Feed-forward Prediction of Product Qualities in Enzymatic Protein Hydrolysis of Poultry By-products: A Spectroscopic Approach.

Sileshi Gizachew Wubshet[a]*, Jens Petter Wold[a], Nils Kristian Afseth[a], Ulrike Böcker[a], Diana Lindberg[a], Felicia Nkem Ihunegbo[a], Ingrid Måge[a]

[a]Nofima AS - Norwegian Institute of Food, Fisheries and Aquaculture Research, PB 210, N-1431 Ås, Norway

Corresponding author. Tel.: +47 909 17 126

E-mail address: sileshi.wubshet@nofima.no (Sileshi G. Wubshet).

Key words: Enzymatic protein hydrolysis; Feed-forward process control; Sequential Orthogonalised-Partial Least Squares (SO-PLS); Spectroscopy; Poultry by-products

Abstract

Enzymatic protein hydrolysis (EPH) is one of the industrial bioprocesses used to recover valuable constituents from food processing by-products. Extensive heterogeneity of by-products from, for example, meat-processing is a major challenge in production of protein hydrolysates with stable and desirable quality attributes. Therefore, there is a need for process control tools for production of hydrolysates with defined qualities from such heterogeneous raw materials. In the present study, we are reporting a new feed-forward process control strategy for enzymatic protein hydrolysis of poultry by-products. Four different spectroscopic techniques, i.e., NIR imaging scanner, a miniature NIR (microNIR) instrument, fluorescence and Raman, were evaluated as tools for characterization of the raw material composition. Partial least squares (PLS) models for ash, protein and fat content were developed based on Raman, fluorescence and microNIR measurements, respectively. In an effort to establish feed-forward process control tools, we developed statistical models that enabled prediction of end-product characteristics, i.e., protein yield and average molecular weight of peptides (*Mw*), as a function of raw material quality and hydrolysis time. A multi-block sequential orthogonalised-PLS (SO-PLS) model, where spectra from one or more techniques and hydrolysis time were used as predictor variables, was fitted for the feed-forward prediction of product qualities. The best model was obtained for protein yield based on combined use of microNIR and fluorescence (R^2 = 0.88 and RMSE = 4.8). A Raman-based model gave a relatively moderate prediction model for *Mw* (R^2 = 0.56 and RMSE = 150). Such statistical models based on spectroscopic measurements of the raw material can be vital process control tools for EPH. To our knowledge, the present work is the first example of a spectroscopic feed-forward process control for an industrially relevant bioprocess.

1. Introduction

By-products from industrial food processing constitute one of the major figures in the global food wastage footprint (Kummu et al. 2012; Parfitt et al. 2010). Recent advance in biotechnology opened several possibilities for utilization of by-products from industrial processing of milk, seafood and several meat products. Enzymatic protein hydrolysis (EPH), i.e., protease catalyzed digestion of proteins, is one of the well-recognized technologies in valorization of by-products and side-streams (Tavano 2013; Aspevik et al. 2017). This technology has been widely successful in production of high-value products, such as sports nutrition and infant formulations, from diary processing by-products (Nasirpour et al. 2006; Tang et al. 2009). However, similar successes are largely unaccomplished in the meat industry and, in many cases, by-products from this sector are regarded as high-volume-low-value side streams. This is, to a significant extent, because meat-processing by-products are very heterogeneous matrices characterized by high degree of composition variation. Hence, controlled production of protein hydrolysates with stable quality attributes is one of the major challenges in valorization of meat-processing by-products.

The existing strategies in controlling quality of protein hydrolysates, nearly exclusively, rely on analysis of the end-product. Degree of hydrolysis (DH), molecular weight distribution (MWD), amino acid composition and total protein recovery are the key industrially relevant quality parameters that determine intrinsic properties of a given product (Li-jun et al. 2008; Rutherfurd 2010; Fountoulakis and Lahm 1998). While amino acid composition is mostly dependent on the raw material, the other three quality parameters (DH, MWD and yield) are strongly dependent on controllable process parameters such as hydrolysis time. Traditionally, these parameters are measured offline using chemical and chromatographic methods. Recent applications using Fourier transform infrared (FTIR) spectroscopy have also been reported as a potential technique for rapid monitoring of protein chain reduction in enzymatic hydrolysis of

dairy, fish and poultry by-products (Poulsen et al. 2016; Bocker et al. 2017; Wubshet et al. 2017). The general limitation of all the methods is, however, that the measurements are performed either during hydrolysis or on the products. This means, process settings, such as enzyme concentration, cannot be adjusted to maintain a specific product quality. Therefore, new strategies for process modeling and control are highly relevant to ensure stable production of protein hydrolysates.

Understanding the underlying relationship of raw material composition and process settings with end-product quality is a crucial aspect of controlling an industrial production (Jørgensen and Næs 2004). In a process where raw material variation is apparent, maintaining a stable product quality requires continuous monitoring of quality changes at a critical point in the production line. In this regard, rapid spectroscopic characterization of raw materials presents a valuable tool that allows progressive control measures of adjusting process settings in a feed-forward manner. A feed-forward approach is one of the advanced process control strategies where a sensor installed upstream in the production line is used to measure raw material variability and transmit signal in real time to downstream unit operation to take the required compensative actions (Muteki et al. 2012). Such strategies are widely acknowledged in pharmaceutical production in the context of process analytical technology (PAT) (Hinz 2006). In contrast, despite few examples, application of PAT for process control is less established in food industries (van den Berg et al. 2013; Sahni et al. 2004; Tajammal Munir et al. 2015). Reduced degree of automation and lack of analytical method have been identified as some of the major challenges associated with the significantly less frequent application of PAT in food industries (Bernd et al. 2015). Therefore, there is a need for new analytical technologies and strategies that can serve as elements of PAT for food sectors.

Rapid, real-time measurement with spectroscopic sensors represents a valuable tool for characterizing food matrices, and it is therefore a key element in the PAT. In the current work,

we have developed new feed-forward process control tools for enzymatic hydrolysis of poultry processing by-products based on rapid spectroscopic measurements. By-products from poultry slaughterhouses span significant variation in major constituents such as protein, ash and fat. Therefore, production of protein hydrolysates with predefined qualities such as molecular weight distribution and protein yield is a challenging task. The present work demonstrates the potential of four different spectroscopic techniques, i.e., NIR scanner, microNIR, fluorescence and Raman, as tools for rapid characterization of the raw materials. Finally, the spectroscopic techniques were evaluated as feed-forward process control tools for their efficiency in predicting product qualities (i.e., average molecular weight (*Mw*) and protein yield). Such feed-forward prediction of product qualities using rapid spectroscopic measurements represents a novel analytical strategy in enzymatic protein hydrolysis. A summary of the concept and the major processes involved in enzymatic hydrolysis of by-product is illustrated in Fig 1.

2. Material and methods

2.1. *Chemicals and sample materials*

Sample materials were collected from both filleting (skin, thighs and carcasses) and mechanical chicken deboning (deboning residues and deboned meat) department of a Norwegian poultry processing plant (Nortura, Hærland, Norway). All samples are by-products either collected directly from the production line or were prepared by mixing in a defined proportion (see Table 1). Samples were coarsely ground using a meat grinder prior to further analysis. A total of 32 samples, 500 g each, were prepared and stored at -4 °C until further analysis. Protease from *Bacillus licheniformis* (Alcalase 2.4 L, 2.4 U/g; EC number 232-752-2), analytical grade acetonitrile, trifluoroacetic acid, monosodium phosphate and analytical standards for molecular weight calibration (Wubshet et al. 2017) were purchased from Sigma-Aldrich (St. Louis, MO). Water used for HPLC was purified by deionization and 0.22 μm membrane filtration (Millipore, Billerica, MA).

2.2. Reference measurements (percentage protein, fat and ash)

Percentages of protein, fat and ash were measured for the 32 raw materials. Total nitrogen of each sample was measured according to Nordic Committee on Food Analysis (NMKL) method number 6 (NMKL 2003). A nitrogen to protein conversion factor of 6.25 was used to calculate the percentage of protein. The fat and ash percentage was determined according to NMKL 131 and NMKL 173, respectively (NMKL 1989; NMKL 2005). Protein yield (%) was calculated as follows:

$$\text{Protein yield (\%)} = \frac{\text{Total N in the water phase of the product}}{\text{Total N in the ground raw material}} \times 100$$

2.3. Spectroscopic measurements

Spectroscopic measurements of the 32 ground raw materials were performed using fluorescence, Raman, a miniature NIR instrument (microNIR) and an NIR imaging scanner. For acquisition of the spectra, 500 g of sample were arranged in cuboid shape with approximate dimension of 21 cm × 28 cm × 2 cm. Schematic summary of the four setups and sampling trails are presented in Fig 2. Fluorescence measurements, in front face mode, were performed using a Fluoromax-4 spectrofluorometer (Horiba Scientific, Kyoto, Japan) equipped with a FL-300/FM4-3000 bifurcated fiber-optic probe. The scanning speed of the instrument was 80 nm/second. The excitation light was 340 nm and emission spectra were recorded in the range 360-650 nm. For each sample, spectra were acquired at five different spots and an average of the five spectra were used for the statistical analysis.

The microNIR measurements were performed using a hand-held spectrometer by VIAVI (Santa Rosa, CA, USA). The spectrometer was positioned approximately 5 mm above the surface of the sample and was manually moved in a zigzag pattern during the acquisition. The measurements were performed in reflectance mode with an acquisition time of 4 seconds. Spectra were acquired between 950 nm and 1650 nm with a wavelength interval of 6.2 nm. All spectra were normalized using standard normal variate (SNV) prior to the statistical analysis.

The on-line NIR system was a QVision500 (TOMRA Sorting Solutions, Leuven, Belgium), an industrial hyperspectral imaging scanner designed for on-line measurement of, typically, fat and protein in meat on conveyor belts (Wold et al. 2016; Wold et al. 2017). The instrument is based on interactance measurements where the light is transmitted into the meat and then back scattered to the surface. Optical sampling depth in the sample material is typically 10-20 mm. Each NIR measurement took about 1 sec. The scanner was placed 30 cm above the conveyor belt with no physical contact between samples and the instrument. The scanner collected hyperspectral images of 15 wavelengths between 760 and 1040 nm with a spectral resolution

of 20 nm. The output per sample scan was an image of the conveyor belt with the sample. Size of the image was 60 pixels in the direction perpendicular to belt movement and 200 pixels in the direction of belt movement. Each pixel represented a spatial area of about 7 mm × 5 mm across and along the conveyor direction, respectively. The imaging capability of the system was used mainly for effective sampling, to obtain one representative mean spectrum from each sample. The individual intensity spectra (int) were subjected to log transformation to obtain the absorption spectra (abs = log (1/int)). The absorption spectra were then normalized by SNV transformation.

The Raman measurements were carried out using a RamanRXN2^{TM} Hybrid system equipped with a non-contact PhAT-probe (Kaiser Optical Systems, Inc., Ann Arbor, MI). The excitation wavelength was 785 nm with a spot size of 6 mm at 25 cm working distance. Raman spectra were collected in range from 175 to 1875 cm^{-1} with an accumulation time of 15 sec × 4. The samples were moved manually as visualized in Fig 2 to secure representative sampling. Before statistical analysis, the Raman spectra were pre-processed using extended multiplicative scatter correction (EMSC) with a sixth order polynomial extension (Liland et al. 2016). The reference mean spectrum used for correction was subjected to asymmetric least squares baseline correction prior to EMSC.

2.4. Enzymatic protein hydrolysis

Of the 32 raw materials, the selected eight (highlighted in bold in Table 1) were subjected to enzymatic protein hydrolysis. The hydrolysis was performed in a Reactor-Ready^{TM} jacketed reaction vessel (Radleys, Saffron Walden, Essex, United Kingdom). Water running through the vessel jacket was kept at 50 °C and delivered using a JULABO circulator pump (Julabo GmbH, Seelbach, Germany). The homogenized samples (500 g) were suspended in 1 L of water and thoroughly mixed until the suspension reached 50±1 °C. The proteolysis was started by addition of 7.5 mL of Alcalase. Each raw material was processed in duplicate for either 10 or 60 min.

This resulted in a total of 32 hydrolysis experiments. Further analysis could not performed on two of the 32 hydrolysate (one of the replicates from 10 min hydrolysis of T5 and T4, see Table 1) due to technical problems. After the specific hydrolysis time the enzyme was thermally inactivated by heating the reaction mixture in a water bath at 95 °C for 15 min. After inactivation, the contents in each batch was transferred to a 2 L centrifuge tube and centrifuged for 15 min at 4600g and 4 °C to afford two phases (i.e., water and fat) of supernatant and segmented solid residue. After volumetric measurements, samples were collected from the protein fraction (i.e., water phase) for total nitrogen and chromatographic measurements.

2.5. *Size exclusion chromatography*

Molecular weight distribution of protein fractions from each of the hydrolysis experiments were analyzed using size exclusion chromatography. Filtrates of the water phase collected from the hydrolysis were directly used as sample injection solution while injection solutions of standards were prepared in water at a concentration of 2 mg/mL. Chromatographic separation of both the standards and the hydrolysates was performed with an Agilent 1200 series instrument (Santa Clara, CA) consisting of a quaternary pump, a degasser, a thermostated column compartment, a photodiode-array detector and an auto sampler. An injection volume of 10 µL was used and separation was performed at 25 °C using BioSep-SEC-s2000 column (Phenomenex, 300 × 7.8 mm). The mobile phase was a mixture of acetonitrile and ultrapure water in a proportion 30:70 (v:v), containing 0.05% trifluoroacetic acid (TFA). Isocratic elution was carried out using a flow rate of 0.90 mL/min for 17 minutes. Between 17.0 and 17.1 minute the mobile phase was changed to NaH_2PO_4 (0.10 M) and maintained up to 20.0 minute for column cleaning. Elution conditions were restored between 20.0 and 20.1 minute and the column was equilibrated for additional 25 minutes. Chromatographic runs were controlled from OpenLAB CDS Rev. C. 01.07 (Agilent Technologies, Inc., Santa Clara, CA, USA). Average molecular weight (*Mw*)

calculations were performed using PSS winGPC UniChrom V 8.00 (Polymer Standards Service, Mainz, Germany) according to a previously published protocol (Wubshet et al. 2017).

2.6. Statistical analysis

Partial Least Squares (PLS) regression is a multivariate regression method that can handle thousands of collinear descriptor variables, and it is routinely used for fitting spectroscopic calibration models (Martens and Næs 1989). The PLS model is linear, and using matrix algebra, it can be written as:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{f} \qquad (1)$$

where **y** is the response, and **X** is a matrix whose rows correspond to samples and columns correspond variables (e.g. spectral wavelengths). The **b** vector contains regression coefficients, reflecting the relationship between each **X**-variable and the response. **f** contains the residuals, i.e. the variation in **y** that cannot be explained by the variables in **X**. The model estimation is based on a decomposition of **X** into a number of latent variables, $\mathbf{X} = \mathbf{TP}'$, where the scores **T** and loadings **P** can be used for model interpretation. In this paper, PLS regression was used to fit spectroscopic calibration models for the content of ash, fat and protein in raw materials, as well as for predicting the end-product yield and *Mw* based on reference measurements and hydrolysis time.

Sequential Orthogonalised PLS (SO-PLS) is an extension of PLS regression for modelling multiple predictor blocks (Næs et al. 2013). The SO-PLS model for *K* predictor blocks can be written as:

$$\mathbf{y} = \mathbf{X}_1\mathbf{b}_1 + \cdots + \mathbf{X}_K\mathbf{b}_K + \mathbf{f} \qquad (2)$$

where each $\mathbf{X}_k$ is a matrix containing multiple collinear variables. The estimation is based on sequential PLS regressions of each predictor block (starting with $\mathbf{X}_1$), followed by an

orthogonalisation of the remaining blocks in each step. In this way, the incremental contribution from each data matrix can be interpreted separately. One advantage of SO-PLS is that it can be used to combine data blocks with very different dimensionalities, such as spectra and process variables. In this paper, SO-PLS was used to combine spectroscopic raw material measurements and hydrolysis time for prediction of protein yield and *Mw.* All the PLS and SO-PLS models were validated by segmented cross-validation, taking out all samples from the same raw material in each segment since these cannot be considered as statistically independent. This is a conservative validation strategy, especially since the raw material qualities are very heterogeneous, and it was chosen to prevent overfitting. The number of model components were selected by evaluating the root mean squared error of cross-validated predictions ($RMSE_{CV}$). A global model selection strategy was used for the SO-PLS models, meaning that all combinations of components were compared in a so-called Måge-plot (Næs et al. 2011). In addition to the regression analysis, fixed effects analysis of variance (ANOVA) including main effects and interactions was used to analyse the effects of raw material (8 levels) and hydrolysis time (2 levels) on protein yield and *Mw*.

All statistical analysis were performed using MATLAB (R2016a, The Mathworks, Inc., Natick, USA, www.mathworks.com). The MATLAB code for estimating SO-PLS models can be downloaded from http://nofimamodeling.org/software-downloads-list/multiblock-regression-by-poso-pls/.

3. Results and discussion

3.1. Raw material composition variations

Descriptive statistics, i.e., minimum, maximum, mean, and standard deviation, of ash, fat and protein content of the 32 raw materials is presented in Table 2. A preliminary observation of the data revealed that the 32 raw materials spanned a significant variation in ash and protein whereas the fat content was relatively similar for the majority of the samples (approximately 20%), Fig 3. Even though these samples were specifically selected or mixed to span as large a variation as possible, all the constitutes are low-value by-products. Therefore, the samples studied here can be regarded as representative examples of the expected raw material variation in industrial processing of poultry byproducts. The observed large variation highlights the significance of monitoring raw material composition and take progressive measures of adjusting process settings to meet desired product characteristics. An overall empirical correlation was observed between ash and protein content, with the sample mainly composed of tendon and skin (C7) outlying the correlation (Fig 3d). No other significant correlations were observed between the measured components. Pesti *et al* have reported a different set of correlations for broiler carcasses, where total nitrogen (i.e., a measure of protein content) is positively correlated with moisture while ash content remains relatively unchanged within fifty samples (Pesti and Bakalli 1997). The different trend observed in the current study is not unexpected since the samples were gathered from different body parts of different fowls and therefore no longer resemble native composition of a single broiler.

3.2. Spectroscopic analysis of raw material

As a first step to establish a feed-forward strategy of controlling product qualities, the four spectroscopic measurements were evaluated for their prediction efficiency of the raw materials compositions. PLS regression was used to correlate the spectra to the contents of protein, fat and ash, see Table 3 for details of the results. For protein estimation, the fluorescence-based

model gave a relatively higher R^2 (0.81) value and lower $RMSE_{CV}$ (0.7). The predicted versus reference values and regression coefficients for the fluorescence model are given in Figs 4a and b, respectively. A contribution from an emission band at 376 nm dominates the regression coefficients of the fluorescence-based model. This band, with a shoulder peak at 440 nm, has previously been assigned as a characteristic signal of collagen in meat (Pu et al. 2013). Due to higher share of connective tissues in the by-products used in the study, a significantly high content of collagen is expected.

For fat content, the best predictions were obtained using the two NIR techniques (see Figs 4c-d for microNIR). The microNIR model was mainly influenced by a band at 1200 nm which was assigned as the second overtone of the CH stretching originating from the fatty acid residues (Burns and Ciurczak 2008). Raman gave the best ash predictions (Figs 4e-f), while microNIR was only slightly behind. The distinct Raman shift strongly influencing the ash prediction model, i.e. approximately at 960 cm^{-1}, was assigned as a phosphate band ($\nu_1 PO_4^{3-}$), a characteristic feature of apatites from bone. Therefore, the significance of the phosphate band in the model for ash prediction is comprehensible, since bone minerals are the major contributors of the total ash content in the samples. Raman profiling has previously been used to qualitatively characterize composition of bone in several medical studies (Mandair and Morris 2015; Morris and Mandair 2011). However, to the knowledge of the authors, quantitative prediction of ash contents in meat-based samples has never been reported before.

3.3. Protein yield and molecular weight distribution

Among the crucial quality attributes of a product from hydrolysis of a protein-based by-products are the protein yield and molecular weight distribution of peptides. The latter is typically measured using size exclusion chromatography as either a distribution profile (Li-jun et al. 2008) or average molecular weight (Wubshet et al. 2017). These two quality parameters are often analyzed at the end of a given process and are distinctly related to both the raw material

composition and process settings (Gilmartin and Jervis 2002; Šližytė et al. 2005). In the present study, eight selected poultry by-products were subjected to enzymatic hydrolysis using two different process settings, i.e., hydrolysis time of 10 and 60 minutes. Selection of the eight raw materials for the hydrolysis was based on a criteria that samples should span a wide variation in composition (i.e., fat, protein and ash) and also equally represent the two animal origins (i.e., chicken or turkey). The selected samples are highlighted in bold in Table 1, and highlighted in red color in Figs 4a, c and e. Protein yield and *Mw* were analyzed as key quality parameters of the products.

Analysis of variance with the two factors, i.e., raw material composition and hydrolysis time, revealed that there are significant effects of both factors on protein yield and *Mw*, and a small but statistically significant interaction effect (see Table 4 and Fig 5). As expected, products acquired from 60 minutes of hydrolysis were characterized by a relatively higher protein yield and lower *Mw* compared to those acquired from 10 minutes of hydrolysis (Fig 5). This was consistent with time dependencies of the two quality parameters where longer hydrolysis time is related to both higher degree of solubilization (i.e., protein yield) and higher degree of hydrolysis (i.e., lower *Mw*). Interestingly, the effect of raw material composition was found to be larger than the effect of hydrolysis time for both of the product quality parameters. This can be seen from the vertical shift of the lines in Fig 5, and confirms the significance of raw material composition. The interaction effect was most prominent for *Mw*, where we clearly see that some of the raw materials have a higher rate of hydrolysis compared with others (represented by the steepness of the lines in Fig 5). The raw materials C4, C7, C11 and T11 have a steeper decrease in *Mw* than the other raw materials.

3.4. Feed-forward prediction of Mw and protein yield

Feed-forward process control requires a model for prediction of end-product characteristics as a function of raw material quality and controllable process parameters. Such a process model

may then be used to decide the optimal settings of process parameters based on the raw materials at hand and the desired product characteristics. As a baseline, we fitted a regular PLS model for protein yield and *Mw* using the reference analysis of 'protein', 'fat' and 'ash' composition in the raw materials together with 'hydrolysis time' as predictor variables. The models gave satisfactory predictions of both protein yield and *Mw,* with cross-validated R^2 of 0.77 and 0.70, respectively. Models with interaction terms were also tested, but we found no improvement in R^2 and $RMSE_{CV}$ values, despite the interaction effect identified by analysis of variance (Table 4). The reported models, therefore, contain linear terms only. The predicted versus measured values and regression coefficients are shown in Fig 6, and the model is also summarized in the first line of Table 5.

As stated in the previous section, the hydrolysis time was found to be highly significant for prediction of both the yield and *Mw.* The regression coefficient 'protein' was found to be negative for the prediction of yield, meaning that a high-protein raw material gives a lower yield. This is probably due to the fact that raw materials with higher content of protein need a longer hydrolysis time in order to afford a higher yield. In contrary, the recovery process from raw materials with lower protein content will be faster and afford a relatively higher yield in a given hydrolysis time. In addition, it should also be noted that raw materials with higher percentage of protein are characterized by higher amount of connective tissues and bones. Therefore, poor extractability of connective tissue proteins is expected to contribute to the lower yield in the case of raw materials with higher percentage of protein. For the *Mw* model, a positive regression coefficient of 'protein' was observed. Another important regression coefficient was 'fat' which was negative in prediction of yield. This is in agreement with the study by Šližytė et al. 2005, which has shown a negative relationship between the amounts of lipids in fish-based by-products, used as raw materials for enzyme-catalyzed hydrolysis, and recovered proteins in the resulting hydrolysates (Šližytė et al. 2005).

To utilize the feed-forward approach in an industrial setting, the analysis of raw materials has to be performed by fast and noninvasive measurements. We, therefore, evaluated two different strategies for replacing the reference analysis with spectroscopic measurements. First, the cross-validated predictions of protein, fat and ash (from the models in Table 3) were plugged into the baseline model. The prediction uncertainties introduced by using spectra instead of lab measurements are then propagated to the predicted yield and *Mw*, resulting in larger prediction errors (see Table 5). Then, the spectra were used directly in the model without relating them to the reference analysis. In addition, SO-PLS was used to fit multiblock models as shown in Equation (2), using spectra in the first **X**-block(s) and hydrolysis time in the last **X**-block. The four spectroscopic techniques were evaluated both separately and in combinations, and the best models are summarized in Table 5.

For feed-forward prediction of protein yield, the multiblock model with blocks microNIR, fluorescence and hydrolysis time gave the best model ($RMSE_{CV}$ of 4.4 and R^2_{CV} of 0.88). The most striking observation was that this model resulted a significantly better prediction of protein yield than the baseline model using reference measurements. This indicates that the spectra contain more useful information than simply the amounts of ash, fat and protein. The loadings from microNIR and fluorescence are shown in Fig 7. Interestingly, the overall profile of the microNIR (Fig 7b) and fluorescence (Fig 7c) loadings from the multiblock model of protein yield are very similar to those observed for the PLS models of fat (Fig 4d) and protein (Fig 4b), respectively. However, these major features of the loadings in the protein yield model have, notably, the opposite sign compared to the regression coefficients for the fat and protein model. This strongly agrees with the previously discussed baseline model (Fig 6), where a clear negative relationship was observed between both fat and protein content of the raw materials and the protein yield.

For *Mw,* the Raman-based model gave similar prediction error as the baseline model whereas the rest of the techniques did not result in adequate model. The molecular weight distribution of peptides, here measured as *Mw*, is a characteristic of a product highly related to the molecular structure of the substrate protein. In the current study, we attempted to predict this parameter based on spectroscopic techniques that are, with the exception of Raman, less sensitive to fine structure of the protein components and are more of a reflection of the gross composition. Therefore, prediction of a detailed molecular information such as molecular weight distribution was expected to be challenging.

Raman spectroscopy provides a platform for probing detailed structural information on protein structure and is, therefore, a promising technique for prediction of *Mw*. However, despite affording the best model in the current study, the obtained R^2 (0.59) and $RMSE_{CV}$ (150) values were considerably moderate. This is mostly because the Raman spectra of the byproducts is dominated by vibrational shifts from the fat component. These vibrational shifts highly overlap with bands related to protein secondary structure, e. g., amide I and III, that are potentially relevant for prediction of *Mw* (Wubshet et al. 2017). Therefore, substantial improvement of the Raman-based model could be achieved by performing the data acquisition on the water phase after the byproducts are solubilized and before addition of the enzyme. This will provide an opportunity to disregard the dominant vibrational shifts from the fat and capture relevant information on the protein component of the byproducts.

Overall, we have demonstrated the application of four rapid spectroscopic techniques as potential process control tools in EPH of poultry by-products. The current study is the first example of a process modelling strategy for EPH, where spectroscopic measurements of raw materials are used to predict important product quality parameters. This study is a vital step in developing an integrated feed-forward process control system where the early predictions can be used as signals to take a compensative action for the inevitable raw material variation. Such

action, for example, can be increasing enzyme concentration or hydrolysis time to meet a higher yield.

4. Conclusions

The present study aimed at developing a feed-forward strategy for prediction of process outcomes from an enzymatic protein hydrolysis of poultry by-products. Four different spectroscopic techniques were employed for measuring gross compositions of the raw materials and evaluated as tools for prediction of *Mw* and protein yield in feed-forward manner. Multi-block regression models, using spectra measured on the raw materials and hydrolysis time as concatenated predictor blocks, were developed for the two product characteristics. The best model was obtained for protein yield using microNIR and fluorescence. There have already been several advancements, in terms of optical engineering and robotics, in the field of both NIR and fluorescence spectroscopy. Hence, these techniques can further be developed and adapted in industrial setting for on-line measurements of raw materials in enzymatic protein hydrolysis. Therefore, the presented approach using microNIR and fluorescence holds strong potential as an industrial process control tools in enzymatic protein hydrolysis of meat-based by-products. However, more research has to be done to develop a robust model covering substantial amount of samples with the relevant composition variation.

Acknowledgments

Bjørg Narum and Marte Ryen Dalsnes are acknowledged for the excellent technical assistance. Financial support from the Norwegian Research Council through the CYCLE project (225349/E40) and the Chickenlysis project (235839), and the Norwegian Agricultural Food Research Foundation through the project FoodSMaCK — Spectroscopy, Modelling and Consumer Knowledge (262308/F40), is greatly acknowledged. Nortura and Norilia are

acknowledged for providing raw materials used in this study and for partnership related to enzymatic hydrolysis.

References

Aspevik, T., Oterhals, A., Ronning, S. B., Altintzoglou, T., Wubshet, S. G., et al. (2017). Valorization of proteins from co- and by-products from the fish and meat industry. Top Curr Chem (Cham). 375(3), 53.

Bernd, H., Ralph, H., Andreas, N., Daryoush, S., Jens, T. & Jarka, G. (2015). Process analytical technologies in food industry – challenges and benefits: A status report and recommendations. Biotechnology Journal. 10(8), 1095-1100.

Bocker, U., Wubshet, S. G., Lindberg, D. & Afseth, N. K. (2017). Fourier-transform infrared spectroscopy for characterization of protein chain reductions in enzymatic reactions. Analyst. 142(15), 2812-2818.

Burns, D. A. & Ciurczak, E. W. (2008). Handbook of near-infrared analysis. Boca Raton: CRC Press

Fountoulakis, M. & Lahm, H-. W. (1998). Hydrolysis and amino acid composition analysis of proteins. Journal of Chromatography A. 826(2), 109-134.

Gilmartin, L. & Jervis, L. (2002). Production of cod (Gadus morhua) muscle hydrolysates. Influence of combinations of commercial enzyme preparations on hydrolysate peptide size range. Journal of Agricultural and Food Chemistry. 50(19), 5417-5423.

Hinz, D. C. (2006). Process analytical technologies in the pharmaceutical industry: the FDA's PAT initiative. Analytical and Bioanalytical Chemistry. 384(5), 1036-1042.

Jørgensen, K. & Næs, T. (2004). A design and analysis strategy for situations with uncontrolled raw material variation. Journal of Chemometrics. 18(2), 45-52.

Kummu, M., de Moel, H., Porkka, M., Siebert, S., Varis, O. & Ward, P. J. (2012). Lost food, wasted resources: Global food supply chain losses and their impacts on freshwater, cropland, and fertiliser use. Science of The Total Environment. 438, 477-489.

Li-jun, L., Chuan-he, Z. & Zheng, Z. (2008). Analyzing molecular weight distribution of whey protein hydrolysates. Food and Bioproducts Processing. 86(1), 1-6.

Liland, K. H., Kohler, A. & Afseth, N. K. (2016). Model-based pre-processing in Raman spectroscopy of biological samples. Journal of Raman Spectroscopy. 47(6), 643-650.

Mandair, G. S. & Morris, M. D. (2015). Contributions of Raman spectroscopy to the understanding of bone strength. BoneKEy Rep. 4.

Martens, H. & Næs, T. (1989) Multivariate calibration. New York: Wiley.

Morris, M. D. & Mandair, G. S. (2011). Raman assessment of bone quality. Clinical orthopaedics and related research. 469(8), 2160-2169.

Muteki, K., Swaminathan, V., Sekulic S. S. & G. L. Reid (2012). Feed-Forward Process Control Strategy for Pharmaceutical Tablet Manufacture Using Latent Variable Modeling and Optimization Technologies. IFAC Proceedings. 45(15), 51-56.

Nasirpour, A., Scher, J. & Desobry, S. (2006). Baby foods: Formulations and interactions (A review). Critical Reviews in Food Science and Nutrition. 46(8), 665-681.

NMKL (1989). Fat, Determination according to SBR (Schmid-Bondzynski-Ratslaff) in meat and meat products. NMKL method No 131, Nordic Committee on Food Analysis, Oslo, Norway

NMKL (2003). Nitrogen, Determination in foods and feeds according to Kjeldahl. NMKL method No 6, 4th Ed. Nordic Committee on Food Analysis, Oslo, Norway

NMKL (2005). Ash, Gravimetric determination in foods. NMKL method No 173, 2nd Ed. Nordic Committee on Food Analysis, Oslo, Norway

Næs, T., Tomic, O., Afseth, N. K., Segtnan, V. & Måge, I. (2013). Multi-block regression based on combinations of orthogonalisation, PLS-regression and canonical correlation analysis. Chemometrics and Intelligent Laboratory Systems. 124, 32-42.

Næs, T., Tomic, O., Mevik, B. H. & Martens, H. (2011). Path modelling by sequential PLS regression. Journal of Chemometrics. 25(1), 28-40.

Parfitt, J., Barthel, M. & Macnaughton, S. (2010). Food waste within food supply chains: quantification and potential for change to 2050. Philosophical Transactions of the Royal Society B: Biological Sciences. 365(1554), 3065-3081.

Pesti, G. & Bakalli, R. (1997). Estimation of the composition of broiler carcasses from their specific gravity. Poultry Science. 76(7), 948-951.

Poulsen, N. A., Eskildsen, C. E., Akkerman, M., Johansen, L. B., Hansen, M. S., et al. (2016). Predicting hydrolysis of whey protein by mid-infrared spectroscopy. International Dairy Journal. 61, 44-50.

Pu, Y., Wang, W. & Alfano, R. R. (2013). Optical detection of meat spoilage using fluorescence spectroscopy with selective excitation wavelength. Applied Spectroscopy. 67(2), 210-213.

Rutherfurd, S. M. (2010). Methodology for determining degree of hydrolysis of proteins in Hydrolysates: a review. J AOAC Int. 93(5), 1515-1522.

Sahni, N., Isaksson, T. & Næs, T. (2004) In-line near infrared spectroscopy for use in product and process monitoring in the food industry. Journal of Near Infrared Spectroscopy. 12(2), 77-83.

Šližytė, R., Daukšas, E., Falch, E., Storrø, I. & Rustad, T. (2005). Characteristics of protein fractions generated from hydrolysed cod (Gadus morhua) by-products. Process Biochemistry. 40(6), 2021-2033.

Tajammal, M. M., Yu, W., Young, B. R. & Wilson, D. I. (2015) The current status of process analytical technologies in the dairy industry. Trends in Food Science & Technology. 43(2), 205-218.

Tang, J. E., Moore, D. R., Kujbida, G. W., Tarnopolsky, M. A. & Phillips, S. M. (2009). Ingestion of whey hydrolysate, casein, or soy protein isolate: effects on mixed muscle protein synthesis at rest and following resistance exercise in young men. Journal of Applied Physiology. 107(3), 987-992.

Tavano, O. L. (2013). Protein hydrolysis using proteases: An important tool for food biotechnology. Journal of Molecular Catalysis B: Enzymatic. 90, 1-11.

van den Berg, F., Lyndgaard, C. B., Sørensen, K. M. & Engelsen, S. B. (2013). Process analytical technology in the food industry. Trends in Food Science & Technology. 31(1), 27-35.

Wold, J. P., Bjerke, F. & Mage, I. (2016). Automatic control of fat content in multiple batches of meat trimmings by process analytical technology. Fleischwirtschaft. 96(11), 100-106.

Wold, J. P., Veiseth-Kent, E., Host, V. & Lovland, A. (2017) Rapid on-line detection and grading of wooden breast myopathy in chicken fillets by near-infrared spectroscopy. PLoS One. 12(3), e0173384.

Wubshet, S. G., Mage, I., Bocker, U., Lindberg, D., Knutsen, S. H., et al. (2017). FTIR as a rapid tool for monitoring molecular weight distribution during enzymatic protein hydrolysis of food processing by-products. Analytical Methods. 9(29), 4247-4254.

Table captions

Table 1. Description of raw materials and their corresponding composition (% ash, % fat and % protein) from reference measurements. The fat and ash percentage was determined according to NMKL 131 and NMKL 173, respectively. While percentage protein was determined using Kjeldahl method (NMKL 6) with nitrogen to protein conversion factor of 6.25.

Table 2. Descriptive statistics [minimum (Min), maximum (Max), mean, and standard deviation (SD)] of the composition (% ash, % fat and % protein) of the 32 raw materials.

Table 3. PLS regression results of prediction models for % ash, % fat and % protein based on the four spectroscopic measurements.

Table 4. Summary of results from ANOVA showing the effects of raw material and hydrolysis time on the protein yield and *Mw*.

Table 5. Summary of results from the feed-forward prediction models. Results from a baseline model using reference analysis, baseline models using predicted values of % ash, % fat and % protein from the four different spectroscopic methods and multi-block models are presented.

Tables

Table 1

Sample ID	Category	By-product composition	% Ash[a]	% Fat[a]	% Protein[a]
C1[¶]	Chicken	MDR	7.46	16.90	16.83
C2[¶]	Chicken	MDR	7.27	16.80	16.23
C3[¶]	Chicken	MDM	0.89	27.30	13.80
C4[¶]	**Chicken**	**MDM**	**0.88**	**26.07**	**13.33**
C5[¶]	Chicken	Skin from thighs	0.62	42.63	14.90
C6[¶]	**Chicken**	**Skin from thighs**	**0.69**	**46.33**	**15.90**
C7[¶]	**Chicken**	**Tendon and skin**	**2.56**	**19.23**	**20.10**
C8[†]	Chicken	90% MDR + 10% MDM	8.15	16.27	16.47
C9[†]	Chicken	80% MDR + 20% MDM	7.65	14.63	17.43
C10[†]	Chicken	70% MDR + 30% MDM	4.73	20.57	16.30
C11[†]	Chicken	60% MDR + 40% MDM	5.46	16.23	16.27
C12[†]	**Chicken**	**50% MDR + 50% MDM**	**3.97**	**16.33**	**16.03**
C13[†]	Chicken	50% MDR + 50% MDM	6.13	15.63	16.37
C14[†]	Chicken	90% MDR + 10% TS	5.85	20.87	15.83
C15[†]	Chicken	80% MDR + 20% TS	7.15	18.43	16.97
C16[†]	Chicken	70% MDR + 30% TS	6.18	20.67	15.80
T1[¶]	Turkey	Carcass	8.41	22.50	18.60
T1[¶]	Turkey	Carcass	7.03	19.50	17.33
T3[¶]	Turkey	MDR	11.30	17.10	20.13
T4[¶]	**Turkey**	**MDR**	**13.93**	**18.40**	**20.00**
T5[¶]	**Turkey**	**MDM**	**1.13**	**22.67**	**16.23**
T6[¶]	Turkey	MDM	1.05	22.17	16.07
T7[†]	Turkey	90% MDR + 10% MDM	9.70	17.93	19.63
T8[†]	Turkey	80% MDR + 20% MDM	8.26	19.37	17.53
T9[†]	Turkey	70% MDR + 30% MDM	10.45	18.10	19.27
T10[†]	Turkey	60% MDR + 40% MDM	8.57	20.20	17.67
T11[†]	**Turkey**	**50% MDR + 50% MDM**	**7.63**	**19.03**	**17.53**
T12[†]	Turkey	50% MDR + 50% MDM	6.57	20.70	17.30
T13[†]	Turkey	90% MDR + 10% TS	10.47	20.13	18.63
T14[†]	Turkey	80% MDR + 20% TS	10.38	23.60	18.27
T15[†]	**Turkey**	**70% MDR + 30% TS**	**10.34**	**25.93**	**18.03**
T16[†]	Turkey/Chicken	⅓ MDR (T) + ⅓ MDR (C) + ⅓ MDM (T)	4.97	19.17	16.67

[¶]Sample acquired directly from the production line
[†]Sample prepared by mixing different by-products
MDM: Mechanically deboned meat; MDR: Mechanical deboning residue; TS: Turkey skin from filleting section; C: Chicken; T: Turkey
Samples selected for enzymatic hydrolysis are highlighted in bold font

Table 2.

Composition	Min	Max	Mean	SD
% Ash	0,62	13,93	6,43	3,53
% Fat	14,63	46,33	21,29	6,82
% Protein	13,33	20,13	17,11	1,67

Table 3.

	Protein			Fat			Ash		
	Number of components	$RMSE_{CV}$	R^2_{CV}	Number of components	$RMSE_{CV}$	R^2_{CV}	Number of components	$RMSE_{CV}$	R^2_{CV}
Fluorescence	2	0.7	0.81	4	2.6	0.86	3	2.6	0.42
Raman	5	1.0	0.69	2	3.6	0.74	1	1.3	0.87
MicroNIR	3	1.1	0.57	2	1.9	0.93	4	1.6	0.80
NIR imaging scanner	3	1.4	0.34	4	2.0	0.92	4*	1.9	0.73

* No preprocessing of the spectra

Table 4.

	d.f.	Protein yield		*Mw*	
		Sum-of-squares (% of total)	p-value	Sum-of-squares (% of total)	p-value
Hydrolysis time (HT)	1	28.8	<0.001	39.3	<0.001
Raw Material (RM)	7	68.6	<0.001	54.2	<0.001
HT x RM	7	1.6	0.028	4.7	0.005
Error	14	1.0		1.8	
R^2_{adj}		0.98		0.96	

Table 5.

	Protein yield			*Mw*		
	Number of components	$RMSE_{CV}$	R^2_{CV}	Number of components	$RMSE_{CV}$	R^2_{CV}
Baseline model (Reference analysis combined with hydrolysis time)	2	6.3	0.77	3	127	0.77
Baseline model using predictions from:						
-best spectroscopic technique*		7.4	0.72		165	0.50
-Fluorescence		8.0	0.68		177	0.43
-Raman		8.4	0.64		170	0.47
-MicroNIR		8.0	0.67		185	0.37
-NIR imaging scanner		10.7	0.41		176	0.43
Multiblock models with blocks:						
-$\mathbf{X}_1$=Fluoresc., $\mathbf{X}_2$=Hydr.time	1+1	8.8	0.60	1+1	172	0.46
-$\mathbf{X}_1$=Raman, $\mathbf{X}_2$=Hydr.time	2+1	11.7	0.30	3+1	150	0.59
-$\mathbf{X}_1$=MicroNIR, $\mathbf{X}_2$=Hydr.time	**	-	-	1+1	182	0.40
-$\mathbf{X}_1$=Fluoresc., $\mathbf{X}_2$=Raman, $\mathbf{X}_3$=Hydr.time	1+1+1	6.9	0.76	-	-	-
-$\mathbf{X}_1$=MicroNIR, $\mathbf{X}_2$=Fluoresc., $\mathbf{X}_3$=Hydr.time	2+1+1	4.8	0.88	-	-	-

* The three constituents (i.e., ash, protein and fat) are predicted from three different models shown in Fig 3.

** Model breaks down because the samples coming from raw material C7 are extreme outliers.

Figure caption

Fig. 1 Schematic presentation of the major processing steps in enzymatic hydrolysis of poultry by-products and illustration of the proof-of-concept where spectroscopic measurements on the raw material are used to predict product quality in a feed-forward approach.

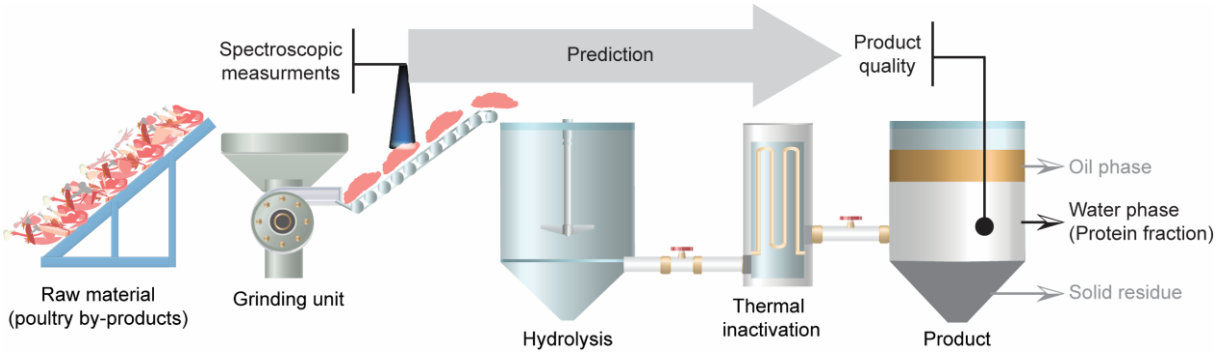


Fig. 2 Illustration of the different setups used for fluorescence (a), Raman (b), microNIR (c) and NIR imaging scanner (d) measurements. All measurements, except the NIR imaging scanner, were performed by manually moving the probe for spatially representative sampling. The measurement location or path is indicated with broken line.

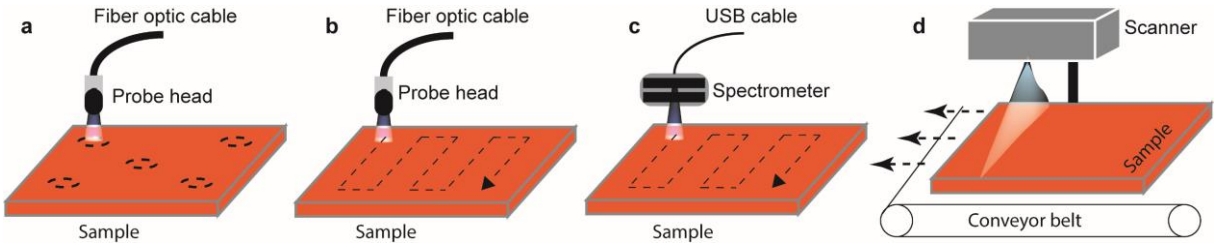


Fig. 3 Distribution of protein (a), fat (b) and ash (c) among 32 poultry-based by-products used in the present study. The graph (d) is illustrating the empirical correlation of ash and protein content within the sample set.

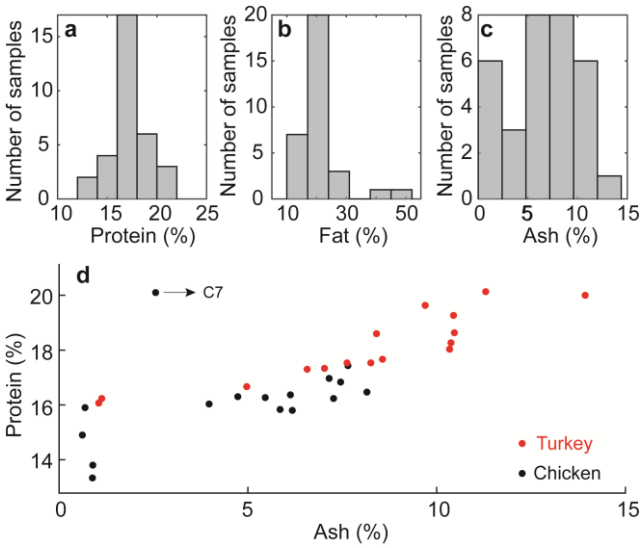


Fig. 4 Predicted versus measured values and regression coefficients from the best prediction models for protein (a-b, fluorescence), fat (c-d, microNIR) and ash (e-f, Raman). The samples that were selected for hydrolysis are highlighted with black solid fill.

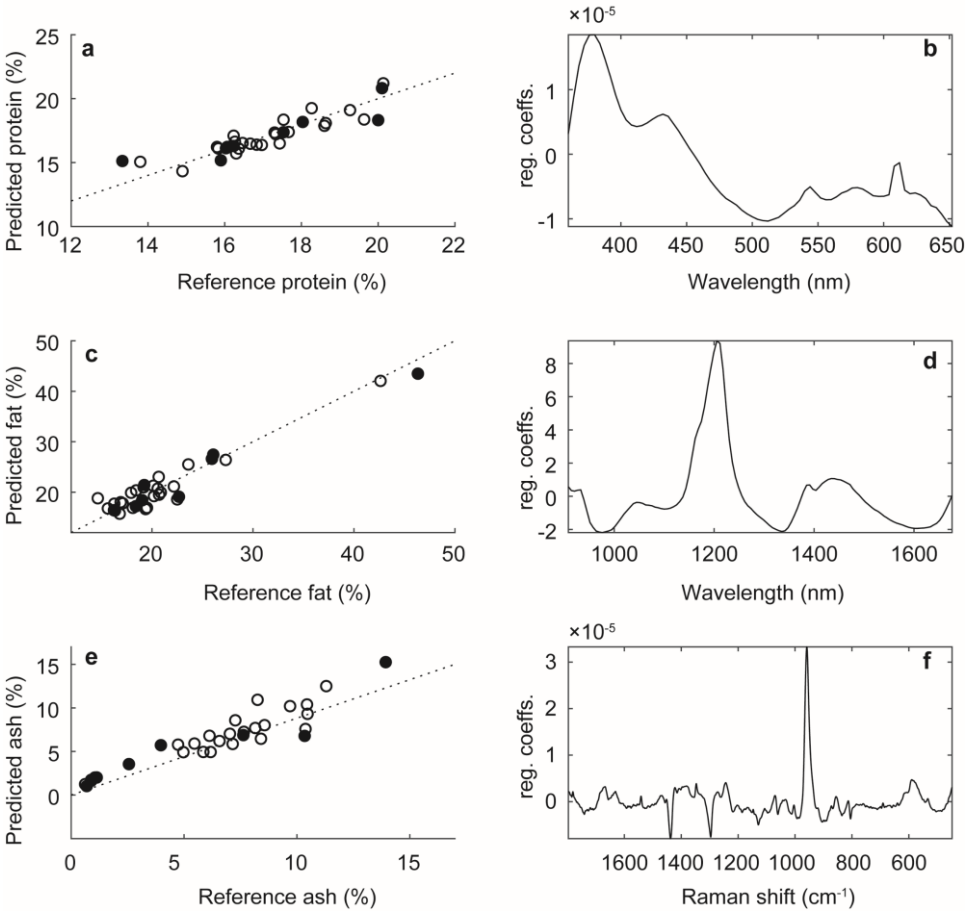


Fig. 5 Interaction plots showing the effects of raw material and hydrolysis time on the protein yield (a) and *Mw* (b).

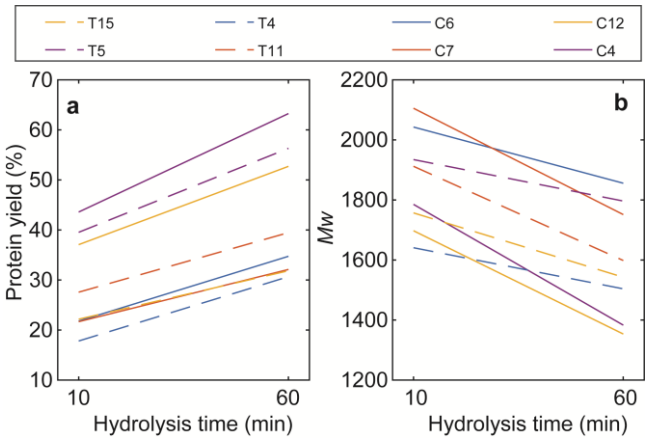


Fig. 6 Plots of predicted versus measured values and regression coefficients for the PLS models using reference values of ash, fat and protein in the raw material and hydrolysis time to predict end-product protein yield (a-b) and *Mw* (c-d).

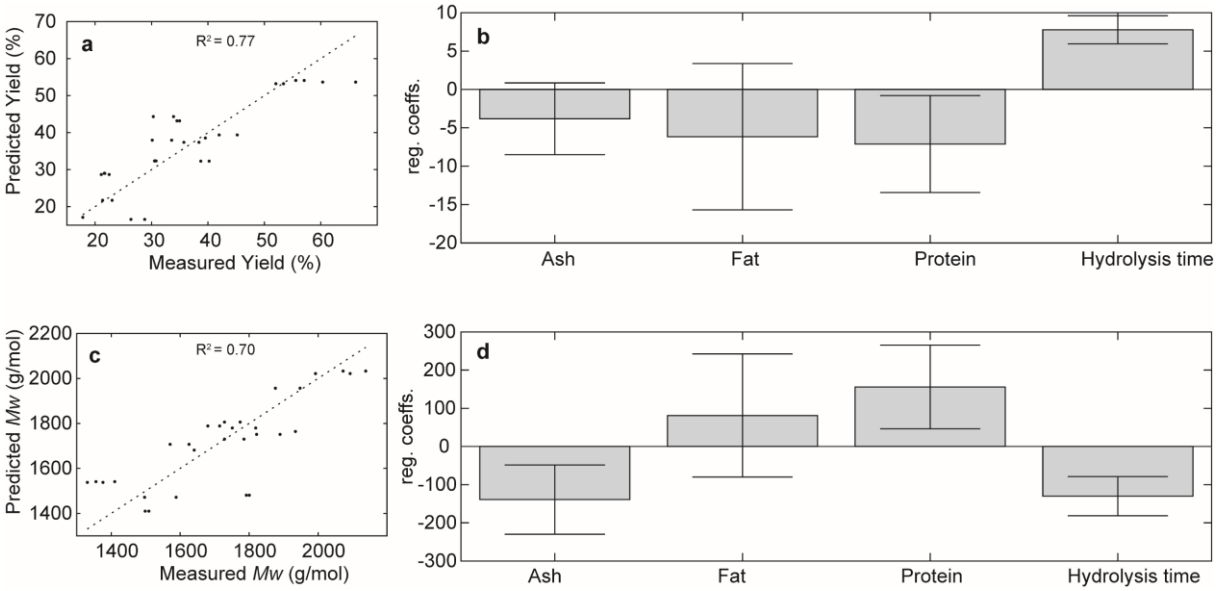
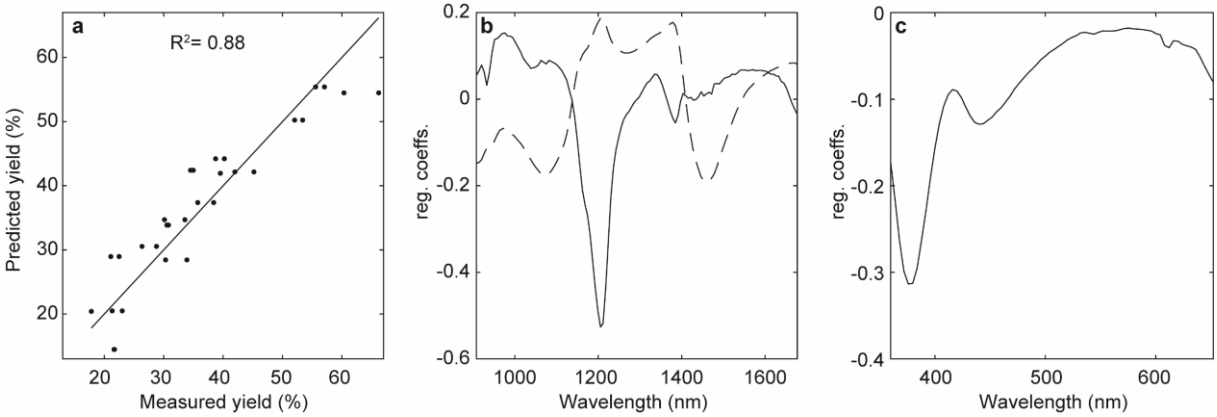


Fig. 7 Predicting protein yield from microNIR, fluorescence and hydrolysis time. Predicted versus measured values (a), loadings from the two microNIR components (b), loadings from the fluorescence component (c).


a
R²= 0.88
Predicted yield (%)
Measured yield (%)
b
reg. coeffs.
Wavelength (nm)
c
reg. coeffs.
Wavelength (nm)