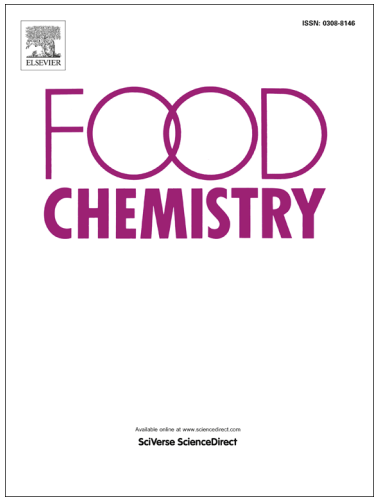
Journal Pre-proofs



Average molecular weight, degree of hydrolysis and dry-film FTIR fingerprint of milk protein hydrolysates: Intercorrelation and application in process monitoring

Kenneth Aase Kristoffersen, Nils Kristian Afseth, Ulrike Böcker, Diana Lindberg, Heleen de Vogel-van den Bosch, Mari Linnéa Ruud, Sileshi Gizachew Wubshet

PII: S0308-8146(19)31933-8
DOI: https://doi.org/10.1016/j.foodchem.2019.125800
Reference: FOCH 125800

To appear in: *Food Chemistry*

Received Date: 30 May 2019
Revised Date: 23 October 2019
Accepted Date: 24 October 2019

Please cite this article as: Aase Kristoffersen, K., Kristian Afseth, N., Böcker, U., Lindberg, D., de Vogel-van den Bosch, H., Linnéa Ruud, M., Gizachew Wubshet, S., Average molecular weight, degree of hydrolysis and dry-film FTIR fingerprint of milk protein hydrolysates: Intercorrelation and application in process monitoring, *Food Chemistry* (2019), doi: https://doi.org/10.1016/j.foodchem.2019.125800

Average molecular weight, degree of hydrolysis and dry-film FTIR fingerprint of milk protein hydrolysates: Intercorrelation and application in process monitoring

Kenneth Aase Kristoffersen[a,b,*Kenneth.kristoffersen@nofima.no], Nils Kristian Afseth[a], Ulrike Böcker[a], Diana Lindberg[a], Heleen de Vogel-van den Bosch[c], Mari Linnéa Ruud[b] and Sileshi Gizachew Wubshet[a]

[a]Nofima - Norwegian Institute of Food, Fisheries and Aquaculture Research, P.O. Box 210, N-1431 Ås, Norway

[b]Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, N-1432 Ås, Norway

[c]BioSensing & Diagnostics, Wageningen Food & Biobased Research, Wageningen University & Research, Bornse Weilanden 9, 6708 WG, Wageningen, the Netherlands

*Corresponding author.

Graphical Abstract

Highlights

DH% and M_w were monitored during enzymatic protein hydrolysis of dairy proteins.

Intercorrelation of DH% vs M_w were studied for 60 milk protein hydrolysates.

A PLSR model was developed for prediction of DH% based on molecular weight distribution profiles.

A multivariate model based on dry-film FTIR was developed for dual prediction of DH% and M_w.

Abstract

Fourier-transform infrared (FTIR) spectroscopy was applied to predict the degree of hydrolysis (DH%) and weight-average molecular weight (M_w) in milk protein hydrolysates. Both DH% and M_w are important quality parameters of protein hydrolysates. Measuring these parameters and following their development during proteolytic reactions is therefore essential for process control and optimization in industry. In the present study the intercorrelation and the complimentary nature of these parameters were investigated and a partial least square regression (PLSR) model was developed for the prediction of DH% from molecular weight distributions. Finally, we developed PLSR models based on dry-film FTIR spectroscopy for the prediction of both DH% and M_w. Here spectral changes in the amide region were found to be important for the two calibration models, underlining the advantage of dry-film FTIR measurement. This shows that dry-film infrared spectroscopy is a promising tool for dual prediction of DH% and M_w.

1 Introduction

Enzymatic protein hydrolysis (EPH) has become an attractive biotechnological process for the recovery of value-added peptides and amino acids from a range of food processing by-products. Bioactive peptides, infant formulas and nutritional supplements are among the most common products based on protein hydrolysates from by-products from fish, meat and dairy processing (Aspevik, Oterhals, Rønning, Altintzoglou, Wubshet, Gildberg, et al., 2017; Lordan, Ross, & Stanton, 2011; Martinez-Maqueda, Miralles, Recio, & Hernandez-Ledesma, 2012). The key advantage of the biotechnological process EPH is the possibility of adapting the process parameters to meet a specific product quality. A typical EPH process consists of protease-catalyzed degradation of dietary proteins, i.e., proteolysis, in batch or continuous reactors. One of the major processing parameters used as a variable to tailor-make specific products is the extent of proteolysis as a function of hydrolysis time. Consequently, monitoring of proteolysis during the EPH process is an essential element for successful production.

Recently, FTIR spectroscopy has been proposed as an industrially relevant rapid tool for monitoring protein degradation during EPH. Amide I and II bands originating from the protein backbone, together with vibrational bands from terminal COO^- and NH_3^+ have been identified as diagnostic features in the infrared (IR) spectra related to protein breakdown during EPH (Böcker, Wubshet, Lindberg, & Afseth, 2017; Guler, Dzafic, Vorob'ev, Vogel, & Mantele, 2011; Guler, Vorob'ev, Vogel, & Mantele, 2016; Ruckebusch, Duponchel, & Huvenne, 2001; Ruckebusch, Duponchel, Huvenne, Legrand, Nedjar-Arroume, Lignot, et al., 1999; Ruckebusch, Sombret, Froidevaux, & Huvenne, 2001). Poulsen et al. proposed liquid IR measurement for monitoring DH% of the hydrolysis of milk proteins (Poulsen, Eskildsen, Akkerman, Johansen, Hansen, Hansen, et al., 2016). However, due to the absorption of water, important IR features such as the amide I band were lost. Later, Wubshet et al. introduced the

first example of dry-film FTIR for prediction of M_w of peptides (Wubshet, Mage, Böcker, Lindberg, Knutsen, Rieder, et al., 2017).

A major challenge associated with the measurements of both DH% and M_w is the extensive sample preparation and long analysis times. Nevertheless, the two parameters are commonly used to characterize protein degradation in EPH, and they have been used interchangeably for process understanding and product characterization (Chi, Cao, Wang, Hu, Li, & Zhang, 2014; Li, Wang, Chi, Gong, Luo, & Ding, 2013; Sbroggio, Montilha, de Figueiredo, Georgetti, & Kurozawa, 2016; Slizyte, Mozuraityte, Martinez-Alvarez, Falch, Fouchereau-Peron, & Rustad, 2009). The measurement of DH% is most often accomplished using 2,4,6-trinitrobenzenesulfonic acid (TNBS), o-phthaldialdehyde (OPA) or the pH-stat method (Spellman, McEvoy, O'Cuinn, & FitzGerald, 2003). While these methods provide the means to monitor a given EPH process, they have notable shortcomings related to the instability of reagents, non-specific derivatizations and lack of standardization (Rutherfurd, 2010; Spellman, McEvoy, O'Cuinn, & FitzGerald, 2003). M_w on the other hand is derived from molecular distribution profile, typically obtained from size exclusion chromatography (SEC). DH% is a parameter used as a process monitoring variable reflecting the relative extent of reaction progress, while M_w is a direct measure of peptide composition reflecting product quality at a specific time during the course of the reaction. In some cases, it is therefore important to measure both DH% and M_w and explore their complimentary nature.

Beaubier et al. demonstrated the use of molecular weight distribution (MWD) profiles for predicting both *Mw* and DH%. This was achieved only for pure protein hydrolysates (Beaubier, Framboisier, Ioannou, Galet, & Kapel, 2019). This allowed for integration of clearly resolved peptide and undigested protein regions of the MWD profile. The ratio of these areas was used in a derived equation for the prediction of DH%. However, such an approach is arguably limited

to pure protein hydrolysates, since chromatographic resolution and selective integration of digested and undigested protein will not be feasible for complex hydrolysates.

In the present work, we demonstrate the first application of dry-film FTIR for dual prediction of DH% and M_w as key process monitoring parameters. In addition, a multivariate calibration model was developed for prediction of DH% from MWD profiles of milk protein hydrolysates.

2 Materials and methods

2.1 Materials

The two enzymes used, protease from *Bacillus licheniformis* (Alcalase, 2.4 U/g) and *Aspergillus oryzae* (Flavourzyme) by Novozymes (Bagsværd, Denmark), and the chemicals used for the DH% measurements and SEC analysis were all purchased from Sigma-Aldrich (St. Louis, MO, USA). This includes the TNBS, tris hydrochloride (Tris-HCl), analytical grade acetonitrile, trifluoroacetic acid and monosodium phosphate and disodium phosphate. The molecular weight standards used for the SEC analysis, (Albumin from chicken egg white, carbonic anhydrase from bovine erythrocytes, lysozyme, aprotinin from bovine lung, insulin chain B oxidized from bovine pancreas, renin substrate tetradecapeptide porcine, angiotensin II human, bradykinin fragment 1-7, [D-Ala2]-leucine encephalin, Val-Tyr-Val and tryptophan) were also purchased from Sigma-Aldrich. The water used for the HPLC mobile phase was purified by deionization and 0.22 μm membrane filtration (MilliporeSigma, Burlington, MA, USA).

2.2 Substrate raw materials

Three protein-rich materials derived from dairy sources were used to produce the hydrolysate samples. This includes milk whey protein concentrate powder (WPC80), whole milk powder (WMP) and milk whey powder (WPO), containing 77%, 28% and 12% protein respectively. The materials were provided by TINE (Oslo, Norway).

2.3 Enzymatic hydrolysis and sampling

The enzymatic hydrolysis reactions were performed in 250 mL bottles. Substrate raw materials and water were mixed in 3% w/v protein to liquid (i.e.,WPC80 (8 g), WMP (22 g) and WPO (52 g) to 200 mL water) to allow stirring in all solutions. All reaction mixtures were placed in a GFL® water bath (Thermo Fisher Scientific, Waltham, MA, USA) and mixed using a Variomag® magnetic stirrer (Thermo Fisher Scientific) until the suspension reached 50±1 °C. The reaction mixtures were then kept at the set temperature followed by addition of 2% enzyme w/w to protein content in the substrate. The reaction times for enzymatic hydrolysis for all combinations of enzyme and substrate were set to 2, 5, 10, 15, 30, 45, 60, 90, 120 and 180 minutes. When the desired reaction time was reached, the enzyme was thermally inactivated by fast microwave heating to minimum 90 °C followed by treatment for 15 minutes in a water bath which was kept at 90 °C. The reaction mixtures were then allowed to cool to room temperature followed by centrifugation using an Avanti® J-26 XP, Beckman Coulter™ (Beckman Coulter Life Sciences, Indianapolis, IN, USA) at 5000 RPM. The supernatant was collected and stored at -20 °C. The frozen samples were then freeze-dried and kept at -20 °C until analysis.

2.4 DH%-assay (TNBS)

The DH% was measured using a TNBS method based on descriptions by Satake et al. and Adler-Nissen (Adler-Nissen, 1979; Satake, Okuyama, Ohashi, & Shinoda, 1960). The buffer (0.21 M sodium phosphate buffer; pH 8.2) was prepared and stirred for 60 minutes at room temperature. Calibration solutions were prepared by a dilution series containing 0, 0.075, 0.15, 0.3, 0.6, 0.9, 1.2 and 1.5 mM Leucine in 1% SDS solution. The samples were prepared by dissolving 10 mg/mL hydrolysate powder in 0.1 M Tris-HCl pH 8.0 buffer followed by a dilution in 1% SDS-solution to 0.5 mg/ml. All samples and calibration solutions were measured in triplicate in Pierce™ 96-Well Polystyrene Plates, Corner Notch (Thermo Fisher Scientific). 15 µL of sample (reference or calibration solution) was added per well followed by the addition of 45 µL 0.21 M sodium phosphate buffer (pH 8.2) and 45 µL of a TNBS solution (0.05% w/v in water). The plate was sealed with a sticker and wrapped in aluminum foil to avoid UV degradation during the one hour incubation time at 50 °C. After incubation, 90 µL 0.1 M HCl was added to all wells before absorbance was measured at 340 nm using a BioTek Synergy™ H1 spectrophotometer (BioTek Instruments, Winooski, VT, USA). All measurements were performed in triplicates. The DH% values were then calculated according to Equation 1, using h_{tot} estimated from literature values and protein content measurements from Dumas combustion analysis (Church, Swaisgood, Porter, & Catignani, 1983; Simonne, Simonne, Eitenmiller, Mills, & Cresman, 1997; Spellman, McEvoy, O'Cuinn, & FitzGerald, 2003). The measured DH% values and the protein content analysis data is presented in supporting information (SI) Table S-2 and S-4.

$$\mathrm{DH\%} = \frac{h}{h_{tot}} \times 100\% \quad (1)$$

158 **2.5 Size exclusion chromatography**

159 SEC was preformed according to Wubshet et al. using 2 mg/mL solutions of standards and
160 rehydrated hydrolysate samples (1% w/v, filtrated using Millex-HV PVDF 0.45 µm 33 mm
161 filters (MilliporeSigma, Burlington, MA, USA)) as injection solutions (Wubshet, et al., 2017).
162 Chromatographic separation of standards and samples was performed with a Thermo Scientific
163 Dionex UltiMate 3000 Standard System (Thermo Fisher Scientific). The injection volume was
164 10 µL for the standards and 15 µL for samples. Separation was performed at 25 °C using a
165 BioSep-SEC-s2000 column (300 × 7.8 mm, Phenomenex, Torrence, CA, USA). The mobile
166 phase consisted of a mixture of acetonitrile and ultrapure water in a proportion of 30:70 (v/v),
167 containing 0.05 % trifluoroacetic acid. Isocratic elution was carried out using a flow rate of 0.9
168 mL/min for 20.0 minutes. Between 20.0 and 20.1 minutes the mobile phase was changed to
169 NaH_2PO_4 (0.10 M) and maintained until 23.0 minutes for column cleaning. Elution conditions
170 were restored between minute 23.0 and 23.1 and the column was equilibrated for an additional
171 27 minutes. Chromatographic runs were controlled from Chromeleon™ Chromatography Data
172 System (CDS) software (Thermo Fisher Scientific). From chromatographic runs of both the
173 standards and hydrolysates, a UV trace of 214 nm was used. The retention times of analytical
174 standards were obtained by manual peak-picking. The retention times of the standards were
175 used to construct a third polynomial (r^2=0.97) fitted calibration curve (Vander Heyden,
176 Popovici, & Schoenmakers, 2002). The retention times for the standards are presented in SI
177 Table S-1. Finally, M_w were calculated using PSS winGPC UniChrom V 8.00 (Polymer
178 Standards Service, Mainz, Germany) for each chromatogram. The calculation from the software
179 was based on a slicing method, similar to those previously used for analysis of protein
180 hydrolysates (Hsieh, Lin, Lang, Catsimpoolas, & Rha, 1979). The calculated M_w values are
181 presented in SI Table S-3.

182

2.6 Reproducibility study

Reproducibility studies were performed on both the DH% and the SEC methods by re-analyzing a selection of samples (5, 30, and 120 min samples for all the six hydrolysis time-series) intraday (n=6) and interday (n=6). All measurements were performed according to the description provided in chapter 2.3 and 2.4. The average, standard, and relative standard deviations are presented in SI Table S-5 and S-6.

2.7 FTIR spectroscopy

The samples for FTIR measurements were prepared by rehydration and filtration of the supernatant (5% w/v, filtrated using Millex-HV PVDF 0.45 μm 33 mm filters (MilliporeSigma, Burlington, MA, USA)). For all of the hydrolysates aliquots of 5 μL were deposited on 96-well IR-transparent Si-plates (Bruker, Billerica, MA, USA) and dried at room temperature for at least 30 minutes to form dry-films as described by Böcker et al. (Böcker, Wubshet, Lindberg, & Afseth, 2017). From each hydrolysate sample, five aliquots were deposited to allow for replicate measurements. FTIR measurements were performed using a High Throughput Screening eXTension (HTS-XT) unit coupled to a Tensor 27 spectrometer (Bruker, Billerica, MA, USA). The spectra were recorded in the region between 4000 and 400 cm^{-1} with a spectral resolution of 4 cm^{-1} and an aperture of 5.0 mm. For each spectrum, 40 interferograms were collected and averaged. Data acquisition was controlled using Opus v6.5 (Bruker, Billerica, MA, USA).

2.8 Data analysis

Pre-processing of FTIR spectra was performed using Savitzky-Golay algorithm with a polynomial degree of two and a window size of 13 points. The second-derivative spectra were

then normalized by applying extended multiplicative signal correction (EMSC) (Afseth & Kohler, 2012). The pre-treated data sets from the different samples were subjected to partial least squares regression (PLSR) modeling to predict DH% and M_w. For the PLSR, the spectral region from 1800-700 cm^{-1} was used. The regions from 5-20 min of the SEC chromatogram were also subjected to pre-processing using area normalization. The pre-treated datasets from 5-16 min of the SEC chromatograms were then subjected to PLSR modeling to predict DH%. Data processing and analysis were carried out using The Unscrambler® X v10.3 (CAMO Software AS, Oslo, Norway).

3 Results and discussion

Three dairy protein sources (WPC80, WMP and MPO) were hydrolyzed for up to three hours using two commercially available enzyme products (Alcalase (A) and Flavourzyme (F)). The reactions were stopped by thermal inactivation of the enzymes before samples were prepared for FTIR, SEC, nitrogen content and DH% analysis. The raw materials used are known to consist of a mixture of many different proteins. Whole milk protein, for example, consists of approximately 20% whey proteins with major components α-lactalbumin and β-lactoglobulin. The rest, 80% casein, constitutes major subclasses α- (α_{S1}- and α_{S2}-), β-, and κ-casein (Gellrich, Meyer, & Wiedemann, 2014).

3.1 Degree of hydrolysis (DH%) and weight average molecular weight (M_w)

In this study, two very different protease products were chosen based on their differences in the main mode of action during hydrolysis. Alcalase consists of mainly endopeptidases while Flavourzyme mainly contains exopeptidases (Merz, Claaßen, Appel, Berends, Rabe, Blank, et

al., 2016). The endopeptidase mode of action is to cut within the peptide chains, whereas exopeptidases cut at the very ends of peptide chains. As such, these are representative of the two main groups of proteases when investigating the development in DH% and M_w during the course of the hydrolysis reaction. DH% was measured using a TNBS method developed for 96 well plates. This type of method is well established and is commonly used to follow EPH reactions of food proteins (Rutherfurd, 2010; Spellman, McEvoy, O'Cuinn, & FitzGerald, 2003). The results displayed in Fig. 1A and 1B, show the development of DH% as a function of the hydrolysis time. One important observation from this data is the similarities of trajectories for all three raw materials hydrolyzed with Alcalase (Fig. 1A), despite an expected difference in the composition of the hydrolysates from the three different raw materials. This is related to the fact that DH%, while showing the overall reaction progress, does not reflect the actual composition of the hydrolysates at a given time. A direct comparison of two different hydrolysates from different batches of raw materials based on DH% alone is therefore inadequate. The addition of a complementary parameter reflecting the actual composition of the hydrolysates in terms of MWD could provide more comprehensive information of the EPH process.

In order to obtain this complementary process monitoring parameter to the DH%, all the hydrolysates were subjected to SEC analysis. Here, M_w derived from the SEC profiles was monitored as a function of hydrolysis time (Fig. 1C and 1D). M_w has previously been used as both a process monitoring and product quality parameter in EPH (Li, Wang, Chi, Gong, Luo, & Ding, 2013; Wubshet, et al., 2017). In contrast to the DH%, a clear distinction could be observed for M_w trajectories for hydrolysates of the different raw materials (see Fig. 1C and 1D). This shows that the combined use of DH% and M_w as process monitoring parameters not only reflects the hydrolysis progression, but also gives additional information of MWD during the course of hydrolysis.

3.2 DH% vs. molecular weight distribution

While DH% and M_w have been independently used to monitor proteolysis, the direct relationship between these two parameters has not been studied previously. In the present study, a general trend of an inverse correlation was observed when plotting M_w against DH% (Fig. 2), and each of the six hydrolysis reactions follows a different exponential decay trend. As previously described, this can also be explained by the fact that Alcalase mainly consists of endopeptidases which results in a relatively fast drop in M_w with increasing DH% in the start of the reaction. Flavourzyme on the other hand, consists mostly of exopeptidases which digest terminal peptide bonds, results in a slower M_w reduction. This difference can easily be observed at any specific DH% value (Fig. 2). Here, Flavourzyme samples always have higher M_w compared to the corresponding sample hydrolyzed with Alcalase e.g. WMP samples at DH% 10. At this DH% value WMPF has a M_w of about 5500 g/mol while WMPA has a M_w of approximately 2700 g/mol. This and the development of the parameters with time shows that relationships between DH% and M_w could provide important insight to the enzymatic mode of action (endopeptidase vs. exopeptidase activity) in a given reaction system. The vital observation in the data presented in Fig. 2 is the non-linear relationship between DH% and M_w. This is an important aspect to consider when using the two parameters interchangeably for monitoring an enzymatic protein hydrolysis. A video showing the data in 3D plot with time as a third dimension is presented in SI.

After observing the non-linear univariate correlation between M_w and DH% (Fig. 2), a multivariate correlation was assumed between the MWD profile and DH%. A PLSR model where the entire chromatographic profile was used as a predictor of DH%, was created. The results displayed in Fig. 3A and 3B show that it is possible to predict DH% from a size distribution profile with high accuracy (root mean square error of the cross-validated

(RMSECV) of 0.86 % and a coefficient of determination of the cross-validated (R^2) of 0.97). The general negative correlation with higher molecular weights and positive with the lower in the regression coefficient shows a close link between MWD derived from SEC chromatography and DH% (Fig. 3B). There are, however, some limitations related to the underestimation of the higher DH% values and the overestimation of the lower DH% values. This is most likely linked to the limitation of the column and instrumental setup for the SEC measurements. SEC columns in general have a limited exclusion range dependent on the mobile and stationary phase (Hong, Koza, & Bouvier, 2012). The BioSep-SEC-s2000 column has an exclusion range between 200-300000 Da (Ahmed & Modrek, 1992). This can result in errors in the measured MWD of samples containing larger amounts of molecules outside the exclusion range. Another major factor responsible for the limitation of the SEC measurements is the detection method used. UV detection at 214 nm will result in a systematic underestimation of free amino acids and overestimation of proteins and peptides (Kuipers & Gruppen, 2007). The limitation in exclusion range and the detection can therefore explain why the PLSR model curves at the lowest and highest DH% values.

3.3 DH% and M_w reproducibility studies

A validation study for the two methods used to measure DH% and M_w, i.e., TNBS and SEC, was performed to evaluate intra- and interday (n=6) reproducibility. A set of samples were evaluated from each of the six hydrolysis time series. The results presented in Table 1 show that the intra- and interday percentage relative standard deviation (%RSD) for DH% measurement ranges from 0.89 to 12.18, and 2.27 to 11.36, respectively. Likewise, the intra- and interday %RSD for the M_w measurements varies between 0.17 to 2.03, and 0.46 to 3.62, respectively. Hence, the DH% measured resulted in generally higher %RSD as compared to M_w. Technical errors and small changes in the reaction conditions are likely to be the major

reasons for the higher %RSD in DH% measurements. The non-systematic variation of the %RSD for the DH% in terms of enzyme type, raw material and hydrolysis time supports this. For M_w on the other hand, the %RSD shows a clear pattern in the intra- and interday measurements. The intraday variation in %RSD is generally lower as compared to the interday variation, explained by small changes in the SEC system over time (e.g. reduction in plate number).

3.4 Multivariate calibration of FTIR for prediction of DH% and M_w

Multivariate calibration models for predicting both DH% and M_w from FTIR spectra were developed. For the prediction of M_w, three outlying samples from EPH reactions using Flavourzyme were removed from the model. These three samples, all collected after two minutes of hydrolysis, contained a large proportion of proteins and peptides with high molecular weight. These molecules are outside the exclusion range of the column, as observed in a relative sharp increase in absorption at the void volume (approximately 6 min, see Fig. S-1 in SI). The calculated M_w value for these samples are therefore underestimated and not a good representation of the actual MWD.

The FTIR-based PLSR prediction models displayed in Fig. 4 provided prediction models with an RMSECV of 1.3 % for DH% and of 373 g/mol for M_w. These errors are low compared to the total variation range for both parameters (i.e., 6.1% for DH% and 7.7% for M_w). The R^2 obtained was 0.93 for DH% and 0.91 for M_w, which is also comparable to those previously reported using FTIR to predict both parameters in EPH samples (Poulsen, et al., 2016; Wubshet, et al., 2017). The prediction errors in the PLSR model for M_w are higher for the lower M_w values (Fig. 4B). This can be explained by the same factors discussed in section 3.2, where similar effects were observed in the PLSR prediction model of DH% when the whole SEC

chromatograms were used. This effect was not observed in the PLSR model where DH% was predicted from FTIR spectra, indicating that the limitations of predicting M_w are mostly related to the reference method.

The regression coefficients of the two PLSR models are provided in Fig. 5. From a comparison of the regression coefficients, both distinct differences and similarities are revealed. Firstly, the spectral region around the amide I (~1650 cm^{-1}) and amide II band (~1550 cm^{-1}) is the dominating feature in both PLSR models. In addition, relative to the M_w model, the regression coefficients of the DH% model show higher contributions of the NH_3^+ (~1516 cm^{-1}) and COO^- (~1400 cm^{-1}) bands. This is very interesting, and in accordance with the fact that the amount of N-terminals and C-terminals are directly proportional to the total cleaved peptide bonds (a principal phenomenon measured as DH%). The spectral region around the amide I and II bands contains information related to peptide backbone, and it has been shown that changes in this region during proteolysis can be used to predict M_w (Wubshet, et al., 2017). There are, however, some limitations as this region also contains more complex features related to protein and peptides secondary structures (Yang, Yang, Kong, Dong, & Yu, 2015). Different raw materials and the use of different enzymes in EPH processes will therefore result a high degree of variation in the amide region, which in turn will affect the robustness of a regression model. A way to overcome this challenge is to use a two-level PLSR model, where FTIR spectra are classified according to raw material and the enzyme used prior to regression tuned to specific raw materials and enzymes (Kristoffersen, Liland, Böcker, Wubshet, Lindberg, Horn, et al., 2019).

From the previous discussion, it is clear that DH% and M_w provide complementary information on EPH processes. Therefore, simultaneous measurement of the two provides a powerful analytical platform in process monitoring and product characterization. In the present work, in addition to studying the general correlation of DH% and M_w, multivariate models were

developed for prediction of DH% values directly from the MWD profiles. Such prediction models can provide an alternative means of obtaining DH% thereby avoiding the need to perform the lengthy derivatization experiment. The intercorrelation of DH% and M_w was studied and we attempted to predict the two parameters from dry-film FTIR fingerprints. Due to the inherent correlation of the two parameters, the PLSR models obtained from FTIR spectra are generally similar, with slight differences in the regression coefficients.

4 Conclusion

The aim of the study was to explore the relationship between two vital parameters of an EPH reaction (i.e., M_w and DH%) and develop dry-film FTIR based models for their rapid and simultaneous prediction. M_w and DH% are important quality parameters containing complimentary information and can therefore be used to adequately characterize the state of the protein hydrolysis reaction, both during the course of the process and in the final product. The results show that the spectral changes found in the FTIR region between (1800-700 cm^{-1}) during EPH reactions of milk proteins can be used to predict both parameters with a relatively high accuracy (RMSECV of 373 g/mol and 1.3 % for M_w and DH%, respectively). As the correlations are satisfactory, we conclude that dry-film FTIR is a promising tool for the prediction of both M_w and DH% in milk protein hydrolysates. The current study represents a first application of dry-film FTIR for the dual prediction of M_w and DH% of dairy protein hydrolysates. With growing interest in production of protein hydrolysates with high reproducibility and quality, for example hydrolysates intended for infant formulas, industrially relevant process monitoring tools are indispensable. Therefore, the demonstrated dry-film FTIR based process monitoring can further be developed to serve as on- or at-line process monitoring tool.

5 Acknowledgments

Anne-Grethe Johansen and TINE provided the dairy raw materials and Carl Emil Aae Eskildsen assisted in making the graphics. They are all greatly acknowledged for their contributions. Financial support from the Norwegian Fund for Research Fees for Agricultural Products through the projects "FoodSMaCk" (no. 262308) and "SunnMat" (no. 262300), the Norwegian Research Council through the project "Notably" (no. 280709) and "PepFishing" (no. 261849), and internal financing from Nofima through the project "PepTek" is also greatly acknowledged.

References

Adler-Nissen, J. (1979). Determination of the degree of hydrolysis of food protein hydrolysates by trinitrobenzenesulfonic acid. *Journal of Agricultural and Food Chemistry, 27*(6), 1256-1262.

Afseth, N. K., & Kohler, A. (2012). Extended multiplicative signal correction in vibrational spectroscopy, a tutorial. *Chemometrics and Intelligent Laboratory Systems, 117*, 92-99.

Ahmed, F., & Modrek, B. (1992). Biosep-SEC-S high-performance size-exclusion chromatographic columns for proteins and peptides. *Journal of Chromatography, 599*(1-2), 25-33.

Aspevik, T., Oterhals, Å., Rønning, S. B., Altintzoglou, T., Wubshet, S. G., Gildberg, A., Afseth, N. K., Whitaker, R. D., & Lindberg, D. (2017). Valorization of Proteins from Co- and By-Products from the Fish and Meat Industry. *Topics in Current Chemistry, 375*(3), 53.

Beaubier, S., Framboisier, X., Ioannou, I., Galet, O., & Kapel, R. (2019). Simultaneous quantification of the degree of hydrolysis, protein conversion rate and mean molar

406 weight of peptides released in the course of enzymatic proteolysis. *Journal of*
407 *Chromatography B, 1105*, 1-9.
408 Böcker, U., Wubshet, S. G., Lindberg, D., & Afseth, N. K. (2017). Fourier-transform infrared
409 spectroscopy for characterization of protein chain reductions in enzymatic reactions.
410 *Analyst, 142*(15), 2812-2818.
411 Chi, C. F., Cao, Z. H., Wang, B., Hu, F. Y., Li, Z. R., & Zhang, B. (2014). Antioxidant and
412 Functional Properties of Collagen Hydrolysates from Spanish Mackerel Skin as
413 Influenced by Average Molecular Weight. *Molecules, 19*(8), 11211-11230.
414 Church, F. C., Swaisgood, H. E., Porter, D. H., & Catignani, G. L. (1983). Spectrophotometric
415 Assay Using o-Phthaldialdehyde for Determination of Proteolysis in Milk and Isolated
416 Milk Proteins. *Journal of Dairy Science, 66*(6), 1219-1227.
417 Gellrich, K., Meyer, H. H. D., & Wiedemann, S. (2014). Composition of major proteins in cow
418 milk differing in mean protein concentration during the first 155 days of lactation and
419 the influence of season as well as short-term restricted feeding in early and mid-
420 lactation. *Czech Journal of Animal Science, 59*(3), 97-106.
421 Guler, G., Dzafic, E., Vorob'ev, M. M., Vogel, V., & Mantele, W. (2011). Real time observation
422 of proteolysis with Fourier transform infrared (FT-IR) and UV-circular dichroism
423 spectroscopy: Watching a protease eat a protein. *Spectrochimica Acta Part A:*
424 *Molecular and Biomolecular Spectroscopy, 79*(1), 104-111.
425 Guler, G., Vorob'ev, M. M., Vogel, V., & Mantele, W. (2016). Proteolytically-induced changes
426 of secondary structural protein conformation of bovine serum albumin monitored by
427 Fourier transform infrared (FT-IR) and UV-circular dichroism spectroscopy.
428 *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy, 161*, 8-18.
429 Hong, P., Koza, S., & Bouvier, E. S. P. (2012). A review Size-exclusion chromatography for
430 the analysis of protein biotherapeutics and their aggregates. *Journal of Liquid*
431 *Chromatography & Related Technologies, 35*(20), 2923-2950.
432 Hsieh, D. S. T., Lin, C., Lang, E. R., Catsimpoolas, N., & Rha, C. K. (1979). Molecular-weight
433 distribution of soybean globulin peptides produced by peptic hydrolysis. *Cereal*
434 *Chemistry, 56*(4), 227-231.
435 Kristoffersen, K. A., Liland, K. H., Böcker, U., Wubshet, S. G., Lindberg, D., Horn, S. J., &
436 Afseth, N. K. (2019). FTIR-based hierarchical modeling for prediction of average
437 molecular weights of protein hydrolysates. *Talanta, 205*, 120084.
438 Kuipers, B. J. H., & Gruppen, H. (2007). Prediction of molar extinction coefficients of proteins
439 and peptides using UV absorption of the constituent amino acids at 214 nm to enable
440 quantitative reverse phase high-performance liquid chromatography-mass spectrometry
441 analysis. *Journal of Agricultural and Food Chemistry, 55*(14), 5445-5451.
442 Li, Z., Wang, B., Chi, C., Gong, Y., Luo, H., & Ding, G. (2013). Influence of average molecular
443 weight on antioxidant and functional properties of cartilage collagen hydrolysates from
444 Sphyrna lewini, Dasyatis akjei and Raja porosa. *Food Research International, 51*(1),
445 283-293.
446 Lordan, S., Ross, R. P., & Stanton, C. (2011). Marine Bioactives as Functional Food
447 Ingredients: Potential to Reduce the Incidence of Chronic Diseases. *Marine Drugs, 9*(6),
448 1056-1100.
449 Martinez-Maqueda, D., Miralles, B., Recio, I., & Hernandez-Ledesma, B. (2012).
450 Antihypertensive peptides from food proteins: a review. *Food & Function, 3*(4), 350-
451 361.
452 Merz, M., Claaßen, W., Appel, D., Berends, P., Rabe, S., Blank, I., Stressler, T., & Fischer, L.
453 (2016). Characterization of commercially available peptidases in respect of the
454 production of protein hydrolysates with defined compositions using a three-step
455 methodology. *Journal of Molecular Catalysis B: Enzymatic, 127*, 1-10.

Poulsen, N. A., Eskildsen, C. E., Akkerman, M., Johansen, L. B., Hansen, M. S., Hansen, P. W., Skov, T., & Larsen, L. B. (2016). Predicting hydrolysis of whey protein by mid-infrared spectroscopy. *International Dairy Journal, 61*, 44-50.

Ruckebusch, C., Duponchel, L., & Huvenne, J. P. (2001). Degree of hydrolysis from mid-infrared spectra. *Analytica Chimica Acta, 446*(1-2), 257-268.

Ruckebusch, C., Duponchel, L., Huvenne, J. P., Legrand, P., Nedjar-Arroume, N., Lignot, B., Dhulster, P., & Guillochon, D. (1999). Hydrolysis of hemoglobin surveyed by infrared spectroscopy II. Progress predicted by chemometrics. *Analytica Chimica Acta, 396*(2-3), 241-251.

Ruckebusch, C., Sombret, B., Froidevaux, R., & Huvenne, J. P. (2001). On-line mid-infrared spectroscopic data and chemometrics for the monitoring of an enzymatic hydrolysis. *Applied Spectroscopy, 55*(12), 1610-1617.

Rutherfurd, S. M. (2010). Methodology for determining degree of hydrolysis of proteins in hydrolysates: A review. *Journal of AOAC International, 93*(5), 1515-1522.

Satake, K., Okuyama, T., Ohashi, M., & Shinoda, T. (1960). The spectrophotometric determination of amine, amino acid and peptide with 2,4,6-trinitrobenzene 1-sulfonic acid. *Journal of Biochemistry, 47*(5), 654-660.

Sbroggio, M. F., Montilha, M. S., de Figueiredo, V. R. G., Georgetti, S. R., & Kurozawa, L. E. (2016). Influence of the degree of hydrolysis and type of enzyme on antioxidant activity of okara protein hydrolysates. *Food Science and Technology, 36*(2), 375-381.

Simonne, A. H., Simonne, E. H., Eitenmiller, R. R., Mills, H. A., & Cresman, C. P. (1997). Could the dumas method replace the Kjeldahl digestion for nitrogen and crude protein determinations in foods? *Journal of the Science of Food and Agriculture, 73*(1), 39-45.

Slizyte, R., Mozuraityte, R., Martinez-Alvarez, O., Falch, E., Fouchereau-Peron, M., & Rustad, T. (2009). Functional, bioactive and antioxidative properties of hydrolysates obtained from cod (Gadus morhua) backbones. *Process Biochemistry, 44*(6), 668-677.

Spellman, D., McEvoy, E., O'Cuinn, G., & FitzGerald, R. J. (2003). Proteinase and exopeptidase hydrolysis of whey protein: Comparison of the TNBS, OPA and pH stat methods for quantification of degree of hydrolysis. *International Dairy Journal, 13*(6), 447-453.

Vander Heyden, Y., Popovici, S. T., & Schoenmakers, P. J. (2002). Evaluation of size-exclusion chromatography and size-exclusion electrochromatography calibration curves. *Journal of Chromatography A, 957*(2), 127-137.

Wubshet, S. G., Mage, I., Böcker, U., Lindberg, D., Knutsen, S. H., Rieder, A., Rodriguez, D. A., & Afseth, N. K. (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.

Yang, H., Yang, S., Kong, J., Dong, A., & Yu, S. (2015). Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. *Nature Protocols, 10*, 382.

Fig. 1. DH% development (over time) for hydrolysis reactions of three different substrates using A) Alcalase and B) Flavourzyme. M_w development (over time) for hydrolysis reactions of three different substrates using C) Alcalase and D) Flavourzyme.

Fig. 2. M_w vs DH% for six hydrolysis reactions produced using two different enzymes.

Fig. 3. A) PLSR correlation plot of DH% for six hydrolysis reactions. Predicted from area normalized and mid-centered chromatograms. B) The normalized regression coefficients of the chromatograms for the PLSR model.

Fig. 4. A) PLSR correlation plot DH%, 60 samples B) PLSR correlation plot of M_w, 57 samples.

Fig. 5. Normalized regression coefficients for PLSR model of DH% (Fig. 5A) and M_w (Fig. 5B).

Table 1. An overview of %RSD for the intra- and interday validation study of the methods used to measure DH% and M_w

%RSD	Hydrolysis time (min)	Intraday (n=6)		Interday (n=6)	
		DH%	M_w	DH%	M_w
WPC80 5 A	5	6.31	0.75	4.70	2.48
WPC80 30 A	30	3.80	1.24	6.08	1.53
WPC80 120 A	120	2.74	0.76[a]	3.54	1.53
WPC80 5 F	5	6.01	0.17	3.94	0.46
WPC80 30 F	30	3.68	0.39	3.67	2.29
WPC80 120 F	120	0.89	0.21	2.79	3.14
WMP 5 A	5	12.18	0.67	11.36	1.21
WMP 30 A	30	4.77	0.49	7.90	1.20
WMP 120 A	120	3.00	0.37	6.25	0.91
WMP 5 F	5	6.79	0.43	5.98	3.62
WMP 30 F	30	3.83	0.66	5.15	1.44
WMP 120 F	120	2.13	1.43	4.20	1.34
MPO 5 A	5	9.21	0.75	8.68	1.74
MPO 30 A	30	5.86	0.76	6.14	1.29
MPO 120 A	120	6.13	0.96	7.94	2.55
MPO 5 F	5	6.53	2.03	8.69	3.28
MPO 30 F	30	2.19	1.13	9.08	1.76
MPO 120 F	120	2.34	1.16	2.27	1.96

[a]One of the measurements were removed due to injection error (n=5).