

Prediction of α -Lactalbumin and β -Lactoglobulin Composition of Aqueous Whey Solutions Using Fourier Transform Mid-Infrared Spectroscopy and Near-Infrared Spectroscopy

Margherita Tonolini¹ , Klavs Martin Sørensen¹,
Peter B. Skou² , Colin Ray², and Søren Balling Engelsen¹

Applied Spectroscopy
2021, Vol. 75(6) 718–727
© The Author(s) 2021



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0003702820979747
journals.sagepub.com/home/asp



Abstract

Characterization and quantification of individual whey proteins are of crucial importance to many industrial dairy processes. Labor intensive wet-chemical methods are still being used for this purpose, but a rapid quantification method for individual whey proteins is highly desired. This work investigate the utility of Fourier transform mid-infrared spectroscopy and Fourier transform near-infrared spectroscopy for rapid quantification of the two main whey proteins (β -lactoglobulin and α -lactalbumin) in complex aqueous whey solutions simulating production process streams. MIR and NIR spectra obtained on whey samples with known and varying amounts of the proteins of interest and are used to develop partial least squares prediction models. Selection of informative wavelength regions allowed for prediction of β -lactoglobulin and α -lactalbumin concentrations with very high precision and accuracy (root mean square error of cross-validation, or RMSECV, of 0.6% and R^2 of 0.99 for NIR), demonstrating the potential of being implemented for rapid in-line monitoring of protein composition in industrial whey streams. Two-dimensional MIR-NIR correlation spectroscopy is used to identify the most informative parts of the NIR spectra in relation to protein secondary structure. In addition multivariate curve resolution is applied to the MIR data to resolve mixture spectra and to elucidate the spectral ranges that were most useful in distinguishing between the two whey proteins. The study concludes that NIR spectroscopy has potential for accurate in-line protein quantification and overall secondary protein structure quantification which open new possibilities for in-line industrial applications.

Keywords

Multivariate curve resolution, 2D correlation spectroscopy, multivariate analysis, whey protein, protein secondary structure, near-infrared spectroscopy, mid-infrared spectroscopy

Date received: 26 June 2020; accepted: 17 October 2020

Introduction

Alpha-lactalbumin (α -La) and beta-lactoglobulin (β -Lg) are the two most abundant whey proteins found in milk. These proteins can be individually isolated from dairy streams and they have interesting nutritional properties. Quantification of individual whey protein fractions in whey streams is useful to monitor and improve the efficacy of whey fractionation processes. Time-consuming wet chemistry methods, such as reverse phase high performance liquid chromatography¹ are still widely used for the individual quantification of α -La and β -Lg. Although several spectroscopic techniques have been tested for their applicability, an efficient and reliable rapid method for quantification of

individual whey proteins in dairy streams that potentially can be utilized in- or at-line has yet to be established. In this work, mid-infrared (MIR) and near infrared (NIR) spectroscopy were examined as rapid methods for the analysis of the two main whey proteins, α -La and β -Lg, in aqueous whey solutions. The potential application of the presented

¹Chemometrics and Analytical Technology, Department of Food Science, Faculty of Life Sciences, University of Copenhagen, Copenhagen, Denmark
²Arla Foods Ingredients Group P/S, Videbæk, Denmark

Corresponding author:

Margherita Tonolini, University of Copenhagen, Rolighedsvej 26, Frederiksberg C, Copenhagen DK-1958, Denmark.
Email: margherita.tonolini@food.ku.dk

method as in- or at-line measurements of the two target proteins in whey stream is demonstrated. Alpha-lactalbumin and beta-lactoglobulin have different native secondary structures. Based on X-ray scattering, the secondary structure of β -Lg is made of around 15% α -helix and 50% β -sheet,² while the α -La is folded mainly as α -helix (47%) with around 6% β -sheet.³ Utilizing these structural differences, the two proteins can be distinguished and quantified in various dairy matrices using vibrational spectroscopic methods. For structural analysis of proteins, the amide I band (1600–1700 cm^{-1}) and the amide II band (1500–1600 cm^{-1}) of the mid-infrared spectral range have been shown to be particularly sensitive to the distribution of the proteins secondary structure.⁴ MIR spectroscopy is therefore a valid option for investigating and quantifying whey protein secondary structure and accordingly MIR spectroscopy has been used to quantify individual whey proteins with moderate success. Sturaro et al.¹ developed a mid-infrared spectroscopy-based prediction model for the rapid quantification of protein in sweet whey, using reversed phase high performance liquid chromatography (HPLC) as a reference method. In another study, the precision of the regression models obtained from MIR spectra was found to be greatly impacted by the complexity of the sample matrix, and samples with chemically complex matrix resulted in weaker regression models.⁵ Schwaighofer et al.⁶ applied external cavity-quantum cascade laser mid-infrared spectroscopy as a rapid method for protein analysis of bovine milk and quantitation of total protein content as well as content of individual proteins (casein, β -Lg, α -La), showing excellent predictions even for protein concentrations less than 1%. Although NIR spectroscopy has been used for monitoring total protein content of various dairy products,^{7,8} there are to our knowledge no studies available demonstrating that NIR can be used to distinguish between and quantify β -Lg and α -La. However, Izutsu et al.⁹ suggested that proteins with different secondary structure compositions give different NIR spectra. The NIR spectra of proteins in aqueous solutions suggest that some bands are indicative for α -helix (4090, 4365–4370, 4615, and 5755 cm^{-1}) and others for β -sheet (4060, 4405, 4525–4540, 4865, and 5915–5925 cm^{-1}) structures. The aim of this work is to investigate if NIR spectroscopy can be considered suitable for quantifying individual whey proteins at-line and potentially in-line. An aqueous model whey matrix was chosen as the subject of this study in order to mimic the chemical composition of whey streams. Principal component analysis (PCA) was used to visualize how samples were varying with respect to each other,¹⁰ and partial least squares (PLS) regression¹¹ was used to investigate how well the spectroscopic measurements can be utilized to predict the α -La and β -Lg concentrations in the samples. Finally, this work explores if the NIR region can provide quantitative information about protein secondary structure. This will be investigated by studying the deconvolution of MIR

spectra by MCR¹² and by applying two-dimensional (2D) MIR-NIR correlation spectroscopy¹³ to the whey samples. The 2D correlation spectroscopy analysis on MIR and NIR spectra illustrate how the fundamental vibrations in the MIR region covary with the overtones and combinations bands in the NIR region.¹⁴ In this study, 2D correlations are used to map how the protein secondary structure variations found in the MIR region translate to the NIR spectra. MCR is a method used to solve the mixture analysis problem. The result of MCR provides chemically meaningful information on the contributions of the pure compounds involved in the system.¹² In spectroscopy, MCR models provide information about the spectra of the individual sample constituents. In this study, MCR was used to obtain chemically interpretable profiles of the two main whey proteins.¹⁵

Experimental

Preparation of Aqueous Whey Solutions

The α -La and β -Lg powders used in this study both contained about 93% of protein, 1.8% ash, 0.1% lactose, and 4% water (powders were provided by Arla Foods Ingredients). The secondary structures present in the powders are considered native, and throughout this experiment we assume that protein secondary structure remains unchanged during whey fractionation processes. These powders were added as-is to a model whey solution for the experiments conducted. The whey solution was prepared using fresh whole milk (3% fat, obtained on the day of the experiment in a local supermarket), which was heated to 40 °C and coagulated with rennet (CHY-MAX, Chr. Hansen, Hørsholm, Denmark). After coagulation (6 min) the curd was cut, agitated, and gravity-filtered through a Whatman 40 filter. The aqueous phase was retained and used to prepare the samples for this study. This authentic whey produced by lab scale renneting acts as a complex matrix to which α -La and β -Lg were added in increasing concentrations. We estimate that this model whey is complex as it contains proteins, lactose, and minerals. The protein concentrations in the whey background were not included in the protein concentration calculations, since they are considered constant across all samples. Two sets of thirteen samples were produced with spiked α -La and β -Lg concentrates, respectively, in different proportions (0%, 0.1%, 0.25%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, and 10% w/w). Subsequently, 13 further mixtures of the α -La and β -Lg whey dilutions were prepared (Fig. 1).

Spectroscopic Measurements

Fourier Transform Mid-Infrared Spectroscopy

The MIR spectroscopy was performed using an MB-Series Spectrometer MB-104/MB-104PH (Bomem) equipped with

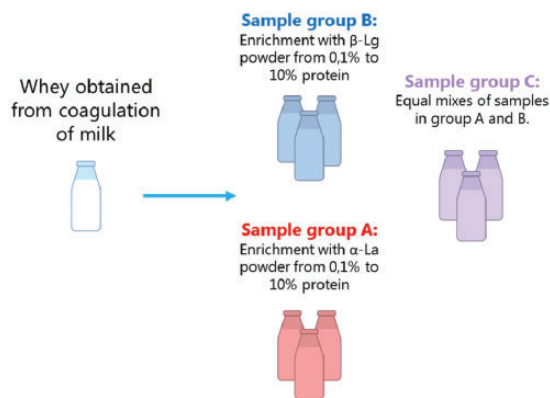


Figure 1. Study design.

a liquid attenuated total reflection (ATR) interface, 12 bounces (Gateway ATR, SpectraTech Inc.). Sixty-four scans were recorded and averaged for each sample at 8 cm^{-1} resolution in the range $4000\text{--}9000\text{ cm}^{-1}$. The ATR crystal was cleaned with water and ethanol between measurements, and the cleaned empty crystal was used as background measurement. Every sample was measured twice, along with a background, and all measurements were made in random order within 12 h.

FT-NIR Spectroscopy

Near-infrared spectra were collected using an MPA II FT-NIR Instrument (BRUKER Optics). Spectra were acquired in transmittance mode with a 1 mm pathlength cuvette. Spectra were measured in the range $4000\text{--}10000\text{ cm}^{-1}$ ($500\text{--}2500\text{ nm}$), a total of 64 scans were averaged for each sample and the spectral resolution was 8 cm^{-1} . Every sample was measured twice, and all samples measured in randomized order within 12 h. The cuvette was cleaned with a cuvette cleaning agent (Hellmanex III, Sigma-Aldrich) between measurements, and was used as the transmission background.

Data Processing

The spectra were analyzed in Matlab (R2019b, The Mathworks, Inc.) with the PLS toolbox 7.5 (Eigenvector Research, Inc.), supplemented by in-house algorithms. All data for PCA and PLS models were mean centered prior to analysis.

PCA and PLS Regression

Principal component analysis was used to visualize how samples were varying with respect to each other. PLS regression was used to investigate how well the spectroscopic measurements predicted the α -La and β -Lg

concentrations in the samples. With respect to MIR data, only the region corresponding to the amide I and II band regions, $1450\text{--}1700\text{ cm}^{-1}$, as for instance demonstrated by Barth⁴ were selected. Prior to PCA and PLS analysis, several preprocessing methods were tested on MIR data (not shown), and it was found that multiplicative scatter correction (MSC)¹⁶ and mean centering yielded models of the highest quality for interpretation. For NIR data, PCA and PLS regression models were focused solely on the NIR region $4250\text{--}4800\text{ cm}^{-1}$ in accordance with previous studies of protein secondary structure with NIR,¹⁷ which proved to give the best results for quantifying α -La and β -Lg. Prior to PCA and PLS analysis, several preprocessing methods (not shown here) were tested on NIR data.¹⁸ A Savitzky–Golay¹⁹ second derivative filter of second-order with a width of 21 points yielded models of the highest quality for interpretation. Leave-one-sample concentration-out (all replicates of the same concentration) cross-validation was applied for PLS models.

2D Correlation Spectroscopy

The 2D correlation method used in this work is similar to the one reported by Ringsted et al.²⁰ In the present study, a 2D correlation map was produced by calculating the correlation between all combination of variables from the MIR and NIR datasets, respectively. The resulting map is presented using color gradients to indicate correlation strength (heat map). The NIR spectra used for the 2D correlation were preprocessed with Savitzky–Golay second derivative as for the PLS model, using the wavelength range $4000\text{--}9000\text{ cm}^{-1}$. The MIR spectra were preprocessed with MSC and the spectral range $1000\text{--}1700\text{ cm}^{-1}$ was used in the correlation map.

Multivariate Curve Resolution

For estimating the MCR models, non-negativity constraints were applied to both MIR spectra and concentrations modes. The resolved pure spectra were normalized to unit area to enable direct comparison. For the MCR analysis, the raw non-preprocessed spectra were used.

Results and Discussion

Assignment and Exploration of the MIR Data

Exploratory analysis of the MIR spectra (Fig. 2a) shows a noisy region caused by the strong absorbances of the O–H stretches at $3000\text{--}3600\text{ cm}^{-1}$. A small peak is visible around 1080 cm^{-1} , which is commonly associated with C–OH vibrations of carbohydrates (such as lactose). The most important region of the spectra, for the purpose of this study, is the one associated to the characteristic MIR bands of proteins and peptides. These includes the amide

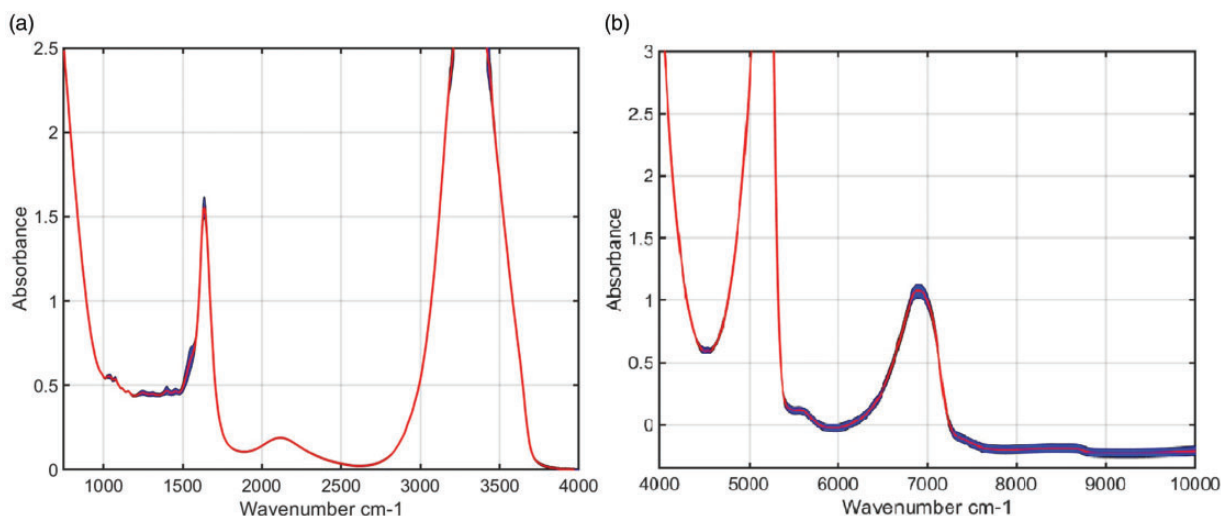


Figure 2. (a) Average MIR spectrum of all samples (in red) and \pm standard deviation of the spectra (in blue). The absorbance has been truncated at 2.5. (b) Average NIR spectrum of all samples (in red) and \pm one standard deviation of the spectra (in blue). The absorbance has been truncated at three.

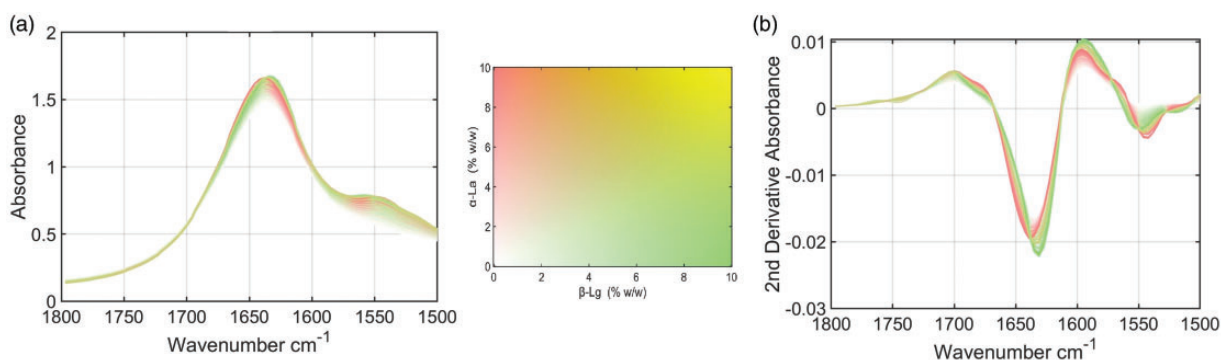


Figure 3. (a) Spectra of the selected MIR region. (b) Spectra of the selected MIR region preprocessed with Savitzky–Golay second derivative. The spectra are colored according to the calculated secondary structure composition. The 2D color-coding is shown in the square color bar in the center.

I band from 1600 to 1690 cm^{-1} , the amide II band from 1480 cm^{-1} to 1575 cm^{-1} , and the amide III band from 1229 to 1301 cm^{-1} .²¹ The proteins of interest vary primarily in their α -helix and β -sheet compositions. The amide I band of polypeptides is known to be sensitive to the secondary structure and β -sheet structures can be found around 1624 – 1642 cm^{-1} , while α -helix can be found at 1656 – 1663 cm^{-1} . Some information about the secondary protein structure distribution can also be found in the amide II band, with β -sheets information at 1530 cm^{-1} and α -helix at 1545 cm^{-1} .²² Figures 3a and 3b show how the position of the α -helix and β -sheet peaks changes as a function of secondary structure found in the samples. Figure 4a shows a PCA score plot for PC1 versus PC2 for the MIR data. Over 99% of the spectral variance is explained by the two principal components. PC1 variance seems to primarily explain the total protein concentration in the samples, as PC1 scores increases concomitantly with the total protein

concentration. The second principal component (PC2) seems to exhibit protein specific information, as the samples from the three sample groups (α -La, β -Lg and mixes) can be easily distinguished. The two protein samples groups (α -La and β -Lg pure samples) yield a comfortable margin of separation. In contrast, the scores of the samples belonging to the mixes group are systematically located between pure α -La and β -Lg samples. The respective PCA loadings reported in Fig. 4c show peaks at 1525 , 1560 , 1620 , and 1660 cm^{-1} , which confirms that there is a correlation between amide I and II bands and the principal components selected.

Assignment and Exploration of the Near-Infrared Data

Inspection of the NIR spectra shows, at 4000 cm^{-1} , the presence of the shoulder with declining absorbance (data not shown) which is due to the O–H stretching vibrations

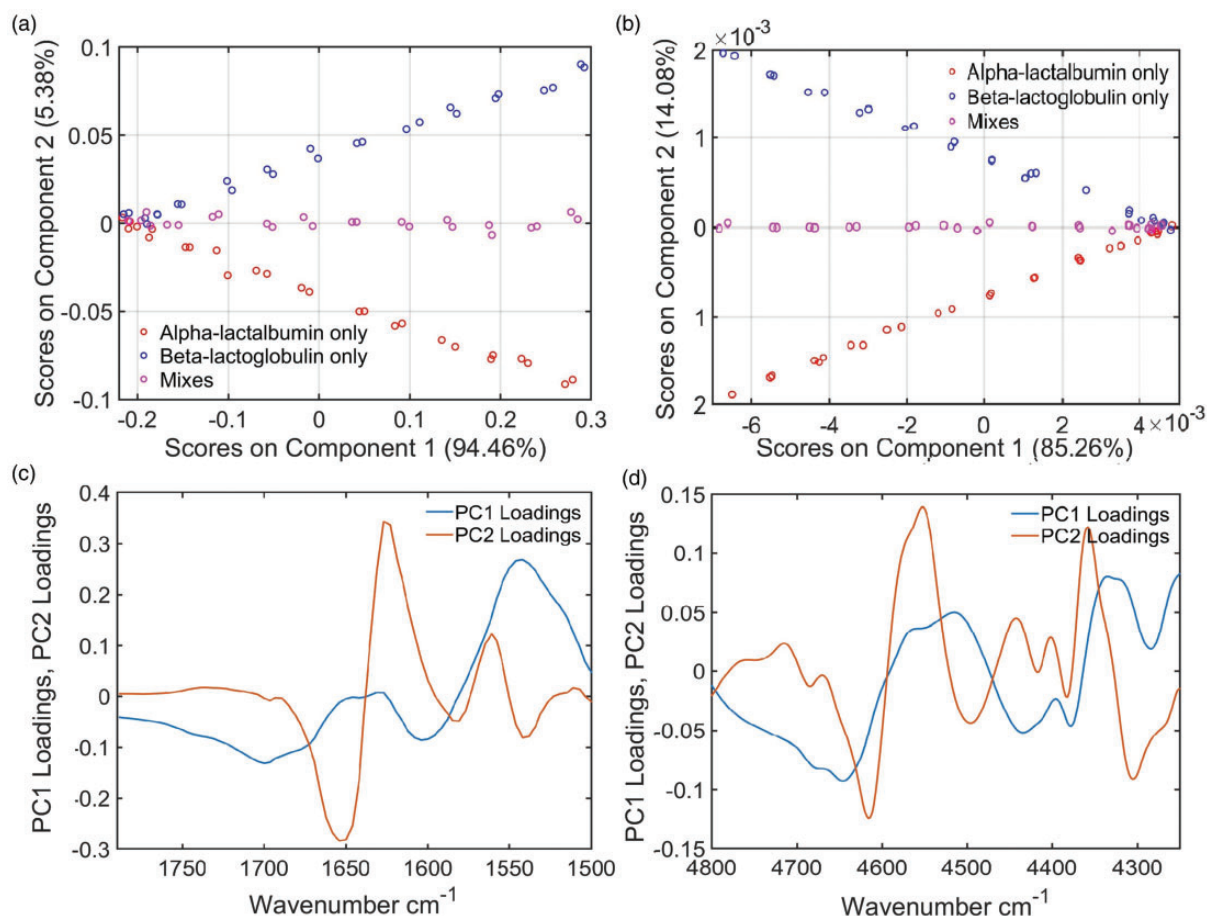


Figure 4. (a) PCA Scores plot for PC1 versus PC2 for the MIR data. The number in parenthesis indicates the explained variance for each component. (b) Scores plot for PC1 versus PC2 for the NIR data. The associated PCA loadings are shown in (c) and (d), respectively.

peaking around 3400 cm^{-1} in the MIR region. Then, we observe a strong absorption peak around 5300 cm^{-1} due to a water combination tone (Fig. 2b). A strong peak is centered around 6800 cm^{-1} , which is the first overtone of the O–H stretching vibration. The spectral window in the peak valley $4250\text{--}4800\text{ cm}^{-1}$ showed to provide the best results for quantifying α -La and β -Lg. This selected region is in accordance with literature data on studying protein secondary structure with NIR.⁹ Figure 4b shows the result of a PCA on the NIR data. Over 99% of the total variance of the spectra is explained by two principal components. Similar to the PCA model of the MIR spectra, the PC1 variance primarily explains the total protein concentration of samples, while PC2 explains protein specific information. The respective PCA loadings shown in Fig. 4d show peaks at 4300 , 4350 , 4500 , and 4550 cm^{-1} which is consistent with the secondary structure bands described by Izutsu et al.⁹ The PCA of the NIR spectra also shows significant correlation with the quality parameters, concentration of α -La and β -Lg.

Quantitative Protein Analysis Employing PLS Regression

PLS regression models were established for each of the two proteins and for each of the spectroscopic techniques. The data set was divided into a calibration set (75% of the samples) and a validation set (25% of the samples). Based on the cross-validated and the test-set prediction errors, three components were chosen for all four models.

The root mean square error of calibration (RMSEC) and of cross-validation (RMSECV) for the PLS model using the NIR data were 0.06% w/w and 0.06% w/w, and 0.07% w/w and 0.06% w/w of β -Lg and α -La, respectively, while the R^2 was over 0.99 for both proteins. While the PLS model on the MIR data resulted in RMSEC and RMSECV values were 0.13% w/w and 0.17% w/w and 0.13% w/w and 0.18% w/w of β -Lg and α -La, respectively. R^2 was better than 0.99 for both analytes (see Table I and Fig. 5). Moreover, the models for α -La predictions showed to be successful in predicting concentrations next to 0% w/w for the sample group

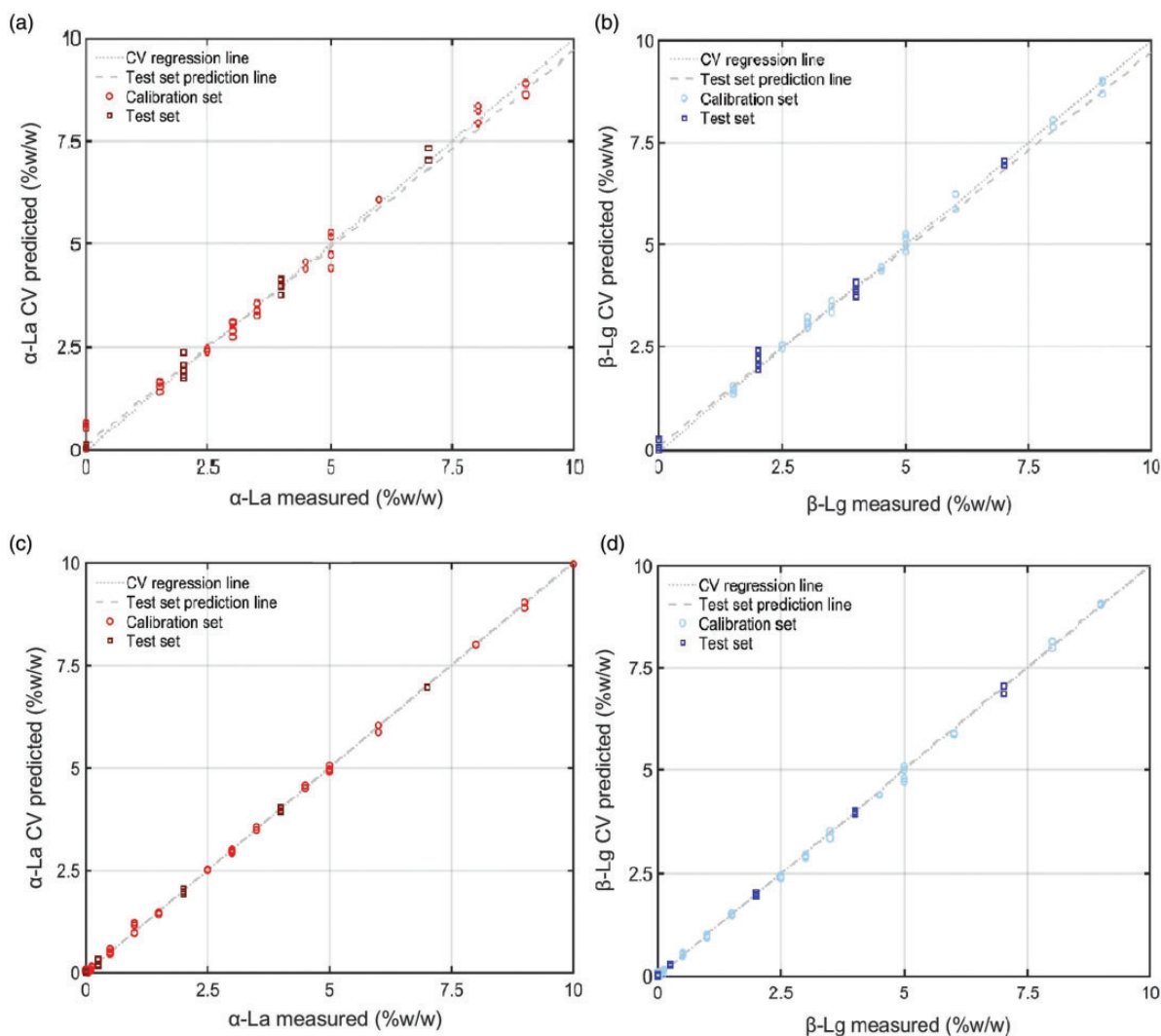


Figure 5. (a) and (b): Predicted vs. measured α -La and β -Lg concentrations from a two-component PLS model on MIR spectra. (c) and (d): Predicted versus measured α -La and β -Lg concentrations from a two-component PLS model on NIR spectra.

containing only β -Lg and vice versa (data not shown). It is remarkable that the prediction errors for the MIR data models are higher than for the NIR data models for both analytes. This may be explained by interference with other unrelated sample constituents in the MIR region, compounds in the background whey solutions, and/or by a better measurement reproducibility with FT-NIR (quartz cuvette versus ATR). Moreover, the prediction errors are also found to be slightly higher for α -La than for β -Lg. This difference is considered non-significant.

Two-Dimensional Correlation Spectroscopy on Near- and Mid-Infrared Spectra

Figure 6 shows the resulting covariance heatmap resulting from the 2D correlation analysis. The figure shows that high correlations are found between the MIR amide II region and

the NIR region between 4000 and 5000 cm^{-1} . This is a surprising result since the amide II region is rarely used for quantifying individual proteins. In contrast, the same strong correlation is not found between the amide I region and the NIR spectra, most likely due to strong overlapping water absorbance at 1640 cm^{-1} . Several of the high correlated areas are in good accordance with the findings of Izutsu et al.⁹ The regions between 4300 and 4400 cm^{-1} , 4600 and 4800 cm^{-1} , and 5800 and 5900 cm^{-1} show particularly strong correlations with the MIR amide II region. Some of these correlations can be explained by the presence of the second overtone of C–H deformations in the proteins, N–H symmetrical stretching overtones, and amide I, II, and III overtones in the 4600–4800 cm^{-1} range. There are also strong correlations to the amide I band in the NIR ranges from 5100 to 5800 cm^{-1} and 6500 to 6700 cm^{-1} which may be assigned to the water

Table I. Summary of results from the four PLS models.

Preprocessing method	NIR data		MIR data	
	Savitzky–Golay second derivative		MSC	
No. of PCs	3		3	
RMSEC	0.06	0.06	0.17	0.13
RMSECV	0.06	0.07	0.18	0.13
R ²	0.99	0.99	0.99	0.99

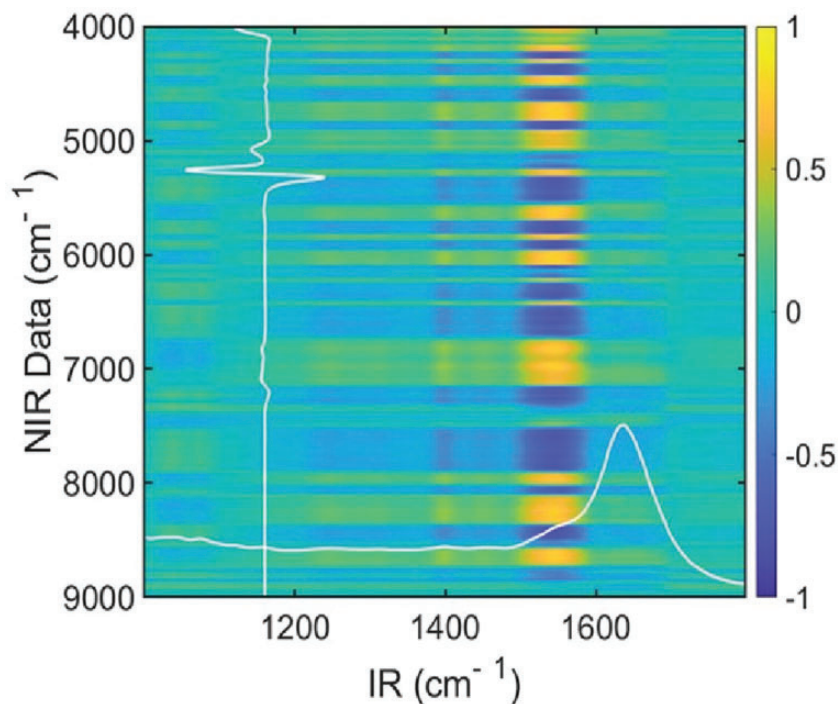


Figure 6. 2D NIR-IR correlation map. The absorbance bands which are positively correlated are shown with yellow color and negative correlations with blue color. The average spectrum of MIR is shown horizontally and the average second derivative NIR spectrum is shown vertically.

combination band and the first overtone of the N–H stretching in proteins.

Multivariate Curve Resolution

The MCR modeling of the MIR spectra resulted in a three-component model. The first component is attributed to water signals in the aqueous whey solution (Fig. 7). The spectra of the second and third components seem to correspond to the pure spectra of β -sheet structure and α -helix structure, respectively.

The estimate of pure α -helix spectra shows maxima at 1645 cm^{-1} , 1470 cm^{-1} , 1545 cm^{-1} , and 1580 cm^{-1} . The pure β -sheet spectrum (MCR loading) features maxima at 1630 cm^{-1} , 1560 cm^{-1} , and 1530 cm^{-1} . Accordingly, secondary structure features seem to be present both in the

amide I region and in the amide II region. While some of the shapes and peak positions of the pure spectral profiles obtained by the MCR are in agreement with the reported vibration bands of various protein structural elements,^{22,21} other characteristics, such as the peaks at 1580 cm^{-1} , 1560 cm^{-1} , and 1470 cm^{-1} are new and not previously reported in literature. In order to seek to validate these MCR results, the theoretical concentration of secondary structure elements (α -helix and β -sheet) was calculated from literature as described by Creamer et al.² and Chandra et al.³ and plotted against the scores of the MCR model in order to investigate if the individual protein structures can be correlated to the MCR components (Fig. 8).

Figure 8a shows an almost perfect linear relationship between MCR component I and the water concentration

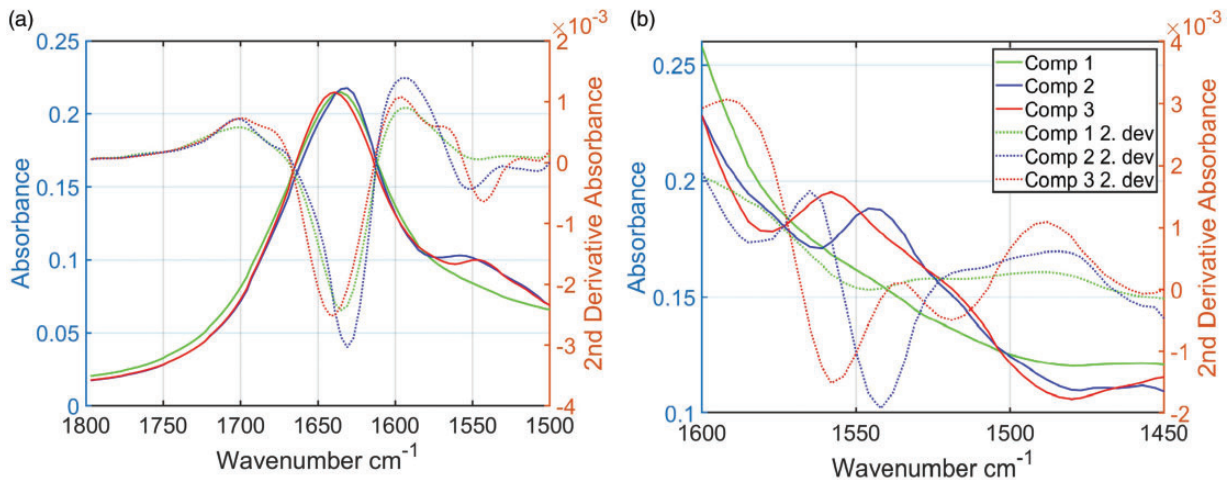


Figure 7. The MCR model of the MIR spectra in the region containing the amide I and amide II bands. (a) The MCR resolved pure spectra. (b) A zoom on the amide II band region. The stipulated lines show the corresponding second derivative spectra for improved interpretation.

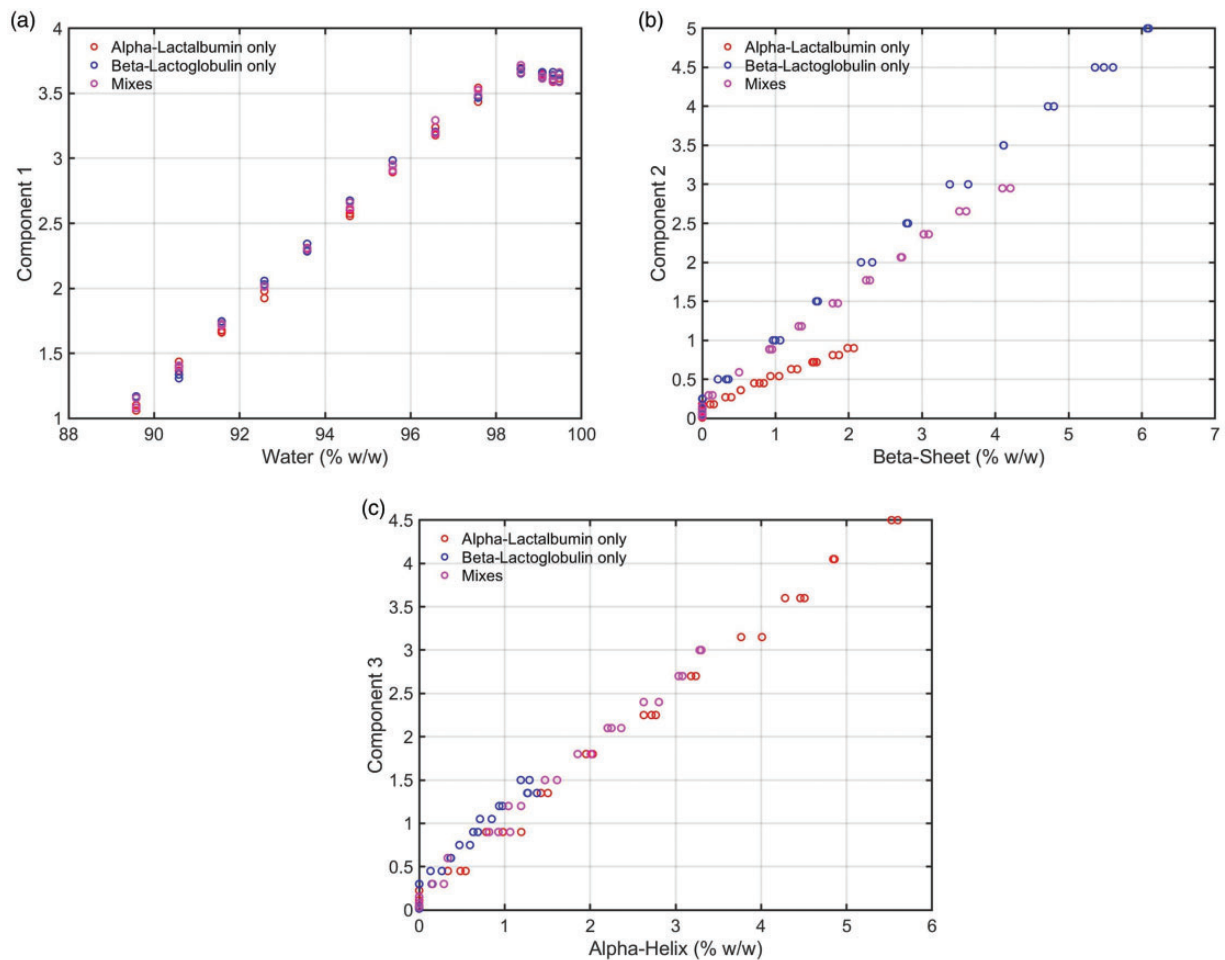


Figure 8. MCR Components scores vs. calculated amount of secondary structure. (a) MCR Component 1 versus water concentration. (b) MCR Component 2 versus β -sheet concentration. (c) MCR Component 3 versus α -helix concentration.

in the samples (R^2 of 0.99) confirming the previous assignment. The MCR scores of component 1 require a scaling to be able to accurately predict the water concentration (or by closure the dry matter content). MCR component 3 (Fig. 8c) shows a high correlation with the theoretical α -helix concentration in the samples (R^2 of 0.98), including also the samples containing only β -Lg (which have low amount of α -helix). Again, a simple scaling of MCR component 3 will be able to give a near perfect prediction of the α -helix content. Last but not least, MCR component 2 (Fig. 8b) correlates to β -sheet structures (R^2 of 0.96), but here it is observed that the sample group with only α -La shows lower correlations in comparison to the β -Lg containing sample groups, indicating that β -sheet concentration seem to be less accurately estimated than α -helix concentrations. This might be due to the presence of other secondary structures that can be found in α -La and β -Lg, such as reverse turn structures that account for roughly 20% of β -Lg structure, and which are not considered in the MCR model. There might also be an issue with the compound purity of the whey protein powders, which were used to prepare the samples, as the secondary structure concentrations for each sample were calculated with reference values that might be subject to small, or unknown, variations. Another reason for the discrepancy in the correlation results could be an erroneous theoretical content of α -helix and β -sheet in the α -La and β -Lg powders. The theoretical secondary structure concentrations were calculated from X-ray diffraction studies of ultra pure compounds in solid state. The combined results obtained from the PLS and MCR models and from the 2D MIR-NIR correlation spectrum suggest that both MIR and NIR are suitable analytical technologies for quantifying α -La and β -Lg in aqueous solutions. The application of NIR spectroscopy might be of potential interest as a screening tool in the food industry, since an in-line NIR spectrometer could be applied on a production scale for monitoring of quality parameters (concentration of individual proteins) during the filtration process.

Conclusion

The study presents novel results of the application of FT-NIR for the quantitative determination of individual whey protein concentrations in a model whey solution. By carefully selecting the appropriate spectral region and spectral preprocessing, α -La and β -Lg in aqueous whey samples were successfully predicted from the NIR and MIR spectra using PLS regression. Furthermore, NIR and MIR spectra were used in conjunction to produce a 2D MIR-NIR correlation spectrum in order to understand the most beneficial parts of the NIR spectra for quantification of individual proteins. The amide II band of the MIR spectra and the NIR range from 4000 to 5000 cm^{-1} was found to be particularly useful for quantifying protein secondary structure,

apparently because the amide I band is strongly overlapped by the water H–O–H bending absorption. Since the use of the amide II band for quantifying single proteins is not well reported in literature, an MCR model was developed for the MIR spectra in order to elucidate the information that could be extracted in the amide II band. The presented results support that the protein secondary structure can be successfully quantified from the MIR amide II band. The use of NIR for protein structural quantification opens new possibilities for at- and in-line predictions, enabling process analytical technology to be applied for ensuring product quality and quantity.^{2,3}

Acknowledgments

The whey protein powders used in this project were provided by Arla Food Ingredients (Videbæk, Denmark).



Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 801199. The whey protein powders used in this project were provided by Arla Food Ingredients (Videbæk, Denmark). The study was supported by Arla Food Ingredients (Videbæk, Denmark) and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 801199.

ORCID iDs

Margherita Tonolini  <https://orcid.org/0000-0002-3688-1335>
Peter B. Skou  <https://orcid.org/0000-0003-2540-5496>

References

1. A. Sturaro, M. De Marchi, A. Masi, M. Cassandro. "Quantification of Whey Proteins by Reversed Phase-HPLC and Effectiveness of Mid-Infrared Spectroscopy for Their Rapid Prediction in Sweet Whey". *J. Dairy Sci.* 2016. 99(1): 68–76.
2. L.K. Creamer, D.A.D. Parry, G.N. Malcolm. "Secondary Structure of Bovine β -lactoglobulin B". *Arch. Biochem. Biophys.* 1983. 227(1): 98–105.
3. N. Chandra, K. Brew, K.R. Acharya. "Structural Evidence for the Presence of a Secondary Calcium Binding Site in Human Alpha-lactalbumin". *Biochemistry.* 1998. 37(14): 4767–4772.
4. A. Barth. "Infrared Spectroscopy of Proteins". *Biochim. Biophys. Acta.* 2007. 1767(9): 1073–1101.
5. M.J.M. Rutten, H. Bovenhuis, J.M.L. Heck, J.A.M. Van Arendonk. "Predicting Bovine Milk Protein Composition Based on Fourier Transform Infrared Spectra". *J. Dairy Sci.* 2011. 94(11): 5683–5690.
6. A. Schwaighofer, J. Kuligowski, G. Quintás, H.K. Mayer, B. Lendl. "Fast Quantification of Bovine Milk Proteins Employing External Cavity-Quantum Cascade Laser Spectroscopy". *Food Chem.* 2018. 252: 22–27.

7. S. Kucheryavskiy, C. Lomborg. "Monitoring of Whey Quality with NIR Spectroscopy: A Feasibility Study". *Food Chem.* 2015. 176: 271–277.
8. P.D. Ingle, R. Christian, P. Purohit, V. Zarraga, et al. "Determination of Protein Content by NIR Spectroscopy in Protein Powder Mix Products". *J. AOAC Int.* 2016. 99(2): 360–363.
9. K. Izutsu, Y. Fujimaki, A. Kuwabara, Y. Hiyama, et al. "Near-Infrared Analysis of Protein Secondary Structure in Aqueous Solutions and Freeze-Dried Solids". *J. Pharm. Sci.* 2006. 95(4): 781–789.
10. S. Wold, K. Esbensen, P. Geladi. "Principal Component Analysis". *Chemom. Intell. Lab. Syst.* 1987. 2(1-3): 37–52.
11. S. Wold, H. Martens, H. Wold. "The Multivariate Calibration Problem in Chemistry Solved by the PLS Method". *Matrix Pencils.* 1983. 973: 286–293.
12. A. Juan, J. Jaumot, R. Tauler. "Multivariate Curve Resolution (MCR). Solving the Mixture Analysis Problem". *Anal. Methods.* 2014. 6(14): 4964–4976.
13. I. Noda. "Generalized Two-Dimensional Correlation Method Applicable to Infrared, Raman, and Other Types of Spectroscopy". *Appl. Spectrosc.* 1993. 47(9): 1329–1336.
14. F.E. Barton, D.S. Himmelsbach. "Two-Dimensional Vibrational Spectroscopy II: Correlation of the Absorptions of Lignins in the Mid- and Near-Infrared". *Appl. Spectrosc.* 1993. 47(11): 1920–1925.
15. A. Schwaighofer, M.R. Alcaraz, L. Lux, B. Lendl. "pH Titration of β -lactoglobulin Monitored by Laser-based Mid-IR Transmission Spectroscopy Coupled to Chemometric Analysis". *Spectrochim. Acta, Part A.* 2020. 226: 117636.
16. P. Geladi, D. MacDougall, H. Martens. "Linearization and Scatter-Correction for Near-Infrared Reflectance Spectra of Meat". *Appl. Spectrosc.* 1985. 39(3): 491–500.
17. B.G. Osborne, T. Fearn, P.H. Hindle. *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis.* Harlow, UK: Longman Scientific and Technical, 1993. Pp. 26–33.
18. Å. Rinnan, F. Van Den Berg, S.B. Engelsen. "Review of the Most Common Pre-processing Techniques for Near-infrared Spectra". *TrAC, Trends Anal. Chem.* 2009. 28(10): 1201–1222.
19. A. Savitzky, M.J.E. Golay. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures". *Anal. Chem.* 1964. 36(8): 1627–1639.
20. T. Ringsted, H.W. Siesler, S.B. Engelsen, "Monitoring the Staling of Wheat Bread Using 2D MIR-NIR Correlation Spectroscopy". *J. Cereal Sci.* 2017. 75: 92–99.
21. W.K. Surewicz, H.H. Mantsch, D. Chapman. "Determination of Protein Secondary Structure by Fourier Transform Infrared Spectroscopy: A Critical Assessment". *Biochemistry.* 1993. 32(2): 389–394.
22. J. Kong, S. Yu. "Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures". *Acta Biochim. Biophys. Sin.* 2007. 39(8): 549–559.
23. F. Van den Berg, C.B. Lyndgaard, K.M. Sørensen, S.B. Engelsen. "Process Analytical Technology in the Food Industry". *Trends Food Sci. Technol.* 2013. 31(1): 27–35.