J. Dairy Sci. 101:1–7
https://doi.org/10.3168/jds.2017-14343

Development of Fourier-transformed mid-infrared spectroscopy prediction models for major constituents of fractions of delactosated, defatted milk obtained through ultra- and nanofiltration

Marco Franzoi,*[1] Carmen L. Manuelian,* Luigi Rovigatti,† Emanuela Donati,† and Massimo De Marchi*
*Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy
†Granarolo SpA, Via Cadriano 27/2, Bologna, Italy

ABSTRACT

Milk filtration procedures are gaining relevance in the dairy industry because milk ultra- and nanofiltrates are used to increase milk processing efficiency, and as additives for products with improved nutraceutical properties. This study aimed to develop Fourier-transformed mid-infrared spectroscopy calibrations for ultra- and nanopermeate and retentate fractions of defatted and delactosated milk. A total of 154 samples from different milk fractions were collected and analyzed using reference methods to determine protein, solids-not-fat, glucose, and galactose content. The obtained values were matched with their respective Fourier-transformed mid-infrared spectroscopy spectra to develop new prediction models. Calibrations for each trait were built following 3 different approaches to get the best prediction models: (1) using the entire data set, (2) using 3 subsets based on component concentrations (level approach), and (3) using hierarchical clusters calculated with pairwise Mahalanobis distance among spectra (cluster approach). Calibrations were developed using partial least squares regression, after removing low signal-to-noise ratio wavelengths, and validated through a leave-one-out cross-validation procedure. In addition, the accuracy of the predicted values within each fraction was checked for each approach. Dividing the data set into subsets improved prediction models for each trait and for the samples in each milk fraction. Without considering milk fraction, the best improvement was observed for glucose and galactose. Glucose ratio performance deviation in cross-validation (RPD) increased from 7.42 to 11.31 and 11.06, for cluster and level approaches, respectively, whereas galactose RPD increased from 8.86 to 11.69 and 11.27 for cluster and level approaches, respectively. Considering milk fractions, the best improvement was observed for protein content, where RPD ranged from 0.08 to 6.06 for the whole data set calibration, whereas it ranged from 0.43 to 40.34 for the subset calibration approaches. Cluster and level approaches to build calibration models were comparable for samples from different fractions, suggesting that the 2 subsetting protocols should be both investigated to get the best prediction performances.
Key words: filtration, clustering, glucose, galactose

INTRODUCTION

Milk ultra- and nanofiltration are common stages in dairy processes, applied both for milk concentration and for the recovery of nutritional constituents from milk and whey (Khatkar et al., 2014; Sturaro et al., 2014; Agarwal et al., 2015; Moreno-Montoro et al., 2015). Actually, filtration procedures are preferred for the recovery of expensive or heat-sensitive components as an alternative to unit operations such as centrifugation and evaporation (Pouliot, 2008). Moreover, filtration procedures allow the development of new products and valorization of dairy industry side-products (Kumar et al., 2013). The most common filtration procedures are based on ceramic or polymeric membranes with different pore sizes, where permeate is separated from retentate applying a defined pressure to the system (Kumar et al., 2013). Based on nominal pore size of membranes, procedures can be divided into (1) microfiltration, for the separation of cells, fat globules, and bacteria; (2) UF, with most proteins found in the retentate; (3) nanofiltration, which is selective for molecules from sugars to salts; and (4) reverse osmosis, which is able to decrease water content of the product (Brans et al., 2004). Considering the increasing market for products derived from milk filtration, the validation of on-line and at-line methods for the evaluation of retentate and permeate composition is essential to develop their industrial applications.

Received December 24, 2017.
Accepted April 3, 2018.
[1] Corresponding author: marco.franzoi@studenti.unipd.it

Fourier-transformed mid-infrared spectroscopy (**FT-MIR**) analyses are a cost-effective method for at-line analysis of dairy products (Goulden, 1956; De Marchi et al., 2014; Wang et al., 2017). Moreover, FT-MIR has demonstrated its effectiveness in predicting milk technological traits, in addition to its composition (De Marchi et al., 2014; Visentin et al., 2016, 2017; McDermott et al., 2017). Nevertheless, the development of new prediction models is required when new products or new traits need to be analyzed. Few data are available regarding infrared applications for monitoring ultra- and nanofiltration processes (Barbano and Dellavalle, 1987; Solís-Oba et al., 2011). The predictability of FT-MIR calibration is strongly influenced by the number and variability of the reference data and the quality of the recorded spectra (Naes et al., 2002). Moreover, inclusion in the prediction model of sources of variability, such as the use of different standardized instruments, demonstrated an increase in calibration reliability (McKenna, 2001; Agelet and Hurburgh, 2010). This paper aimed to develop and compare 3 different FT-MIR prediction methods to determine glucose, galactose, protein, and SNF content in products derived from subsequent filtrations of defatted, delactosated milk.

MATERIALS AND METHODS

Sample Collection

Fractionation of bulk defatted milk (**DFM**) is routinely done by Granarolo S.p.A. (Bologna, Italy), which is the largest Italian dairy company (Benedet et al., 2018), at industrial scale following internal procedures to concentrate milk and to recover nutritional constituents from milk and whey. Briefly, bulk DFM is delactosated (**DLM**) and thereafter ultra- and nanofiltrated. Two fractions are obtained by UF of DLM: permeate (**PUF**) and retentate (**RUF**) of DLM. The PUF fraction is subsequently nanofiltered with a polymeric membrane, again obtaining a retentate and a permeate named **RN1** and **PN1**, respectively. The PN1 fraction is subjected to a final nanofiltration step with a polymeric membrane at very low porosity obtaining the 2 final fractions named **PN2** and **RN2** for the new permeate and retentate, respectively. A total of 154 samples of starting milk or milk fractions were collected in Granarolo S.p.A. Each sample differentiates for fractionation step, starting bulk milk or fraction, or process parameters such as pressure and temperature. For each sample, 3 aliquots were collected: one was immediately frozen and kept at −20°C for reference analysis, and the other 2 aliquots were used for FT-MIR spectra collection.

Infrared Spectra Acquisition

Each sample was analyzed using 2 FT-MIR instruments, Milkoscan FT2 and Milkoscan FT1 (Foss Electric A/S, Hillerød, Denmark) located at Granarolo S.p.A. facilities in Granarolo (Bologna, Italy) and Gioia del Colle (Bari, Italy), respectively, routinely standardized accordingly to the manufacturer. The optical standardization of the instruments, along with the standardization of the sample preparation and presentation, permits to merge in the same database spectra obtained from different instruments to develop a more robust calibration that will work on both instruments adjusting only the slope and the bias (McKenna, 2001). In the case that more than one spectrum from the same sample and same instrument was recorded, only one was randomly selected. Thus, 2 spectra were associated with each reference data point, one from each instrument. Before doing the reference analysis, spectral outliers were checked within each milk fraction using Mahalanobis distance, detecting 5 spectral outliers. The 2 spectra corresponding to the samples of the spectral outliers were removed from the data set. The final data set included 298 spectra from 149 samples. Spectra were recorded at room temperature within 24 h of sample collection. Spectral information of each sample contained 1,060 data points in the region between 5,000 and 900 cm^{-1} and was recorded as log(1/transmittance).

Reference Analysis

Ultrapure laboratory-grade (**mQ**) water was produced with Arium 611 UV (Sartorius, Monza Brianza, Italy), and all chemicals were bought from Sigma-Aldrich (Saint Louis, MO), unless otherwise indicated, at the highest available purity.

The 149 samples were analyzed for glucose and galactose content by HPLC, partially modifying the method proposed by Indyk et al. (1996). One milliliter of homogenized sample was diluted in 8.5 mL of mQ water, briefly mixed, and then 0.50 mL of Carrez 1 solution {3.60 g of $K_4[Fe(CN)_6]\cdot 3H_2O$ in 100 mL of mQ water} was added. After mixing, 0.50 mL of Carrez 2 solution (7.20 g of $ZnSO_4\cdot 7H_2O$ in 100 mL of mQ water) was added and immediately vortexed for 10 s. Thereafter, the preparation was incubated at room temperature for 30 min and centrifuged at 8,000 × *g* for 10 min at room temperature. The obtained supernatant was diluted 1:10 and passed through a 0.45 μm filter. Finally, 10 μL of the resulted sample was injected in a HPLC Spectra system (Thermo Finnigan, Waltham, MA) equipped with Aminex HPX 87C column (Bio-Rad, Hercules,