Proteomic study of regressions between milk yield and whey protein composition

**L.B. Larsen[1,4], A. Wedholm[2], H.S. Møller[1], A. Andrén[2]
and H. Lindmark-Månsson[3]**

[1]*University of Aarhus, Faculty of Agricultural Sciences
DK-8830 Tjele, Denmark*
[2]*Department of Food Sciences, Swedish University of Agricultural Sciences
SE-750 07 Uppsala, Sweden*
[3]*Swedish Dairy Association
SE-223 63 Lund, Sweden*

ABSTRACT

A decrease in casein number (i.e. the content of casein in relation to total protein) in milk during the last 30 years, as reported by the Swedish Dairy Association, may be related to an increasing milk production. The objective of this investigation was to study the relations between whey protein composition and daily milk yield using proteomics combined with multivariate data analysis. The results indicated that several individual minor whey proteins increased with increasing milk yield. Some of these proteins could be identified by mass spectrometry, including: fatty acid binding protein, apolipoprotein A1, lactoferrin and endopin. The technique used proved to be a useful tool for identification of significant proteins extracted from large data sets.

KEY WORDS: milk production, whey proteins, PLS, proteomics, 2D gels, MALDI-TOF

INTRODUCTION

The composition of cows' milk is of great importance for the dairy industry. Surveys carried out by the Swedish Dairy Association (Lindmark-Månsson et al., 2003) have documented variations in the casein number, both at farm-tank level and in dairy silo-tank milk. The Swedish investigation showed that the concentration of casein had decreased from 2.61%, in 1976, to 2.56%, in 1996,

---

[4] Corresponding author: e-mail: lottebach.larsen@agrsci.dk

and the concentration of whey proteins has increased from 0.73 to 0.81% during the same period (Lindmark-Månsson et al., 2003). The results implicate a change in the mean casein number from 78 to 76 during a 30-years period, and indicate a change in the proteins in the milk, either in the overall casein and whey protein content, or in a change in the mutual content of individual proteins. One reason for this change could be the increased milk yield, but the detailed background for the change in protein composition is not known.

Proteomics is a tool for the simultaneous determination of the protein composition of complex samples. One of the tools in proteomics is to combine the separation power of two-dimensional gel electrophoresis (2-DGE) with mass spectrometry (MS) for identification of separated protein spots. This method has been applied for the characterization of composition of milk powder (Galvani et al., 2000). Multivariate statistics may be used to search for systematic variation in large data sets. By partial least squares regression (PLS), the relation between the spot data and other characteristics such as treatments or quality measurements can be analysed (Jessen et al., 2002).

The objective of this study was to investigate the relations between milk production and whey protein composition using these tools.

MATERIAL AND METHODS

Milk samples

Evening whole milk samples (10 L) from 44 Danish Holstein Friesian (SDM) cows were collected at the Research Centre Foulum at the Faculty of Agricultural Sciences as described earlier (Wedholm et al., 2006). Daily milk yield in kg at the day of sampling was recorded for each cow. Protein concentration of milk samples was determined by Milkoscan FT120 (Foss Electric). After two days of cold storage (4°C), individual milk samples were skimmed twice at $3000 \times g$ for 10 min. The skimmed milk was preheated at 30°C for 30 min, and fractionated into casein and whey fractions by addition of chymosin (2 ml/L of Chy-Max Plus, 190 International Milk Clotting Units/mL, Christian Hansen A/S). The samples were incubated at 30°C for 30 min, centrifuged at $1000 \times g$ for 10 min at 5°C, and the whey fraction was recovered.

Two-dimensional gel electrophoresis and image analysis

A 2-D gel set of 44 gels, one for each individual sample in the experiment, was analysed. For analytical gels subjected to image analysis a volume of the whey protein fractions corresponding to 50 μg was applied, whereas for preparative

gels for MS analysis a volume of sample corresponding to 370 µg was applied. The protein concentrations were based on the Milkoscan determinations and a whey protein:casein ratio of 20:80. Running of the 2-DE analyses was carried out essentially as described by Lametsch and Bendixen (2001). The first dimension of protein separation was carried out in immobilized 11 cm IPG strips (pH 4-7, BioRad), with 8-16% gradient Criterion gels (BioRad) in the second dimension. A Dodeca cell (BioRad) was used, allowing the simultaneous separation of 12 gels. Analytical gels were silver stained according to Lametsch and Bendixen (2001), and the preparative gels were stained according to Shevchenko et al. (1996). The gels were analysed using ImageMaster 2D analyser software (Amersham Pharmacia). After the automatic spot detection carried out by the software, the annotated spots were manually edited. Most protein spots assigned were detected in all gels, i.e. were present in all animals, but some spots were missing in some gels. After manual editing, spots missing in some gels were considered not to be the result of technical error, and were assigned a value of zero (Grove et al., 2006). The relative spot volumes were calculated for each separate gel to overcome gel-to-gel variations in spot intensities and sample loadings.

In gel digestion, desalting and concentration of protein spots

Protein spots of significance were subjected to in-gel digestion by addition of trypsin essentially as described by Jensen et al. (1998), and the peptides were desalted and concentrated as described by Lametsch et al. (2002) by elution in 0.5 µL matrix solution (15-20 g/L of α-cyano-4-hydroxycinnamic acid in 70% acetonitrile) directly onto the MALDI target plate (Bruker Daltonics GmBH).

Partial least squares regression analysis

Partial Least Square regression 1 (PLS1) analysis was carried out by explorative data analysis using the software Unscrambler ver. 9.0 (CAMO ASA) to look for interactions between spot variations and the response Y-variable daily milk production in kg for each cow. The relative spot volumes of the whey proteins comprised the continuous X-variables in the PLS models. Standardized (centred: $\mu=0$, and normalized: 1/SD) variables and full cross validation was used. Protein spots contributing little to the model (i.e. small regression coefficients) were removed through variable selection, and based on the remaining spots (33 contributing factors) significant ($P<0.05$) regression coefficients were identified.

Identification of milk proteins by MALDI-TOF MS

Mass spectra were obtained using a Bruker Ultraflex MALDI-TOF tandem mass spectrometer (Bruker Daltonik GmBH) in reflection mode. The ion accelerating voltage was 25 kV with a delay time of 40 ns. The laser frequency was 50 Hz and 200 laser shots were accumulated for each spectrum. Peptides were identified by mass searches in the database Swiss Prot (Swiss Institute of Bioinformatics) using the MS/MS ion search programme Mascot (Matrix Science).

RESULTS

The daily milk yields of the cows in the experiment varied between 11 and 40 kg, with a mean value of 24.0 ± 7.8 kg.

Two-dimensional gel electrophoresis

A typical gel of the separated whey proteins from one of the samples is shown in Figure 1. By this analysis the majority of the whey proteins were separated. Basic whey proteins with a pI above 7 (like e.g., lactoperoxidase) could, however, not be seen at the selected conditions. The positions in this 2-DGE system of some of the major whey proteins, i.e. α-lactalbumin, β-lactoglobulin, bovine serum albumin and proteose peptone component 3 (PP3)/lactophorin were determined by MALDI-TOF MS, and are shown in Figure 1. Image analysis of the 44 whey protein 2-D gels included annotation of 444 individual protein spots.
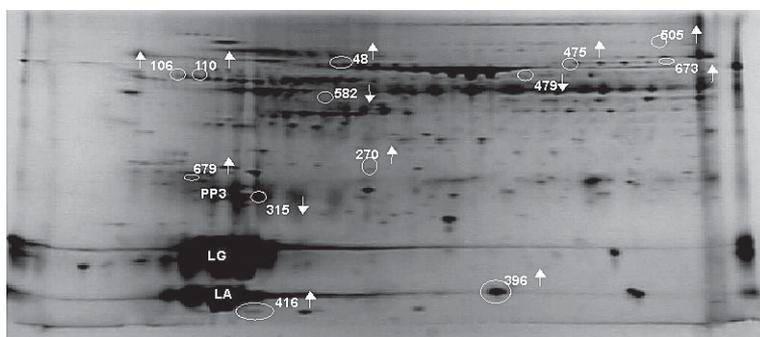

Figure 1. Two-dimensional gel electrophoresis of whey proteins from milk from a single cow in the experiment. The proteins were silver stained. The pH gradient of the first dimension is indicated. The result of the PLS1 analysis is indicated. The ID's of protein spots being positively or negatively associated (P<0.05) with daily milk yield are indicated by their number in the image analysis and by a white circle. The positions of the major whey proteins (LA - α-lactalbumin, LG -β-lactoglobulin, PP3 - proteose peptone component 3/lactophorin and BSA: bovine serum albumin) are indicated

PLS1

Result from the PLS1 analysis is presented in Figure 2. It was found that 10 proteins were significantly positively associated with increasing milk yield, whereas 3 proteins were significantly negatively associated with the same trait. These results indicated that more whey proteins were positively associated with high milk production, than negatively associated, and therefore indicate that more whey proteins increased at increasing milk yield. The positions of these spots on the 2D gels are shown in Figure 1. The value of R^2 between measured and predicted Y values in the model was 0.95, indicating a good model.
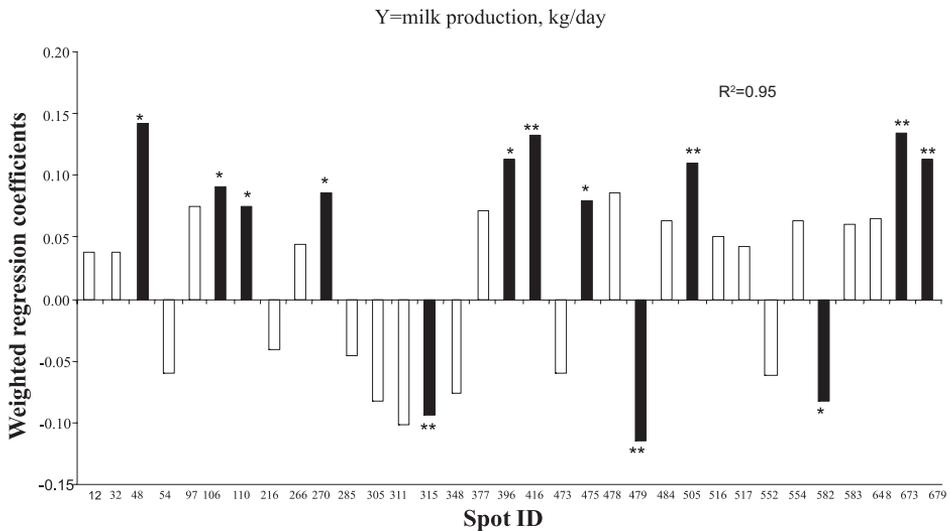

Figure 2. Results from the partial least squares regression analysis presented as the weighted regressions between x-variables of relative spot volumes and the y-variable milk production (kg/day). Black bars represent significant regression coefficients. R^2 is the correlation coefficient between measured and predicted y variables.* $P<0.05$; ** $P<0.01$

MALDI-TOF MS identifications

The significant protein spots were excised from preparative gels, and subjected to MALDI-TOF MS. By this analysis the following proteins, which were all positively associated with milk yield, were identified (with their accesskeys in the SwissProt database indicated): spot 48: lactoferrin (P24627); spots 106 and 110: endopin (Q9TTE1), spot 270: apolipoprotein A1 (P15497) and spot 506: fatty acid binding protein (FABP, P10790). It was unfortunately not possible to identify

other significant spots, probably, in part, due to the fact that many of these were minor spots, and therefore not present in substantial amounts sufficient for MS analysis. The three spots negatively associated with milk yield were present in low amounts, and these were not identified.

DISCUSSION

This investigation showed that a large amount of whey proteins could be separated by the gel-based 2-DGE system employed, with a total of 444 individual spots annotated in the image analysis. Of these proteins, many were in a similar position according to molecular mass, but with different pI, lying on a row. This confirms the presence of many post-translational modifications in the whey proteins, e.g., phosporylations and other modification not resulting in significant mass changes.

The results from the PLS1 analysis indicated, that a range of individual whey proteins were positively associated with milk production, i.e. their concentrations were positively associated with milk yield. Some of these proteins were identified by MS, and included lactoferrin, FABP, apolipoprotein A1 and endopin, in addition to other minor whey proteins not identified. FABP participates in the intracellular transport of long-chain fatty acids and their acyl-CoA esters, and in relation to milk in the uptake of preformed fatty acids from feed or body reserves. A possible explanation for the increase in FABP could be linked to an increased output of fatty acids in the milk with increasing milk production. Apolipoprotein A1 is the major protein of high-density lipoprotein (HDL), and participates in cholesterol transport, and therefore also linked to milk fat synthesis. The biological explanation for the increase in lactoferrin and endopin, a proteinase inhibitor, is not known. The increase in concentration of several individual whey proteins, found in this study, could partly explain the decrease in casein number observed along with increasing milk yield, as reported by Lindmark-Månsson et al. (2003).

It should be noted, however, that an experience from this 2-DGE and image analysis, is that the quantification of the large protein spots was difficult to achieve due to the limited dynamic range of the silver staining, and the results therefore should be interpreted according to this caution. A better dynamic range could probably be achieved by the use of fluorescent staining methods for 2-DGE gels, and therefore might result in identification of further major spots associated with milk yield. The investigation demonstrates an application of proteomics integrated with multivariate statistics for the identification of significant protein spots extracted from large datasets.

CONCLUSIONS

The results from the present study showed that several individual whey proteins were positively associated with increasing daily milk yield, whereas a smaller number was negatively associated with the same trait. The proteins that could be identified by MS included two proteins related to milk fat, FABP, apolipoprotein A1, in addition to lactoferrin and endopin.

REFERENCES

Galvani M., Hamdan M., Righetti G., 2000. Two-dimensional gel electrophoresis/matrix-assisted laser desorption/ionisation mass spectrometry of a milk powder. Rap. Commun. Mass Spec. 14, 1889-1897

Grove H., Hollung K., Uhlen A.K., Martens H., Færgestad E.M., 2006. Challenges related to analysis of protein spot volumes from two-dimensional gel electrophoresis as revealed by replicate gels. J. Proteome Res. 5, 3399-3410

Jensen O.N., Larsen M.R., Roepstorff P., 1998. Mass spectrometric identification and microcharacterization of proteins from electrophoretic gels: strategies and applications. Proteins, Suppl. 2, 74-89

Jessen F., Lametsch R., Bendixen E., Kjærsgård I.V.H., Jørgensen B.M., 2002. Extracting information from two-dimensional electrophoresis geld på partial least squares regression. Proteomics 2, 32-35

Lametsch R., Bendixen E., 2001. Proteome analysis applied to meat science: characterizing *post mortem* changes in porcine muscle. J. Agr. Food Chem. 49, 4531-4537

Lametsch R., Roepstorff P., Bendixen E., 2002. Identification of protein degradation during post-mortem storage of pig meat. Agr. Food Chem. 50, 5508-5512

Lindmark-Månsson H., Fondén R., Petterson H.-E., 2003. Composition of Swedish dairy milk. Int. Dairy J. 13, 409-425

Shevchenko A., Wilm M., Vorm O., Mann M., 1996. Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels. Anal. Chem. 68, 850-858

Wedholm A., Hallén E., Larsen L.B., Lindmark-Månsson H., Karlsson A.H., Allmere T., 2006. Comparison of milk protein composition in a Swedish and a Danish dairy heard using reversed phase HPLC. Acta Agr. Scand, Sect. A, Anim. Sci. 56, 8-15