International Dairy Journal 93 (2019) 81–84

Contents lists available at ScienceDirect

International Dairy Journal

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

Short communication

Development and validation of a front-face fluorescence spectroscopy-based method to determine casein in raw milk

Yizhou B. Ma [a], Inès Birlouez-Aragon [b], Jayendra K. Amamcharla [a,*]

[a] *Department of Animal Sciences and Industry/Food Science Institute, Kansas State University, Manhattan, KS, 66506, USA*
[b] *Spectralys Innovation, 93230, Romainville, France*

ARTICLE INFO

Article history:
Received 11 September 2018
Received in revised form
6 February 2019
Accepted 14 February 2019
Available online 28 February 2019

ABSTRACT

A front-face fluorescence spectroscopy (FFFS)-based method for measuring casein in raw milk was developed and validated. Calibration samples ($n_{cal} = 20$) with different casein contents (0.36–3.7%) were prepared by mixing ultrafiltered retentate (2 × concentrate) and permeate at different ratios. A principal component regression model was developed using the calibration fluorescence spectra, showing a useful prediction power based on a residual prediction deviation of 3.1. The calibration model was independently validated using 20 raw milk samples. The FFFS-based method showed a root mean square error of prediction of 0.15% and 6.7% relative prediction error. A larger sample size should be included in the future to further validate the method and potentially implement it for routine measurement of casein levels in raw milk.

1. Introduction

Spectroscopic methods combined with chemometrics have been used to provide rapid analysis of chemical, physical, and functional properties of dairy foods and dairy processes. Prior to cheese-making, manufacturers often standardise the casein-to-fat ratio in cheese milk to optimise production yield (Guinee, O'Kennedy, & Kelly, 2006). The standard casein measurements from Association of Analytical Chemists (AOAC) involve using hazardous reagents and require trained operators. Currently, milk casein quantification methods have been developed based on Fourier transform mid-infrared spectroscopy (FT-MIR) (McDermott et al., 2016) and near-infrared spectroscopy (NIR) (Barbano & Dellavalle, 1987). The NIR method was based on the indirect measurement of casein by taking the difference between total protein and serum phase protein. The study of McDermott et al. (2016) applied partial least square regression (PLSR) to FT-MIR spectra collected for raw milk to predict casein content, as measured by a high performance liquid chromatography (HPLC)-based method.

Milk contains several intrinsic fluorophores that are suitable for fluorometric analysis. Front-face fluorescence spectroscopy (FFFS) has shown the potential to measure turbid samples directly for classification and quantification of dairy food composition. Tryptophan, as a compound of interest, is commonly used for method development, as its emission spectra can be well-correlated with protein content, acid coagulation, and heat treatment in milk (Andersen & Mortensen, 2008; Birlouez-Aragon et al., 1998). However, FFFS has not been utilised as a tool to measure casein in raw milk, but may provide an alternative casein quantification method for the cheese industry. The objective of this study was to develop and validate a quantification method for casein in raw milk using FFFS and principal component regression.

2. Materials and methods

2.1. Sample preparation for calibration set

Two batches of commingled raw milk were procured from the Kansas State University Dairy Cattle Teaching and Research Unit (Manhattan, KS, USA) on two different random days. The raw skim milk was produced using a bench-top cream separator (Motor Sich-100, Ukraine). For each batch, ultrafiltration of raw skim milk was carried out at 25 °C using a bench-top plate and frame system (SmartFlow Technologies, Apex, NC, USA) equipped with a 10-KDa cut-off polyethersulfone membrane (Hannifin Corp., Oxnard, CA, USA). Approximately, 2 × retentate was produced at a constant transmembrane pressure of 207 KPa. Subsequently, the retentate

---

* Corresponding author. Tel.: +1 785 5321221.
E-mail address: jayendra@ksu.edu (J.K. Amamcharla).

https://doi.org/10.1016/j.idairyj.2019.02.004

and permeate were combined to make ten calibration samples with different casein levels. In total, 20 calibration samples were prepared with a uniform distribution of casein content ranging from 0.36 to 3.7% (w/w) and whey protein-to-casein ratio ranging from 0.22 to 0.27.

2.2. *Reference measurement of casein*

Total nitrogen and non-protein nitrogen of retentate and permeate were analysed using AOAC (2016) standard methods (990.20 and 990.21, respectively). Due to the high protein content in the retentate, the non-casein nitrogen was measured using the method described in Zhang and Metzger (2011). Casein content was obtained from the difference between total nitrogen and non-casein nitrogen multiplied by 6.38.

2.3. *Front-face fluorescence spectroscopy measurement of casein*

Sample preparation prior to acquisition of fluorescence spectra involved a patent-pending precipitation of caseins. Preliminary experiments have revealed spectral difference between casein-precipitated (pH = 4.6) raw milk and raw milk at natural pH (data not shown). Prior to FFFS, 7 mL of calibration sample was taken in a 10-mL test tube and mixed with 0.6 mL of 10% acetic acid (Certified ACS, Fisher Scientific, Hampton, NH, USA) to ensure a pH of 4.6 ± 0.05. The mixture was vortexed for 15 s and transferred immediately into a Quartz cuvette (Starna Cells Inc., Atascadero, CA, USA), ensuring no phase separation. Tryptophan emission spectra were immediately acquired using a spectrofluorimeter fitted with a 1% attenuator (LS-55; Perkin Elmer, Waltham, MA, USA) at an excitation wavelength 280 nm and an emission scan from 300 to 440 nm. Triplicate measurements on freshly precipitated milk samples were performed at 25 °C and averaged to improve signal-to-noise ratio.

2.4. *Construction of a calibration model*

Pre-processing tools such as normalisation and derivation are commonly used to reduce drift noise and reveal spectral overlays prior to the model development. In this study, the first derivative of the tryptophan emission spectra was obtained using the Savitzky–Golay algorithm with 9-point neighbour values (Savitzky & Golay, 1964). No light-scattering effect was observed from preliminary experiments, so the scattering subtraction was not conducted as part of preprocessing. Supervised prediction models were developed from the first derivative spectra using principal component regression (PCR) and evaluated with leave-one-out cross-validation. The model performance was evaluated in terms of root mean square error of cross-validation (RMSECV; Equation (1)):

$$RMSECV\ or\ RMSEP = \sqrt{\frac{\sum_{j=1}^{n}(y_{j}-\widehat{y}_{j})^{2}}{n}} \tag{1}$$

where y is the predicted value from cross-validation or independent validation of the j^{th} sample, $\widehat{y}$ is the reference value of the j^{th} sample, and n is sample size.

Residual prediction deviation (RPD) was calculated from RMSECV divided by the sample standard deviation and used as a parameter to estimate the model prediction power. Statistical model building and evaluation were conducted using RStudio (version 1.1.442; RStudio, Boston, MA, USA) with the Caret package (Kuhn, 2008).

2.5. *Independent validation of the model*

To validate the PCR-based calibration model, raw milk samples were randomly collected from 20 individual cows ($n_{val} = 20$) from the same dairy farm, and casein content was measured by the reference method as described above. Subsequently, tryptophan FFFS spectra were also collected from the validation samples as described above and preprocessed following the same procedure used for calibration samples. The accuracy of the FFFS-based method was evaluated in terms of the mean bias and root mean square error of prediction (RMSEP; Equation (1)).

Relative prediction error (RPE) was also calculated dividing the RMSEP by the average casein content obtained from the reference method (Amamcharla & Panigrahi, 2010). The range error ratio (RER) was calculated by dividing the range of reference casein content by the RMSEP. Graphical representation was prepared by plotting the difference and the averaged casein content measured by the reference and FFFS-based methods. This difference plot approach allowed the comparison of bias and error between the two methods and determine outliers from the measurements (Twomey, 2006).

3. Results and discussion

3.1. *Tryptophan fluorescence spectra of casein in calibration samples*

Fig. 1A shows representative tryptophan fluorescence spectra of low (0.36%) and high (3.7%) casein samples from the calibration set. The remaining calibration samples also produced similar tryptophan spectra ranged between the low and high-casein samples (data not shown). The tryptophan emission maximum was observed at 338.5 ± 1 nm for all the samples. Shaikh and O'Donnell (2017) reviewed tryptophan fluorescence in milk and reported emission maxima around 340 nm with excitation of 290 nm. Fig. 1B shows the first derivative spectra of low and high-casein samples from the calibration set. More spectral variations were observed around emission wavelengths of 313 and 363 nm. The derivation also smoothed the spectra and unified the fluorescence intensity scale and was subsequently used for calibration model development.

3.2. *Calibration model development*

Fig. 2 shows the scatter plot between the casein content obtained from the FFFS-based method and the reference method. For the reference method, the duplicated casein measurements showed an average coefficient of variations of 3.5%. The optimum number of principal components ($n_{pc} = 2$) was chosen based on the lowest RMSECV (= 0.35) of the calibration model. The coefficient of determination (R^2) was found to be 0.89 and the slope for the predicted vs reference casein values was 0.91, showing an increasing negative bias towards the high-casein samples. The calibration samples were evenly distributed with equal sample size for different casein content. This approach was used to minimise the "Dunne effect" which leads the overestimation of the lower range and underestimation of the higher range in calibrations (Williams & Norris, 1987). The intercept of the scatter plot (Fig. 2) was 0.18, indicating that it may result in more bias for low-casein samples. Physical differences in precipitation behaviour of caseins between the high- and low-casein samples could have contributed to the spectral variations. The prediction power of the model was evaluated by RPD, and, according to Williams and Norris (1987), an RPD of 3.1 for the model indicates a useful prediction.

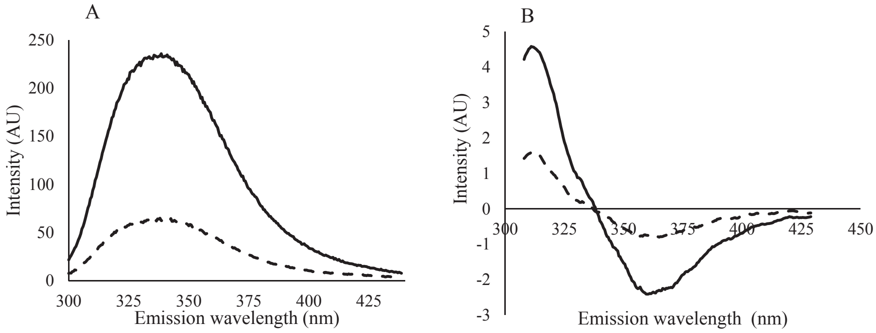

Fig. 1. Representative front-face tryptophan fluorescence spectra (excitation = 280 nm; intensity in arbitrary units, AU) of low (dashed line) and high-casein (solid line) calibration samples (A) and first derivative of spectra (B).

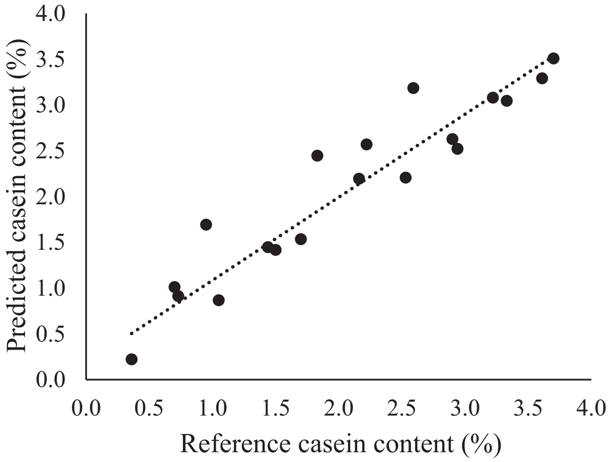

Fig. 2. Principal component regression calibration model: reference versus predicted casein content (%). Number of principal components was 2, coefficient of determination was 0.89, slope was 0.91, intercept was 0.18, root mean square error of cross-validation was 0.35, residual prediction deviation was 3.1.

3.3. Independent validation

According to the difference plot shown in Fig. 3, 19 out of 20 validation samples fall within the 95% confidence interval, with only 1 sample falling beyond the lower confidence interval. This sample had a reference casein value of 2.34% (w/w) and predicted value of 2.01% (w/w). Due to the small cow-to-cow variation in the casein content among validation samples, R^2 may not fully evaluate the model performance (Altman & Bland, 1983). In this study, the mean casein content bias between the reference and FFFS-based methods was 0.01%, indicating a slight underestimation of the FFFS-based method. The RMSEP was 0.15% and the corresponding RPE of the FFFS-based method was 6.7%.

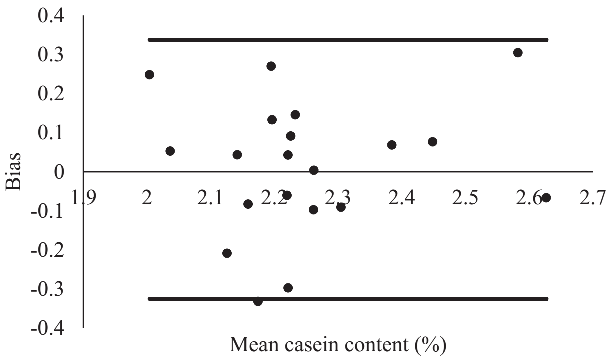

Fig. 3. Difference plot showing the bias between reference and predicted casein contents (n_{val} = 20) against the mean casein content (%). The black bars indicate the 95% confidence intervals. Mean bias: 0.01%, root mean square error of prediction: 0.15%, relative prediction errors: 6.7%, range error ratio: 8.4.

McDermott et al. (2016) conducted FT-MIR measurement of casein in milk, and the external validation of the casein prediction resulted in an RPE of 13% using HPLC-based reference values. Hewavitharana and van Brakel (1997) developed a FT-IR method for casein in raw milk using PLSR and PCR and results showed an RPE between 2.3 and 4.3%. In general, an RPE less than 5% is considered acceptable in most applications, so the FFFS-based method can still be improved to fulfil the accuracy requirement. For the present study, the RER of the validation was 8.4. According to Williams and Norris (1987), RER describes the practicality of such models, and a value above 4.0 is acceptable for sample screening purposes. Therefore, the FFFS-based method developed in this study still has practical use to serve the dairy industry.

4. Conclusions

This study validated a FFFS-based method to quantify casein in raw milk. Tryptophan front-face fluorescence spectra of casein suspension obtained from acidifying raw milk to pH 4.6 was used to develop a calibration model. The model showed practical prediction power and an independent validation confirmed the industry feasibility. For future studies, the FFFS-based method should be validated with a larger sample size including more sources of variation such as season, feeding, and cow breed. The method will be further adapted to a portable fluorescence device and potentially implemented for routine measurement of casein in raw milk.

Acknowledgement

This project is Kansas State Research and Extension contribution number 19-046-J.

References

Altman, D. G., & Bland, J. M. (1983). Measurement in medicine: The analysis of method comparison studies. *Journal of the Royal Statistical Society. Series D (The Statistician), 32*, 307–317.

Amamcharla, J. K., & Panigrahi, S. (2010). Simultaneous prediction of acetic acid/ethanol concentrations in their binary mixtures using metalloporphyrin based opto-electronic nose for meat safety applications. *Sensing and Instrumentation for Food Quality and Safety, 4*, 51–60.

Andersen, C. M., & Mortensen, G. (2008). Fluorescence spectroscopy: A rapid tool for analyzing dairy products. *Journal of Agricultural and Food Chemistry, 56*, 720–729.

AOAC. (2016). *Official methods of analysis (20th ed.). Association of Official Analytical Chemists*. Gaithersburg, MD, USA: AOAC International.

Barbano, D. M., & Dellavalle, M. E. (1987). Rapid method for determination of milk casein content by infrared analysis. *Journal of Dairy Science, 70*, 1524–1528.

Birlouez-Aragon, I., Nicolas, M., Metais, A., Marchond, N., Grenier, J., & Calvo, D. (1998). A rapid fluorimetric method to estimate the heat treatment of liquid milk. *International Dairy Journal, 8*, 771–777.

Guinee, T. P., O'Kennedy, B. T., & Kelly, P. M. (2006). Effect of milk protein standardization using different methods on the composition and yields of cheddar cheese. *Journal of Dairy Science, 89*, 468–482.

Hewavitharana, A. K., & van Brakel, B. (1997). Fourier transform infrared spectrometric method for the rapid determination of casein in raw milk. *Analyst, 122*, 701–704.

Kuhn, M. (2008). Caret package. *Journal of Statistical Software, 28*, 1–26.

McDermott, A., Visentin, G., De Marchi, M., Berry, D. P., Fenelon, M. A., O'Connor, P. M., et al. (2016). Prediction of individual milk proteins including free amino acids in bovine milk using mid-infrared spectroscopy and their correlations with milk processing characteristics. *Journal of Dairy Science, 99*, 3171–3182.

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*, 1627–1639.

Shaikh, S., & O'Donnell, C. (2017). Applications of fluorescence spectroscopy in dairy processing: A review. *Current Opinion in Food Science, 17*, 16–24.

Twomey, P. J. (2006). How to use difference plots in quantitative method comparison studies. *Annals of Clinical Biochemistry, 43*, 124–129.

Williams, P. C., & Norris, K. (1987). *Near-infrared technology in the agricultural and food industry*. St. Paul, MN, USA: American Association of Cereal Chemists, Inc.

Zhang, H., & Metzger, L. E. (2011). Noncasein nitrogen analysis of ultrafiltration and microfiltration retentate. *Journal of Dairy Science, 94*, 2118–2125.