50 kDa. Automatic peak detection was performed with the following settings: (i) signal-to-noise ratio at 4 for the first pass and 2 for the second pass, (ii) minimal peak threshold at 15% of all spectra, (iii) cluster mass window at 0.5% of mass. The resulting file, containing absolute intensities and m/z ratios, was exported into Microsoft Excel (Microsoft, Redmond, WA) for subsequent analysis.

2.5. Reproducibility

Reproducibility was estimated using a control sample which was randomly spotted on each chip to measure the variability of on-chip spotting and data acquisition. The reproducibility analysis was performed for all surface conditions, and spectra were collected following the experimental protocols described above. The coefficients of variation (CV) of 15 reliable peaks were calculated by averaging peak intensity values derived from different runs. The average CVs were 10% and 20%, respectively, for intra- and inter-assay variability, which is consistent with previously reported studies [7–10].

3. Results and discussion

In order to minimize the risk of losing potentially important proteins and peptides from the milk sample in the course of the sample preparation process, we investigated anion-exchange fractionation of milk as an alternative to depletion by immune-affinity or by dye–ligand chromatography. The general steps of the protocol are outlined in Fig. 1. Twenty microliters of human skim milk or whey, containing approximately 190 μg of total protein, were incubated with Q Ceramic HyperD F anion-exchange resin in low protein binding filter plates. Peptides and proteins were eluted with a pH gradient using six different aqueous buffers, followed by a final elution step with an organic wash buffer. The total protein content in each fraction of skim milk was estimated by BCA analysis (Table 1). Interestingly, a relative protein recovery of about 95–98% of total protein was observed, with a higher protein concentration in fraction 1.

SELDI-TOF MS has combined specially designed protein chip arrays with surface ionization mass spectrometry to provide an interesting method for identifying new diagnostic or prognostic tumor markers in complex biological fluids. In most studies, diluted or depleted samples are spotted onto different protein chip arrays

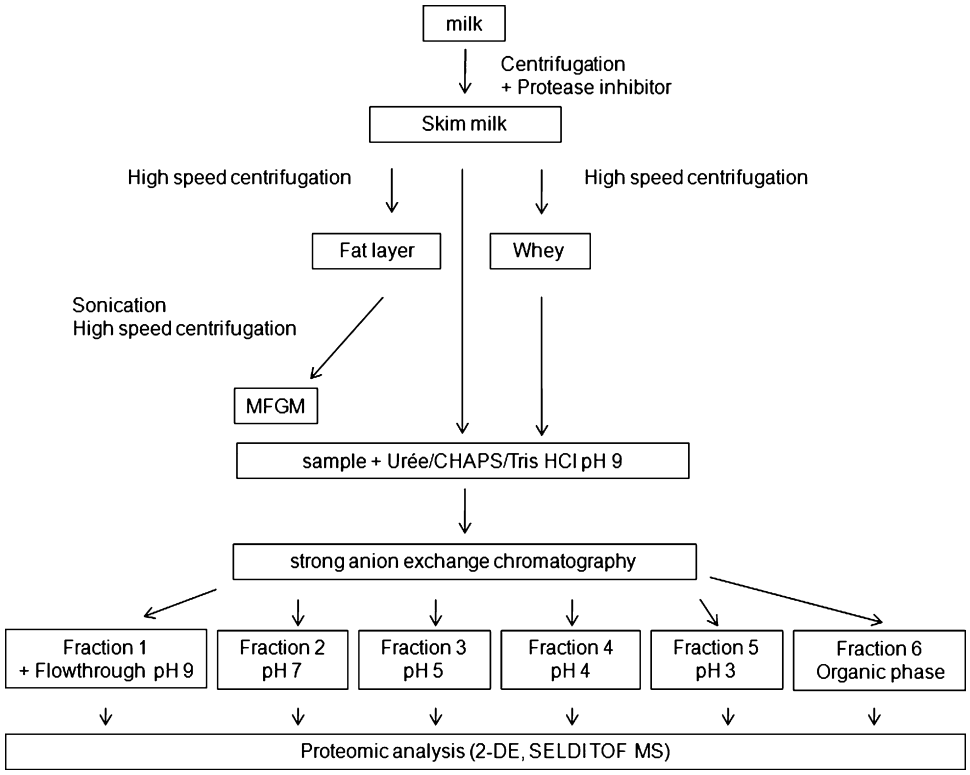

Fig. 1. Flow-chart illustrating the general steps of the protocol for fractionation of milk.

and analyzed by SELDI-TOF MS. However, the limited protein binding capacity of the protein chip arrays has led to the observation that only high- and medium-abundance proteins can actually be detected by this approach [11]. Since our goal was to maximize the number of peptides and proteins under investigation, we assessed the milk fractionation protocol on different chromatographic surface arrays and analyzed the proteins by SELDI-TOF MS. Equal amount of fractionated skim milk samples were applied on four surface arrays with anionic (Q10), cationic (CM10), hydrophobic (H50) or copper metal binding (IMAC-Cu) surface chemistries. A representative view of the distribution of protein/peptides in fraction 3 is shown in Fig. 2. Ionic surfaces chemistries (Q10 and CM10) gave the best results in terms of detectable protein peaks, with more than 500 protein peaks in the six fractions and under each condition, while IMAC-Cu resulted in around 400 protein peaks in the six fractions (data not shown). The last surface array (H50) cannot be considered a relevant array for milk protein detection since only a few protein peaks with low signal-to-noise ratios were observed. We next applied equal amounts of non-fractionated (crude) and fractionated skim milk samples directly onto cationic ProteinChip CM10 arrays. In the analysis of low molecular weight range from 2.5 to 50 kDa, more than 500 peptide and protein peaks were effectively detected in the six fractions, compared to only 90 in the crude skim milk, with an S/N ratio >4. A similar result was also obtained from the fractionated spectra in the high-molecular-mass region (spectra not shown). As an example, Fig. 3A illustrates a representative view between 2.5 and 8 kDa of the molecular mass profile of the different fractions in comparison to the total non-fractionated skim milk sample, and shows that several peptides and proteins appear only in some fractions. Although some overlapping of the detected protein peaks occurred between the different fractions, we estimated that, in each fraction, there was a 20–40% increase in the number of detectable protein peaks compared with those detected in the crude skim milk. Moreover, by combining the results from the

Table 1
Protein content and recovery in crude and fractionated milk.

	Milk 1		Milk 2	
	Protein content (μg/sample)	Recovery (%)	Protein content (μg/sample)	Recovery (%)
Skim Milk	195.3	100.0	179.4	100.0
Fraction 1	69.8	35.7	70.7	39.4
Fraction 2	7.9	4.1	71.2	4.0
Fraction 3	31.7	16.2	27.2	15.2
Fraction 4	23.2	11.9	20.2	11.2
Fraction 5	25.9	13.3	26.4	14.7
Fraction 6	28.9	14.8	23.1	12.9
Total fractions	187.4	95.9	174.7	97.4

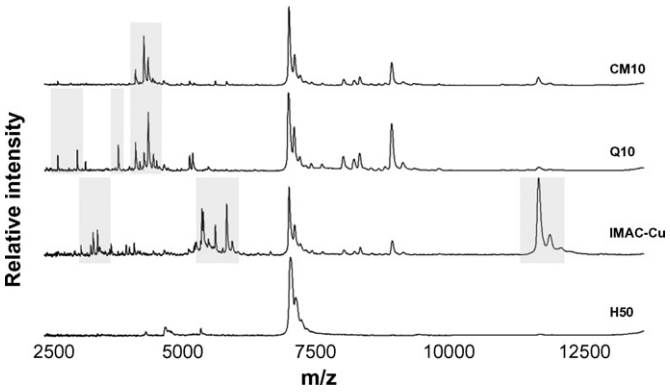

Fig. 2. SELDI-TOF MS profiles of skim milk sample. Five μg of the same skim milk sample was fractionated and incubated in a bioprocessor on CM10, Q10, IMAC-Cu and H50 protein chip arrays and subsequently analyzed by SELDI-TOF MS. Mass spectra were acquired with sinapinic acid absorbing matrix. The figure shows a representative view of fraction 3 obtained on each surface array. The relative intensity is displayed along the *y*-axis, and the mass is given as *m*/*z* ratio on the *x*-axis.

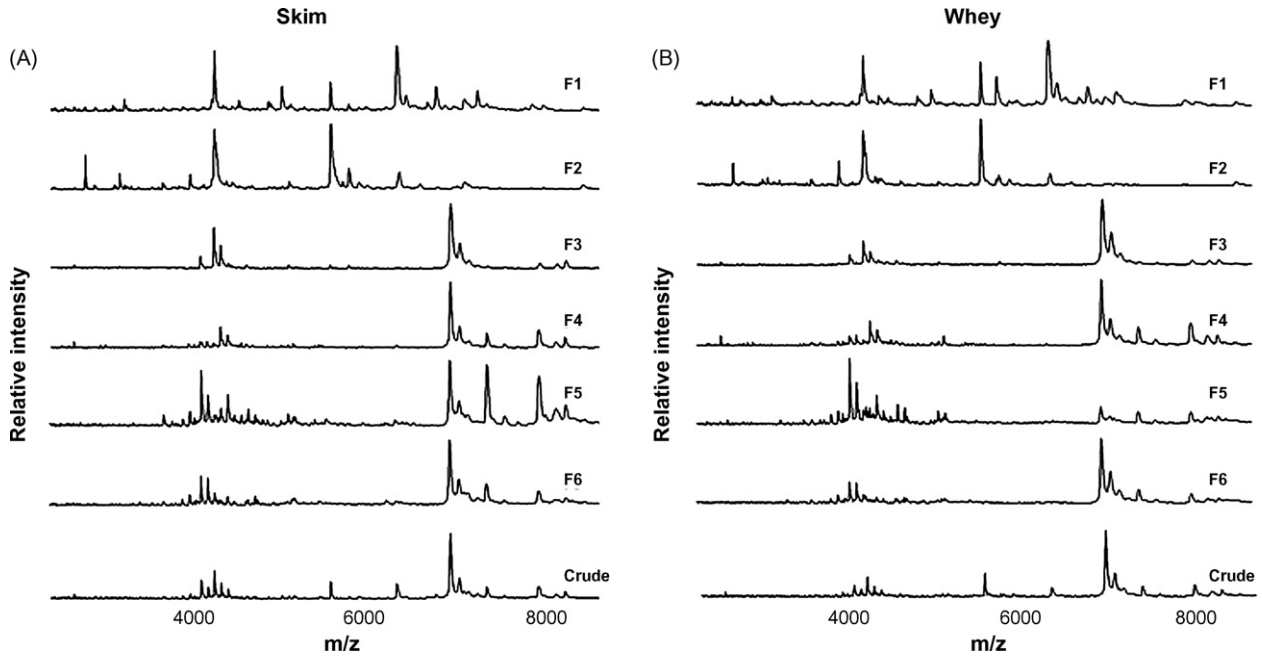

Fig. 3. Comparison of SELDI-TOF MS profiles of crude and fractionated skim milk (A) and crude and fractionated whey (B) from the same milk sample, and incubated on CM10 protein chip arrays. The relative intensity is displayed along the *y*-axis, and the mass is given as *m*/*z* ratio on the *x*-axis.

different affinity capture surface arrays, the total number of protein peaks detected in a single sample was increased by 20–30%.

Using our experimental approach, protein samples undergo a 2D fractionation process that depends, first, on the biochemical properties of proteins for the strong anion exchange chromatography and, second, on the chemical properties of the chromatographic surface used. The denaturating conditions used in our fractionation procedure is also an important step causing the disruption of protein–protein interactions and preventing the elimination of minor proteins that would normally bind to major milk protein species, a problem observed in serum samples with classical affinity-based procedures [12,13]. The result is the ability to obtain a more exhaustive view of the peptides and proteins present in the skim milk sample.

Prior to molecular analysis, milk is usually pre-fractionated by centrifugation into its three major components: soluble whey, skim, and a floating MFGM (Fig. 1). To test whether this supplementary fractionation approach combined with anion-exchange fractionation could improve the peptide mass spectrometry, we performed SELDI-TOF MS on the whey fractions using cationic ProteinChip CM10 arrays (Fig. 3B), and compared spectra to that obtained on the skim milk samples (Fig. 3A). The overlap of the protein profiles was very high, with less than 2% variation between the two samples. Interestingly, the attempt to remove casein by centrifugation was only partially successful, as protein peaks assigned to β caseins were still observed in the whey sample profiles (data not shown). The MFGM fraction consists of the protein-rich bilayer which normally surrounds triglyceride droplets. These membrane proteins are highly informative, since they could be associated with several factors implicated in the signaling and secretory pathways of the mammary gland [14] or prove useful as biomarkers of diseases transmitted by breastfeeding [15,16]. To date, human MFGM has been investigated mainly by 2-DE [15,17–21], however this classical protein approach fails to determine the membrane protein composition of the MFGM. The MGFM samples were analyzed using 4 different surface arrays (data not shown). We observed very poor profiles in terms of the number of protein peaks detected, revealing the need for more sensitive proteomic strategies, notably in the field of mass spectrometry, such as liquid chromatography mass spectrometry for MFGM analysis as described for bovine milk [5].

Finally, according to the *m*/*z* and approximate p*I*, it was possible to suggest identities for several abundant milk proteins detected in skim milk by SELDI-TOF MS analysis using cationic ProteinChip CM10 arrays, and to localize some high-abundance proteins depending on the fraction analyzed (Fig. 4). The 80,135.7 Da protein peak was identified as lactotransferrin (expected molecular weight of 76,165.29 Da, p*I* 8.47) and was mainly detected in fraction 2. The 23,946.2 Da and the 66,312.1 Da protein peaks, detected in fractions 4, 5 and 6, were identified as β-casein (expected molecular weight of 23,857.82 Da, p*I* 5.33) and seroalbumin (expected molecular weight of 66,472.1 Da, p*I* 5.67), respectively. Finally, the 14,032.6 Da protein peak was identified as α lactalbumin (expected molecular weight of 14,078.15 Da, p*I* 5.70) and was observed in fractions 3–6.

Milk proteomic analysis appears to be a promising approach to screening for new biological markers of or potential therapeutic targets for diseases transmitted by breastfeeding such as cytomegalovirus, human immunodeficiency virus, and human T-lymphotropic virus type I [16,22]. SELDI-TOF MS is a powerful profiling technology allowing the discovery of many markers in samples of a large cohort with good statistical relevance. However, it is negatively affected by the presence of high-abundance proteins yielding only a fraction of the peptide and protein signals thought to be present. Our study shows that anion exchange fractionation of milk samples represents an easy-to-use method of increasing the number of unique peptide and proteins detectable by proteomic analyses. Most importantly, our results are the first demonstration that milk fractionation and SELDI-TOF MS analysis could eventually be used to highlight specific profiling associated with the transmission of infectious agents by breastfeeding. Although there have been several controversies surrounding the use of SELDI-TOF, especially with regards to the identification of biomarkers, this technology nevertheless represents a relevant profiling method which will help focus time and research efforts on the most promising biomarker candidates in milk [23,24].

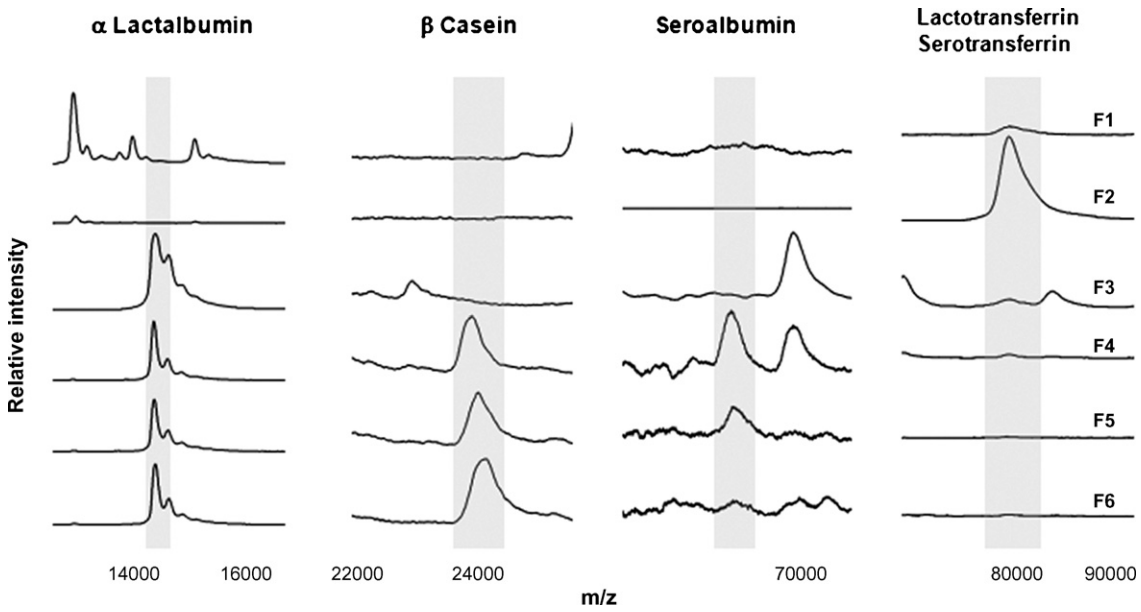

Fig. 4. Localization of some high-abundance proteins on the SELDI-TOF MS profiles obtained from skim milk fractions on CM10 protein chip arrays. Alpha-lactalbumin, β caseins, seroalbumin, lacto- and serotransferrin and immunoglobulins peaks are clearly compartmentalized in specific fractions and "boxed" in grey. The relative intensity is displayed along the *y*-axis, and the mass is given as *m*/*z* ratio on the *x*-axis.

Acknowledgment

We are deeply grateful to Pr. G. Favre and Dr. Ben Allal (INSERM-U563, Toulouse, France) for the use of the SELDI-TOF MS facilities.

References

[1] J. Solassol, W. Jacot, L. Lhermitte, N. Boulle, T. Maudelonde, A. Mange, Expert Rev. Proteomics 3 (2006) 311.
[2] D.J. Palmer, V.C. Kelly, A.M. Smit, S. Kuy, C.G. Knight, G.J. Cooper, Proteomics 6 (2006) 2208.
[3] M. Yamada, K. Murakami, J.C. Wallingford, Y. Yuki, Electrophoresis 23 (2002) 1153.
[4] S.A. Hoffman, W.A. Joo, L.A. Echan, D.W. Speicher, J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. 849 (2007) 43.
[5] T.A. Reinhardt, J.D. Lippolis, J. Dairy Res. 73 (2006) 406.
[6] G. Smolenski, S. Haines, F.Y. Kwan, J. Bond, V. Farr, S.R. Davis, K. Stelwagen, T.T. Wheeler, J. Proteome Res. 6 (2007) 207.
[7] E.F. Petricoin, A.M. Ardekani, B.A. Hitt, P.J. Levine, V.A. Fusaro, S.M. Steinberg, G.B. Mills, C. Simone, D.A. Fishman, E.C. Kohn, L.A. Liotta, Lancet 359 (2002) 572.
[8] A.J. Rai, P.M. Stemmer, Z. Zhang, B.L. Adam, W.T. Morgan, R.E. Caffrey, V.N. Podust, M. Patel, L.Y. Lim, N.V. Shipulina, D.W. Chan, O.J. Semmes, H.C. Leung, Proteomics 5 (2005) 3467.
[9] O.J. Semmes, Z. Feng, B.L. Adam, L.L. Banez, W.L. Bigbee, D. Campos, L.H. Cazares, D.W. Chan, W.E. Grizzle, E. Izbicka, J. Kagan, G. Malik, D. McLerran, J.W. Moul, A. Partin, P. Prasanna, J. Rosenzweig, L.J. Sokoll, S. Srivastava, I. Thompson, M.J. Welsh, N. White, M. Winget, Y. Yasui, Z. Zhang, L. Zhu, Clin. Chem. 51 (2005) 102.
[10] Z. Zhang, R.C. Bast Jr., Y. Yu, J. Li, L.J. Sokoll, A.J. Rai, J.M. Rosenzweig, B. Cameron, Y.Y. Wang, X.Y. Meng, A. Berchuck, C. Van Haaften-Day, N.F. Hacker, H.W. de Bruijn, A.G. van der Zee, I.J. Jacobs, E.T. Fung, D.W. Chan, Cancer Res. 64 (2004) 5882.
[11] T. Linke, S. Doraiswamy, E.H. Harrison, J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. 849 (2007) 273.
[12] N.I. Govorukhina, A. Keizer-Gunnink, A.G. van der Zee, S. de Jong, H.W. de Bruijn, R. Bischoff, J. Chromatogr. A 1009 (2003) 171.
[13] L.F. Steel, M.G. Trotter, P.B. Nakajima, T.S. Mattu, G. Gonye, T. Block, Mol. Cell Proteomics 2 (2003) 262.
[14] T.W. Keenan, J. Mammary Gland Biol. Neoplasia 6 (2001) 365.
[15] J. Charlwood, S. Hanrahan, R. Tyldesley, J. Langridge, M. Dwek, P. Camilleri, Anal. Biochem. 301 (2002) 314.
[16] R.M. Lawrence, R.A. Lawrence, Clin. Perinatol. 31 (2004) 501.
[17] D. Fortunato, M.G. Giuffrida, M. Cavaletto, L.P. Garoffo, G. Dellavalle, L. Napolitano, C. Giunta, C. Fabris, E. Bertino, A. Coscia, A. Conti, Proteomics 3 (2003) 897.
[18] S. Quaranta, M.G. Giuffrida, M. Cavaletto, C. Giunta, J. Godovac-Zimmermann, B. Canas, C. Fabris, E. Bertino, M. Mombro, A. Conti, Electrophoresis 22 (2001 1810).
[19] K. Murakami, M. Lagarde, Y. Yuki, Electrophoresis 19 (1998) 2521.
[20] M. Cavaletto, M.G. Giuffrida, A. Conti, Clin. Chim. Acta 347 (2004) 41.
[21] M. Cavaletto, M.G. Giuffrida, A. Conti, Adv. Exp. Med. Biol. 606 (2008) 129.
[22] G. John-Stewart, D. Mbori-Ngacha, R. Ekpini, E.N. Janoff, J. Nkengasong, J.S. Read, P. Van de Perre, M.L. Newell, J. Acquir. Immune Defic. Syndr. 35 (2004) 196.
[23] G. Malik, M.D. Ward, S.K. Gupta, M.W. Trosset, W.E. Grizzle, B.L. Adam, J.I. Diaz, O.J. Semmes, Clin. Cancer Res. 11 (2005) 1073.
[24] N.T. Zinkin, F. Grall, K. Bhaskar, H.H. Otu, D. Spentzos, B. Kalmowitz, M. Wells, M. Guerrero, J.M. Asara, T.A. Libermann, N.H. Afdhal, Clin. Cancer Res. 14 (2008) 470.