Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography B, 845 (2007) 69–73
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Determination of proteins in infant formula by high-performance liquid chromatography–electrospray tandem mass spectrometry

Soon Sun Hong [a], Jeong Hill Park [a], Sung Won Kwon [a,b,*]

[a] *Research Institute of Pharmaceutical Sciences, College of Pharmacy, Seoul National University, Seoul 151742, Republic of Korea*
[b] *Protein Chemistry Technology Center, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA*

Received 15 May 2006; accepted 23 July 2006
Available online 5 September 2006

Abstract

To determine the protein content of formula, gel electrophoresis was performed on the infant formula samples and the entire protein patterns were analyzed by nano-high performance liquid chromatography–electrospray tandem mass spectrometry (nano-HPLC/ESI/MS/MS). From the commercial infant formula profiled in this study, a total of 154 peptides, corresponding to 31 unique proteins were identified by nano-HPLC/ESI/MS/MS. Each of the identified peptides was reconfirmed by a strict integrated approach using tandem mass spectra. This protein profiling method using gel electrophoresis coupled with nano-HPLC/ESI/MS/MS and manual evaluation is a sensitive and accurate method for protein identification as well as a powerful tool for monitoring various types of food products.

Keywords: Infant formula; High-performance liquid chromatography; Tandem mass spectrometry; Food analysis

1. Introduction

Breast feeding is natural in humans and is continued after childbirth because milk derived from human represents an essential supply of both nutritional and functional food. Milk generated from the mammary glands is a complex colloidal system, containing proteins, minerals, lipids, vitamins, lactose, and salts [1]. Human milk in particular contains a wide spectrum of proteins, some of which have biological activities such as antimicrobials and immune stimulators [2]. Therefore, it represents an essential source of nutrition for the growth of an infant and there is no clear substitute for breast feeding.

When breast feeding is not sufficient, cow's milk which is modified either qualitatively or quantitatively is generally used as a substitute for human milk for promoting the growth of an infant. However, it must be emphasized that there are differences between human and other mammal's milk even though various technologies are used to adjust the composition of infant formulae closer to that of human milk. One such feature is the ratio of casein proteins. Traditional bovine milk which is a blend of nutrients has a higher casein protein ratio than human milk. Literature reports indicate that casein proteins cause a load on the kidney and the uptake of poorly digestible high molecular proteins can lead to metabolic stress [3,4]. Maillard reactions also take place due to the thermal treatment which required for the pasteurization of bovine milk. Products generated by the addition of reducing sugars to lysine in Maillard reactions can be toxic to infants [5]. Therefore, considering the various treatments applied in the preparation of bovine milk, the proteins of an infant formula should be profiled.

The difference in protein components between human milk and infant formula are amenable to a protein profiling by mass spectrometry. To date, the analysis of milk proteins has been achieved using various techniques including classical gel electrophoresis, chromatographic methods, capillary zone electrophoresis, and a mass spectrometry-based proteomic approach [6–11].

In-gel digestion coupled with tandem mass spectrometry has recently emerged as a powerful analytical tool for general protein profiling. In the past, proteins in many infant formulae have been reported using different mass spectrometry and related techniques [12–18]. However, overall, only a small number of proteins have been identified. But a real field application requires exhaustive protein profiling which can be achieved by a more

* Corresponding author. Tel.: +82 2 880 7844; fax: +82 2 886 7844.
E-mail address: swkwon@snu.ac.kr (S.W. Kwon).

1570-0232/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2006.07.036

sensitive and simple technique. Based on the need to identify a high number of proteins, the profiling of infant formulae proteins has been carried out using gel electrophoresis coupled nano-HPLC/ESI/MS/MS and 31 unique proteins have been identified by this rigorous integrated approach. It is also predicted that the determination of proteins by nano-HPLC/ESI/MS/MS will extend our knowledge of protein components between an original food and a substitute.

2. Experimental

2.1. Chemicals and reagents

Commercially available infant formula derived from a bovine source was investigated in the study. The following reagents were purchased from commercial sources and were used without further treatment. HPLC grade-acetonitrile and water were obtained from EM Science (Gibbstown, NJ, USA). Trifluoroacetic acid (TFA) was obtained from Fluka (Buchs, Switzerland). Acetic acid was obtained from Aldrich (St. Louis, MO, USA). Ammonium bicarbonate, bromophenol blue, sodium dodecyl sulfate (SDS), monobasic sodium phosphate, and dibasic sodium phosphate were obtained from Fisher (Fair Lawn, NJ, USA). Ammonium persulfate, acrylamide, bis acrylamide, Tris–HCl, Tris base, glycine, glycerol, 2-mercaptoethanol, and bovine serum albumin (BSA) were obtained from Sigma (St. Louis, MO, USA). TEMED and Colloidal Coomassie Blue staining kit was obtained from Invitrogen (Carlsbad, CA, USA). Protein standards and Bradford protein assay kit were obtained from Bio-Rad (Hercules, CA, USA). Trypsin was obtained from Promega (Madison, WI, USA).

2.2. Sample preparation

The formula sample was dissolved in phosphate buffered saline (PBS), followed by centrifugation at $14{,}000 \times g$ for 30 min at 4 °C. The supernatant was then filtered through a cellulose acetate membrane filter, 0.22 μm pore size (Corning Inc., Corning, NY, USA) to remove insoluble compounds. Protein concentrations were determined to be 4.6 μg/μl by the Bradford assay. Ten micrograms of sample was mixed with mobile phase of sample and incubated for 5 min at 95 °C. After reaching room temperature, the solution was centrifuged at $10{,}000 \times g$ for 1 min and then loaded onto a home-made 12% (w/v) polyacrylamide gel with a protein ladder. The gel was stained using a Colloidal Coomassie Blue staining kit.

2.3. In-gel tryptic digestion

An entire lane was excised into small bands regardless of staining. The height of each gel was about 3–4 mm and the number of bands was 23. Each protein band was cut into 1 mm cubes, and destained with 25 mM NH_4HCO_3 in 50% methanol/50% water (v/v) three times for 5 min. They were then washed with 10% acetic acid/50% methanol/40% water (v/v/v) three times for an hour, and swollen in water twice for 20 min. The gel pieces were dehydrated with acetonitrile and dried in a SpeedVac (Thermo Savant, Holbrook, NY, USA). The gel pieces were again rehydrated with modified porcine trypsin at a concentration of 10 ng/μl in 100 mM NH_4HCO_3 and then subjected to trypsin digestion at 37 °C overnight. Tryptic peptides were sequentially extracted with 50% acetonitrile/45% water/5% TFA (v/v/v) and 75% acetonitrile/24.9% water/0.1% TFA (v/v/v) solutions. The peptide extracts were combined and dried in a SpeedVac. The peptide samples were cleaned with ZipTip C18 (Millipore, Billerica, MA, USA) prior to the nano-LC/MS/MS analysis.

2.4. Nano-HPLC/ESI/MS/MS

The nano-HPLC/ESI/MS/MS analysis was performed in an LCQ DECA XP ion-trap mass spectrometer (ThermoFinnigan, SanJose, CA, USA) equipped with a nano electrospray ionization (ESI) source. The electrospray source was coupled online with an Agilent 1100 series capillary HPLC system (Agilent, Palo Alto, CA, USA). Two microliters of peptide solution in mobile phase A (2% acetonitrile/97.9% water/0.1% acetic acid, v/v/v) was manually loaded into a capillary HPLC column (70 mm length × 75 μm ID) using a Rheodyne manual injector. The HPLC column was packed in-house with C18 resin and the ambient temperature was applied to the column during the analysis. The eluted peptides were electrosprayed directly into the LCQ mass spectrometer. MS/MS spectra were acquired in a data-dependent mode, in which the masses of the parent ions and fragments of three strongest ions were determined. All MS/MS spectra were searched against the NCBI non-redundant human protein sequence database using the Knexus program (Genomic Solutions, Ann Arbor, MI, USA) for protein identification.

2.5. Nano-HPLC conditions

Chromatographic separation was achived with a gradient elution program. The peptides were eluted from the column with a gradient of 15–80% mobile phase B (90% acetonitrile/9.9% water/0.1% acetic acid, v/v/v) in mobile phase A (2% acetonitrile/97.9% water/0.1% acetic acid, v/v/v) for 30 min. The flow rate was 0.2 μl/min and the injected volume of the sample was 1 μl. In order to maximize the number of proteins identified, the injection volume was adjusted to the half of injection loop to reduce laminar flow effects.

2.6. Evaluation of protein identification

MS/MS spectra acquired in a data-dependent mode were used to search the protein database with the automated search program, Knexus. The enzyme was specified as trypsin with one missing cleavage. Then the program was set to find peptides which have no or one incomplete digestion, ending in K or R residue at their C-terminus. The mass error for the parent ion mass was set to ±4 Da and for the fragment ion to ±0.6 Da. Manual analysis was applied to validate protein identification results. The following criteria were used for manually verifying the MS/MS spectra. Isotopically resolved peaks should match the calculated fragment ions because a single peak is less

likely to be relevant to peptide fragments. Only y and b ions and associated peaks arising from the loss of water or amine were considered. y/b ions mean fragment peptide ions which are generated on the right/left side of parent peptide. Specific amino acids such as proline residue, acidic residues, and large hydrophobic residues were also considered for manual verification because these amino acids show relatively strong intensities. Proteins that are identified by at least two peptides were also considered to be true protein identification. Proteins identified from a single spectrum of a peptide were neglected in order to avoid a false-positive identification, even when the spectrum was perfect for identification. All the redundant proteins were removed by confirming unique peptides and keratin proteins. Overlapping peptides in the excised gels were also removed by listing the identified peptides.

3. Results and discussion

For the purpose of unbiased protein profiling, one-dimensional SDS-PAGE coupled with nano-HPLC/ESI/MS/MS was applied and the proteins resolved by SDS-PAGE analysis of infant formula are shown in Fig. 1. Regardless of the visualized bands being separated, 23 gel slices were excised into same size, since an unstained band might contain proteins and were analyzed by nano-HPLC tandem mass spectrometry. If a two-dimensional SDS-PAGE system is applied to the proteomic analysis of infant formula, it would be difficult to perform an unbiased protein profiling due to the number of gel slices. Assuming the number of gel slices of a one-dimensional electrophoresis to be 23, as shown in this study, the total number of gel slices of two dimensional electrophoresis would be about 310. The figure of 310 arises from the calculation of nine lanes plus nine gaps between the lanes. The width of two gaps approximately equals that of one lane. This is the reason why one-dimensional SDS-PAGE was used for the impartial proteomic analysis (Fig. 1).

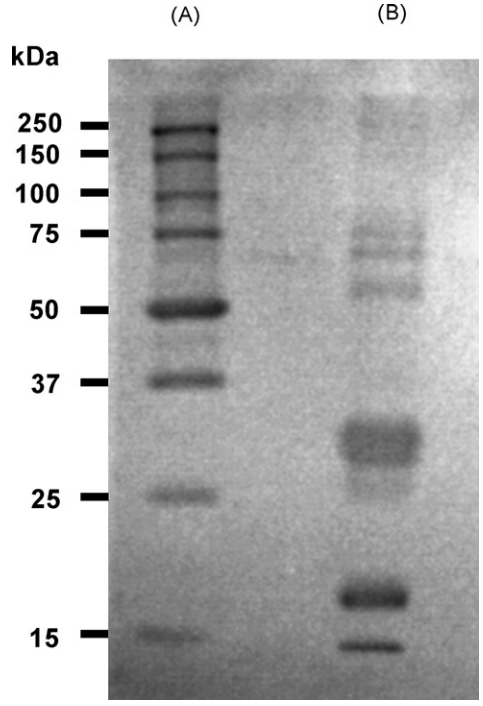

Fig. 1. SDS-PAGE separation of infant formula proteins. The proteins were resolved in 10% (w/v) acrylamide and stained with Simply Blue. (A) The molecular marker is shown in the left lane; (B) the infant formula proteins are shown in the right lane.

One hundred fifty four peptides corresponding to 31 unique proteins were unambiguously identified in this study. Proteins having only one identified peptide were excluded to avoid false-positive protein identification because spectra containing a false-positively identified single peptide were found occasionally even if numerous good matching tandem mass spectra of a single peptide were found. The confirmation of false-positive peptide identification was done by a manual comparison of spectra of synthetic peptides.

Consistent and solid rules were used as criteria in the manual verification. Although manual verification is the most accurate method of confirmation, the reliability of the validation varies with the environment of the analyst, and a systematic approach is needed to evaluate peptides. The rules for the systematic approach were as follows; (1) specific amino acids were also considered for manual verification since some residues showed intense peaks. For example, the fragmentation of a proline containing peptide showed a relatively large fragment pattern at the C-terminal of proline; (2) only y and b ions or the loss of water or amine ions were considered for peptide sequencing. Sometimes an ion can be considered for identification but not in this study because of the chance of false-positive matching ions; (3) five to seven isotopically resolved fragment peaks should match the theoretical fragmentation data; (4) all isotopically resolved peaks with intensities should match the theoretical fragmentation data; (5) the deviation in mass errors among fragment peaks should have tendency, i.e., a negative or positive deviation.

Fig. 2 shows a representative nano-HPLC/ESI/MS/MS analysis of peptides that were derived from proteins. Usually 700–1200 spectra were generated per injection, searched by the automated protein identification program, and examined by manual verification. By this sensitive and accurate proteomic analysis process, a large number of proteins were identified with confidence and this would help us to understand the difference between human milk and infant formula. Changes in proteins after Maillard reactions could also be observed by this technique. Meanwhile not more than 10 proteins from infant formula have been reported in the literature to date. By profiling proteins by nano-HPLC/ESI/MS/MS the adverse effects of food could be monitored for the purpose of clinical studies and an unbiased protein profiling technique based on mass spectrometry could be applied to a wide variety of foods in the near future (Fig. 2).

Table 1 shows the identified proteins and the representative unique peptide of each protein. The unique peptide was selected by searching the amino acid sequences of all identified proteins and the identified proteins were displayed according to their molecular weight. Some of the proteins identified in this study include human-like proteins; apolipoprotein D precursor and human nucleobindin 1 precursor. These two human-like proteins imply two possibilities. One is the human protein might

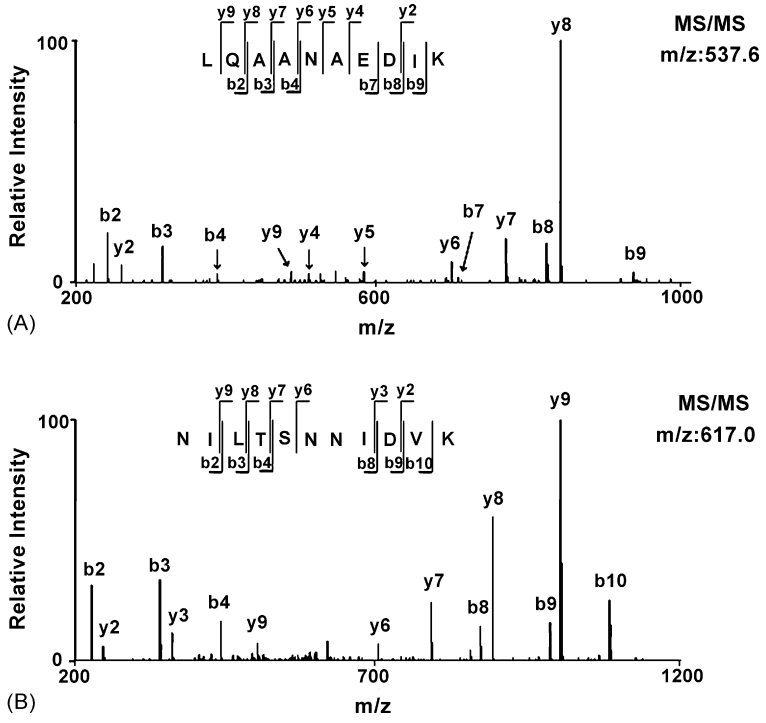

Fig. 2. An example of a nano-HPLC/ESI/MS/MS analysis. (A) Tandem mass spectrum of 537.6 *m/z*, which identified the peptide LQAANAEDIK, unique to the Nucleobindin 1 precursor from *Homo sapiens*; (B) tandem mass spectrum of 617.0 *m/z*, which identified the peptide NILTSNNIDVK, unique to the Apolipoprotein D precursor from *Homo sapiens*.

Table 1
Proteins identified in infant formula by nano-HPLC/ESI/MS/MS

Identified protein	Mass (kDa)	gi number	Representative peptide	Number of peptides
Similar to beta 2-microglobulin	11.0	28189603	HVTLEQPR	3
Chain A, alpha-lactalbumin	14.3	2392358	VGINYWLAHK	2
Complement component 3	17.0	4093220	AGQYSSDLR	2
Glycosylation-dependent cell adhesion molecule 1	17.1	27807339	NLQISNEDLSK	4
Beta-lactoglobulin	18.3	4388846	LSFNPTQLEEQCHI	8
Kappa casein	21.1	226020	HPHPHLSFMAIPPK	2
Apolipoprotein D precursor [Homo sapiens]	21.2	4502163	NILTSNNIDVK	4
Anti-testosterone antibody	24.4	432627	SPPSVTLFPPSTEELNGNK	2
Casein alpha-S1	24.5	30794348	HIQKEDVPSER	9
Beta-casein precursor	25.1	162797	FQSEEQQQTEDELQDK	2
Casein alpha-S2	26.0	27806963	LTEEEKNR	9
IgG2a heavy chain constant region	35.8	1699167	EPQVYVLDPPKEELSK	2
IgG1 heavy chain constant region	35.8	7547266	FSWFVDDVEVNTATTKPR	3
Alpha-2-HS-glycoprotein	38.4	27806751	ALGGEDVR	2
Serine proteinase inhibitor	46.2	31340900	GSTLTEILEGLK	2
Alpha1-antichymotrypsin isoform pHHK12	46.9	535509	MQDLEAK	3
Lactadherin precursor	47.4	2494285	INLFDTPLETQYVR	10
IgM heavy chain constant region	47.9	2232299	AEVLSPVVSVFVPPR	5
Adipose differentiation-related protein	49.3	27806759	EVSDGLLASSK	6
Lipoprotein lipase	50.5	1071864	AQQHYPVSAGYTK	2
PAS-4	52.7	1322373	EVVLEEGTIAFK	2
Nucleobindin 1 precursor [Homo sapiens]	53.8	2506255	LQAANAEDIK	4
MUC1 protein	58.0	18495795	VSLYFLSFR	2
Butyrophilin precursor	59.2	3183510	VAALGSDPHISMK	7
Albumin	66.1	229552	FKDLGEEHFK	15
Serum albumin precursor	69.2	418694	HPYFYAPELLYYANK	3
Transferrin	77.7	29135265	DNPQTHYYAVAVVK	4
Lactoferrin	78.0	30794292	ESPQTHYYAVAVVK	12
Lactoperoxydase	80.6	27806851	ASEQILLATAHTLLLR	4
Polymeric-immunoglobulin receptor precursor	82.4	3914346	SPIFGPEEVTSVEGR	4
Xanthine dehydrogenase	146.7	1620375	VVLGAHNLSR	15

be introduced as a contaminant during the infant formula production or the protein profiling experiment. The other is that these may be new proteins which have not been reported in the database and have a homology with human proteins (Table 1).

In conclusion, the sensitive nano-HPLC/ESI/MS/MS technique being used in this study can be employed as a valid analytical tool in the dairy industry or protein related field. For example, an adulteration in a formula can be investigated by this technique because the characterization of milk from various origins can be carried out.

Acknowledgement

This work was supported by Grant no. R01-2006-000-10593-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

References

[1] S. Rudloff, C. Kunz, J. Pediatr. Gastroenterol. Nutr. 24 (1997) 328.
[2] B. Wroblewska, M. Karamac, R. Amarowicz, A. Szymkiewicz, A. Troszynska, E. Kubicka, Int. J. Food Sci. Tech. 39 (2004) 839.
[3] S. Hellerstein, M. Berenbom, S. Dimaggio, P. Erwin, S.D. Simon, N. Wilson, Pediatr. Nephrol. 19 (2004) 780.
[4] I.M. Schmidt, I.N. Damgaard, K.A. Boisen, C. Mau, M. Chellakooty, K. Olgaard, K.M. Main, Pediatr. Nephrol. 19 (2004) 1137.
[5] E. Ferrer, A. Alegria, R. Farre, P. Abellan, F. Romero, Food Chem. 89 (2005) 639.
[6] M. Strickland, M.E. Johnson, J.R. Broadbent, Electrophoresis 22 (2001) 1510.
[7] R. Pizzano, M.A. Nicolai, R. Siciliano, C. Manzo, M.F. Mazzeo, F. Addeo, J. Agric. Food Chem. 52 (2004) 649.
[8] R. Cozzolino, S. Passalacqua, S. Salemi, P. Malvagna, E. Spina, D. Garozzo, J. Mass Spectrom. 36 (2001) 1031.
[9] P. Manglano, R. Farre, M.J. Lagarda, P. Abellan, F. Romero, G. Clemente, J. Sci. Food Agric. 84 (2004) 1126.
[10] J.L. Chavez-Servin, A.I. Castellote, M.C. Lopez-Sabater, J. Chromatogr. A 1043 (2004) 211.
[11] O. Heudi, M. Trisconi, C. Blake, J. Chromatogr. A 1022 (2004) 115.
[12] M. Galvani, M. Hamdan, P.G. Righetti, Rapid Commun. Mass Spectrom. 15 (2001) 258.
[13] 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.
[14] L.F. Marvin, V. Parisod, L.B. Fay, P.A. Guy, Electrophoresis 23 (2002) 2505.
[15] A. Trujillo, I. Casals, B. Guamis, J. Chromatogr. A 870 (2000) 371.
[16] S. Catinella, R. Seraglia, R. Marsilio, Rapid Commun. Mass Spectrom. 13 (1999) 1546.
[17] J.W. Holland, H.C. Deeth, P.F. Alewood, Proteomics 4 (2004) 743.
[18] M. Galvani, M. Hamdan, P.G. Righetti, Rapid Commun. Mass Spectrom. 14 (2000) 1889.