Manuscript after review

Access to published version: https://www.sciencedirect.com/science/article/pii/S0023643820315802

Raman, near-infrared and fluorescence spectroscopy for determination of collagen content in ground meat and poultry by-products

Olga Monago-Maraña[1]*, Jens Petter Wold[1], Rune Rødbotten[1], Katinka Riiser Dankel[1], Nils Kristian Afseth[1]

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

*corresponding author: olgamonago@gmail.com

+34675673233

Abstract

Raman, near-infrared and fluorescence spectroscopy were evaluated for determination of collagen content in ground meat. Two sample sets were used (i.e. ground beef and ground poultry by-products), and collagen concentrations (measured as hydroxyproline) varied in the ranges 0.1 - 3.3% in the beef samples and 0.4 - 1.5% in the poultry samples. Similar validation results for hydroxyproline were obtained for NIRS (R^2 = 0.82 and RMSECV = 0.11%) and Raman (R^2 = 0.81 and RMSECV = 0.11%) for the poultry samples. For the beef samples, NIRS obtained slightly less accurate results (R^2 = 0.89, RMSECV= 0.25%) compared to Raman (R^2 = 0.94, RMSECV= 0.19%), most likely due to less representative sampling. Fluorescence spectroscopy gave higher prediction errors (RMSECV= 0.50% and 0.13% for beef and poultry, respectively). This shows that Raman spectroscopy employing a scanning approach for representative sampling is a potential tool for on-line determination of collagen in meat.

Keywords: Raman; NIR; fluorescence; collagen; ground meat

1. Introduction

Collagen is the most abundant mammalian and avian fibrous protein. It is predominantly located in the skin (or hide), tendons and bones. Different types of collagen are distinguished by their amino acid composition, with collagen type I-IV being the most abundant. The collagen triple-helix presents a conformation consisting of glycine-X-Y repeating sequences. The X and Y positions can accommodate any amino acid in order to form a stable triple-helix. However, when proline and hydroxyproline are situated in the X and Y positions, respectively, this sequence is the most stabilizing and most commonly found tripeptide unit present in collagen (Persikov, Ramshaw, Kirkpatrick, & Brodsky, 2000). In meat, collagen contributes to quality parameters such as tenderness, texture and sensory properties. In addition, bioprocessing of by-products from fish and poultry is a growing industry (Aspevik et al., 2017), and collagen is an interesting target protein for a range of different markets, from food ingredients to cosmetics (Gomez-Guillen, Gimenez, Lopez-Caballero, & Montero, 2011). Thus, there is a high interest in developing tools for rapid determination of collagen in meat.

The traditional methods for determination of collagen in meat are destructive and time consuming, usually involving the quantification of hydroxyproline by colorimetric (Kolar, 1990) or chromatographic (Colgrave, Allingham, & Jones, 2008) methods after complete proteolysis. Spectroscopic methods, on the other hand, offer fast and non-invasive measurements and can enable effective quality differentiation and process control (Beganovic, Hawthorne, Bach, & Huck, 2019). Near-infrared spectroscopy (NIRS) is one of the most frequently used non-destructive techniques in the meat industries, and NIRS has also been used for determination of collagen (measured as hydroxyproline) in meat. However, in several studies, unsatisfactory prediction results have been found for ground beef and ovine meat (R^2 in the range 0.18 - 0.55) (Alomar, Gallo, Castañeda, & Fuchslocher, 2003; Prieto, Andrés, Giráldez, Mantecón, & Lavín, 2006; Young, Barker, & Frost, 1996). NIRS have been evaluated for the quantification of hydroxyproline in cured pork sausages and dry cured beef with better results (R^2 = 0.77 and standard error of prediction of 0.05%) (González-Martín, Bermejo, Hierro, & González, 2009).

Recently, other authors used this technique to classify sous-vide loins as a function of time of cooking and predicted texture-related parameters of the samples (including hydroxyproline) with R^2 of 0.92 and mean absolute scaled error (MASE) of 0.19 (Perez-Palacios, Caballero, González-Mohíno, Mir-Bel, & Antequera, 2019). In that study, the hydroxyproline concentration range was larger (2.0 - 4.5%) compared to the previously mentioned studies, which probably improved the results.

The rationale behind using fluorescence spectroscopy for determination of collagen is related to the fact that several components present in connective tissue, like collagen crosslinks and components such as pyridinoline and pentosidine, have fluorescing properties (A. J. Bailey, Sims, Avery, & Halligan, 1995; J. Bailey & Light, 1989). Wold et al. (1999) determined hydroxyproline in intact slices of beef, with moderate results due to a narrow range of hydroxyproline (0.4 - 0.9%) (Jens Petter Wold, Kvaal, & Egelandsdal, 1999). However, a considerable improvement was obtained when the range was expanded (0.72 - 7.12%) and samples were homogenized ($R^2 = 0.94$, RMSECV= 0.37%) (J. P. Wold, Lundby, & Egelandsdal, 1999). The potential of fluorescence was further elucidated for the quantification of hydroxyproline in sausage batters (beef and pork) with a large variation in myoglobin content (Egelandsdal, Dingstad, Tøgersen, Lundby, & Langsrud, 2005). The concentration of myoglobin largely affects both the intensity and shape of the fluorescence spectra, and it turned out that prediction errors were reduced slightly when spectra were normalized by multiplying them by a^*, i.e. the measured redness of the samples. The authors found lower prediction errors for fluorescence (0.48%) than for NIRS (0.64%).

Raman spectroscopy has the potential to provide detailed chemical information on protein composition and protein structure (Herrero, 2008). The technique has been employed to obtain biochemical fingerprints of collagen fibers in native aortic heart valve tissues and to monitor the increasing damage of collagen fibers (Votteler et al., 2012). Type I and type IV collagens were characterized by Raman spectroscopy in order to study the relation between aging and cancer progression (Nguyen et al., 2012). Collagen was also quantified in native and engineered cartilage tissues with good results (R^2 = 0.84) (Bergholt, Albro, & Stevens, 2017). Also, Raman

77 spectroscopy has been used for the characterization of structural changes in collagen, which
78 allows a more thorough understanding of disease progression (Martinez, Bullock, MacNeil, &
79 Rehman, 2019). But despite its use in medical diagnostics, only one study reports the
80 determination of collagen (as hydroxyproline) in meat using Raman spectroscopy (Nian et al.,
81 2017). The authors obtained good results ($R^2 = 0.79$, RMSECV = 0.07%). However, interpretation
82 of the regression models reveals that some of the main spectral features used for determination of
83 hydroxyproline was found in a spectral region with no known spectral information related to
84 proteins (i.e. the spectral region between 1800 and 2800 cm^{-1}).

85 Due to the inherent heterogeneity of foods, representative sampling is always a crucial factor in
86 food analysis. Thus, the main objective of this work was to elucidate the feasibility of Raman
87 spectroscopy for rapid and non-destructive quantification of collagen in ground meat using a
88 Raman system equipped with a large volume probe. Two different sample sets were used for this
89 purpose: 1) samples of ground beef, homogenized in the laboratory and 2) samples of poultry by-
90 products, industrially ground resulting in less homogeneous samples. A process Raman
91 instrument was used in scanning mode for all analysis, and to the authors knowledge, this is the
92 first time that a large volume Raman probe was used for this purpose. For both sample sets, the
93 performance of Raman spectroscopy was compared with that of NIRS and fluorescence
94 spectroscopy.

95

96 **2. Materials and methods**

97 **2.1. Samples**

98 **2.1.1. Beef samples**

99 In order to expand the possible range of collagen and other components in beef samples, a mixture
100 of different starting materials was used to make 60 different samples. Different kinds of fats,
101 muscles and tendons were obtained from a commercial slaughterhouse (Furuseth AS, Dal,
102 Norway) and food grade collagen powder usually used for dry sausages, emulsified product etc.

(Collapro Bovine Standard) was supplied from Hulshof Protein Technology, Lichtenvoorde, The Netherlands. Different amounts of the ingredients were blended to obtain a wide range of collagen, fat and total protein content in the sample set. The design of samples can be found in the supplementary material (Table S1). Samples were ground in a laboratory blender to get samples as homogeneous as possible. A total of 60 samples (400 g each) were obtained. Before spectroscopic measurements, the samples were shaped flat with a surface of approximately 160 cm^2 and a thickness of approximately 2 cm.

2.1.2. Poultry by-products

The poultry by-product sample material was collected from a poultry processing plant (Bioco, Nortura Hærland, Østfold, Norway). Five by-product fractions were selected, including chicken skin, and carcasses from both chicken and turkey, before and after mechanical deboning, respectively. In addition to the pure fractions, the remaining of the 52 samples were prepared by manually combining 25%, 50% or 75% of the by-product fractions in a range of possible manners (excluding the 50% -25% -25% versions). The design of samples can be found in the supplementary material (Table S2). The samples were ground on-site and immediately measured with NIRS. 400 g of the sample materials were shaped in the same way as the beef samples and stored at 4 ºC until further analysis by Raman and fluorescence spectroscopy.

2.1.3. Pure turkey collagen

A collagen reference sample was extracted from turkey tendons following a literature procedure (Grønlien et al., 2019). A Raman spectrum was recorded from this sample for comparison with Raman spectra from more complex beef and poultry samples.

2.2. References measurements (Percentage of protein, hydroxyproline and fat)

The references measurements were performed at an external laboratory (ALS laboratory). Two parallels from each sample were analyzed. Dumas method (Dumas, 1826) was used for total N and protein content was determined as 6.25*N-total. An established spectrophotometric method

was used for quantifying the hydroxyproline percentage (BS 4401-11:1995, ISO 3496:1994), generally used as analytical criterion to assess the amount of collagen. In the case of fat content, an internal method at the ALS laboratory was used based on pulsed nuclear magnetic resonance (NMR). Samples were dried in an oven to determine the moisture content. After that, samples were stabilized at 50 ºC and resonance of samples were determined. The fat content was determined automatically by comparing the resonance of the sample with a calibration curve established using a certified olive oil content.

2.3. Spectroscopic measurements

The Raman spectra were collected with a RamanRXN2^{TM} Hybrid system equipped with a non-contact PhAT-probe (Kaiser Optical Systems, Inc., Ann Arbor, MI, USA). A laser with a 785 nm excitation wavelength, and a circular spot size of D = 6 mm at a 25 cm working distance was used. The spectral range was 300-1890 cm^{-1}. Each spectrum was an average of 4×20 sec accumulations. All measurements were performed by moving the samples manually under the laser beam, assuring that large parts of the sample surface were probed. The purpose of this procedure was to obtain representative sampling of the inhomogeneous samples. Each sample was measured in triplicate and the average spectrum was obtained after fluorescence background correction.

For practical reasons, two different instruments were used for the NIRS measurements. The beef samples were measured using a FOSS NIRSystems XDS Optiprobe Analyzer^{TM} (FOSS Analytical A/S, Hillerød, Denmark). Using a fibre-optic probe, the measurements were done in reflectance mode with a spectral range of 400-2500 nm and a resolution of 0.5 nm. The spectra were transformed from reflectance to absorbance units ($A = \log_{10}(1/R)$). The probe head was positioned so that it was in contact with the sample surfaces. Replicate spectra from each sample were acquired in five different spots (D = 1 cm) and the average spectrum was used for further analysis.

In the case of the poultry by-products samples, a Perten DA7440 Process NIR Sensor (Perten Instruments, a PerkinElmer Company, USA) was used to obtain spectra in reflection mode at a 25 cm working distance. The spectral range was 950-1650 nm with a resolution of 5 nm. The

156 samples were spread out on a board and each spectrum was acquired as an average of 10 seconds
157 of acquisition while the samples were moved manually under the spectrometer to scan most of
158 the sample surface. The spectra were transformed from reflectance to absorbance units. Three
159 replicate spectra were obtained for each sample and a different surface was scanned each time to
160 obtain a representative sample spectrum. The average spectrum was used for further analysis.

161 The fluorescence emission spectra were measured in front-face mode using a Fluoromax-4
162 spectrofluorometer (Horiba Scientific, Kyoto, Japan) equipped with a FL-300/FM4-3000
163 bifurcated fiber-optic probe. The probe head was positioned 5 cm from the sample surface to
164 create a 4 cm measurement area. The probe and the sample were shielded from ambient light. The
165 excitation wavelength was set at 340 nm and emission spectra were recorded in the range 360 -
166 600 nm at every 4 nm. The excitation and emission slit widths were 5 nm. For each sample,
167 replicate spectra were recorded in five different spots on the sample surface and the average
168 spectrum was used for further analyses.

169
170 **2.4. Spectral pre-processing and data analysis**

171 The fluorescence background in the Raman spectra was removed from the raw spectra by
172 applying a commonly used background correction approach based on fitting a polynomial to the
173 baseline (Lieber & Mahadevan-Jansen, 2003). The procedure was applied to the range 476 - 1890
174 cm^{-1}. A polynomial degree of 4 was used. The correction was performed using in-house adapted
175 automated Matlab scripts (R2007b, The MathWorks, Inc., Natick, MA, USA).

176 NIR and fluorescence spectra were normalized using standard normal variate (SNV) (Barnes,
177 Dhanoa, & Lister, 1989). For the NIR spectra of the beef samples, a data reduction was performed
178 so that the spectral range and resolution would be identical to that of the poultry by-products
179 samples, i.e. a 950-1650 nm range and a 5 nm resolution.

180 Calibration models were based on partial least-squares regression (PLSR) (Martens & Naes,
181 1989). Full cross-validation was used to determine the number of components to use in the
182 calibration and to evaluate the performance of the models. The SNV pre-processing and

multivariate calibrations were performed using The Unscrambler version 6.11 (CAMO Software AS, Oslo, Norway).

3. Results and discussion

3.1. Sample gross composition

For all samples included in this study (n = 112), collagen (measured as hydroxyproline content), protein and fat contents were determined as percentage of wet weight. An overview of the gross composition of the samples is provided in Table 1. The beef samples span a wider range of hydroxyproline content compared to the poultry by-product samples.

Correlation coefficients (Pearson's r) between the different chemical components are also provided in Table 1. A moderate positive correlation between protein and hydroxyproline content is seen in both sample sets, which is reasonable since collagen is part of the total protein content. A weak negative correlation is seen between protein and fat content in the beef samples. For the poultry by-products, however, a stronger negative correlation between protein and fat is seen. This is due to the fact that the proportion of chicken skin in the samples, which are high in fat and low in protein content, is responsible for the main variation of the fat and protein content in the data set. Finally, extremely weak correlations are seen between hydroxyproline and fat content in both sample sets.

3.2. Spectral information

3.2.1. Raman

Baseline-corrected Raman spectra are presented in Figure 1. As expected, the spectra are dominated by signals originating from fat, with strong Raman bands at 1062, 1129, 1268 (=CH bending, scissoring), 1300 (C-H bending, stretching), 1442 (C-H bending, scissoring), 1655 (C=C stretching) and 1742 (RC=OOR, C=O stretching) cm^{-1}. Some of these bands are related to saturated fatty acids or ester groups (1300, 1442 and 1744 cm^{-1}) whereas others are related to unsaturated fatty acids (1655 and 1268 cm^{-1}) (Lee et al., 2018). No clear visible trend in the spectra

was observed according to the contents of hydroxyproline, which was as expected since hydroxyproline was low in concentration, and since hydroxyproline is a relatively weak Raman scatterer compared to fat. In addition, the collagen bands were partly overlapped by the fat bands.

Clear differences between beef and poultry by-products were observed for the bands 960, 970 and 1269 cm^{-1}, all of which were more pronounced in the poultry by-products. 960 cm^{-1} is the phosphate band (ν_1 PO_4^{3-}), and stems from bone residue (Wubshet, Wold, Böcker, Sanden, & Afseth, 2019). The bands at 970 and 1269 cm^{-1} can be assigned to unsaturated fatty acids and the degree of fatty acid unsaturation (Lee et al., 2018).

3.2.2. NIR

Preprocessed absorption spectra from beef (upper panel) and poultry by-products (lower panel) samples are shown in Figure 2. For both sample sets, the main bands appeared at 1200 and 1450 nm, where water, protein and fat bands overlap. A clear trend for fat content was observed around 1200 nm in both sets (Figure 2A and 2B) assigned to the second overtone of C-H stretching of several chemical groups (-CH_2, -CH_3, -CH=CH-) (Hourant, Baeten, Morales, Meurens, & Aparicio, 2000). No spectral variation due to hydroxyproline contents could be detected by visual inspection of the NIR spectra colored according to hydroxyproline content (Figure 2C and 2D).

3.2.3. Fluorescence spectra

Figure 3 shows the pre-processed fluorescence spectra for beef (upper panel) and poultry by-products (lower panel) samples. For excitation at 340 nm, collagen has a broad emission band peaking at about 400 nm (Wagnières, Star, & Wilson, 1998). It is also well known that the myoglobin in meat reabsorbs the created fluorescence, and valleys therefore appear in the fluorescence spectra at wavelengths where myoglobin has absorption peaks (Egelandsdal et al., 2005). This explains the valleys at around 410, 548 and 579 nm. The position of the myoglobin absorption peak at around 410 nm shifts according to exposure to oxygen and might explain the observed shifts in the corresponding valley in the beef spectra. The reabsorption of myoglobin makes the fluorescence spectra rather complex, with low intensity for samples with much

myoglobin and very strong intensity for samples with little myoglobin. These intensity differences were removed by pre-processing and are not visible in Figure 3. For the beef samples, no clear tendency was observed due to differences in fat and hydroxyproline content. There was a slight shift in the spectra from 450 to 440 nm for those samples with higher hydroxyproline content, but this can be related to the concentration of myoglobin. Maxima for collagen around 390 nm and 450 nm were reported by Wold et al., 1999 (J. P. Wold et al., 1999), and these can be seen quite clearly in the spectra from poultry by-products (Figure 3D). Fat, or adipose tissue, has an emission peak around 475 nm (J. P. Wold et al., 1999), and this can be seen as a shoulder in the region 475 – 525 nm in the spectra from fatty poultry by-products samples (Figure 3B). The origin of this fluorescence is not certain. The cofactor NADH could be a candidate, but this would fade over time (Wu, Dahlberg, Gao, Smith, & Bailin, 2019) and introduce instability in the system. NADH would also be found in other cellular tissues. Lipo-pigments fluoresce in the range 500-600 nm but are mainly products of lipid oxidation and the presence in fresh meat is therefore less likely. The very fat poultry samples contained much poultry skin, and skin contains elastin, which also has a strong fluorescence peaking at 410 nm.

3.3. Regression analysis

The descriptive statistics of the different regression models based on Raman spectra are presented in Table 2. Corresponding “predicted vs. reference"-plots are provided in the supplementary material. High coefficients of determination (R^2) were obtained between measured and estimated hydroxyproline for both the beef samples and the poultry by-products. Lower prediction errors (RMSECV) were obtained for the poultry by-products. The comparatively lower R^2 for the poultry by-products model was ascribed to a narrower range of hydroxyproline concentrations.

The regression coefficients for the models are shown in Figure 4A. Ideally, for simple regression models, coefficients with high values should correspond with spectral bands that carry information about the target component. In this case, some highlighted coefficients clearly correspond to Raman bands from collagen (Figure 4B) extracted from turkey. These peaks were found at 855 (proline ring), 877 (hydroxyproline ring), 922 (proline ring), 936 (C–C stretching

vibration of the backbone formed by the glycine-X-Y sequences), 1004 (phenylalanine), 1031 (phenylalanine), 1242 (Amide III) and 1670 (Amide I band) cm^{-1}. These peaks have previously been identified in Raman spectra from collagen type I and type IV (Herrero, 2008; Nguyen et al., 2012). Furthermore, some negative peaks in the regression coefficients found at 1303, 1438 and 1652 cm^{-1} are associated with fatty acid chains. Even though the correlations between hydroxyproline and fat in these data sets are very weak (as shown in Table 1), the peaks assigned to fat could turn out negative due to the simple fact that the fat peaks dominate the spectra and that they are not related to the contents of hydroxyproline. To verify that these Raman bands did not influence the model, the variables were removed, and new models were obtained (R^2 = 0.92 and 0.82, and RMSECV = 0.20% and 0.11%, for beef samples and poultry by-products, respectively). Since similar results were obtained, this indicates that these fat peaks were not needed to model the collagen content. Table 2 also shows that combining the two data sets into one regression model was possible, providing good results with similar number of PLSR components as for the beef samples. Figure 4A also shows that the regression coefficients of the combined data sets model were comparable to those of the individual models, suggesting a certain robustness of the Raman approach across different species.

Since moderate positive correlations between protein and hydroxyproline content were seen in both data sets, it was important to assure that the calibrations for collagen did not rely on the total protein content. One way of studying this is by investigating the correlations between the predicted values for protein and hydroxyproline contents, respectively (Eskildsen, Næs, Wold, Afseth, & Engelsen, 2019). Thus, PLSR models for protein were obtained, and the correlation coefficients between predicted protein and predicted hydroxyproline was calculated (r= 0.64 and r= 0.78 for beef and poultry by-products, respectively). Since these values are close to the correlations reported in Table 1, we presume that it is possible to predict hydroxyproline independently of changes in protein content.

Results for calibrations based on NIR and fluorescence are presented in Table 2, and the corresponding regression coefficients are provided in Figure 4. NIRS gave quite similar results

for poultry by-products as obtained by Raman spectroscopy. In the case of the beef samples, higher prediction errors were obtained with NIRS than with Raman spectroscopy and also more components were needed to obtain a good model. The slightly poorer result for beef could rely on less representative sampling, since a laboratory system with a fiber-optic probe, with a limited sampling spot size, was used. NIRS results were comparable to other studies on the determination of collagen in pork sausages (González-Martín et al., 2009) and pork loins (Perez-Palacios et al., 2019). The regression coefficients obtained for beef and poultry by-products were quite different from each other, suggesting that NIRS models for the prediction of collagen content are more difficult to use across different species compared to Raman spectroscopy models. Due to the broad NIR bands in the regression coefficients, it was also difficult to make any conclusive interpretations. The correlation coefficients obtained between predicted protein and predicted hydroxyproline from NIRS models were $r = 0.62$ and $r = 0.78$ for beef and poultry by-products, respectively. As in the case of Raman, it is thus reasonable to assume that it is possible to predict hydroxyproline content independently of variations in protein contents.

In the case of fluorescence, the results for beef were not as good as previous work on ground beef and sausage batter, where prediction errors of 0.37% and 0.48%, respectively, were obtained (Egelandsdal et al., 2005; J. P. Wold et al., 1999). A difficulty in this study was the large color difference within the beef samples, spanning from red meat to almost white tissue consisting of mainly fat and connective tissue. As pointed out above, this color variation results in complex spectra largely affected by myoglobin in both shape and intensity, and the close relation to the collagen content is lost. This was partly confirmed when five beef samples, white colored and with very high fat contents were omitted from the data set and the RMSECV was reduced to 0.39%.

The fluorescence results for the poultry samples were better and not that far from the results for NIRS and Raman. Although these samples were very heterogeneous, the color variations were not as pronounced as in the beef samples. The fat contents in these samples were also lower than in the beef, making them less complex. A disadvantage with the fluorescence measurement of the

poultry samples compared to Raman and NIR was that a rather limited part of the samples was measured, i.e. only five small regions. This could result in less representative measurements and a reduced match with the reference values. The regression coefficients were different between the two sample sets (Figure 4D) due to different spectral properties. Therefore, it did not make sense to make a combined model. Due to the complexity of the spectra and some uncertainty with regards to the present fluorophores, it is difficult to interpret the shape of the regression vectors. Weaker correlation was observed between predicted protein and predicted hydroxyproline (i.e. r = 0.19 and r = 0.68 for beef and poultry by-products, respectively). As in the case of Raman and NIRS, it is thus possible to predict hydroxyproline content independently of variations in protein contents. However, it is important to note that protein models obtained with fluorescence had much higher RMSECV compared to Raman and NIRS in the first place.

The samples of the present study were made to span the range of collagen contents, resulting in a slightly wider range of collagen than is normally encountered in industrial samples. However, the results clearly provide relevant knowledge on which spectroscopic methods that are feasible for collagen determination in foods. Due to rather small sample sets, all regression models were validated using full cross-validation, which is normally regarded as a rather optimistic validation approach. All regression models were cross validated with different validation segment sizes in order to test robustness, and similar results were obtained for all regression models (not shown), showing that relevant conclusions can be drawn based on the presented results.

Fluorescence spectroscopy provided the poorest regression results of the three techniques in the study. This can partly be attributed to the color variations seen in the sample sets. For Raman and NIRS, similar regression results were obtained, with Raman providing slightly better results for the beef samples. It is interesting to note that Raman spectroscopy is the only of the three techniques that can provide direct information on hydroxyproline, however, the Raman regression coefficients also show that other protein-related bands are important in the regression models. This could anyway contribute to explain why Raman seems to be better than NIRS for providing generic regression models for collagen contents across different species. Finally, based on the

present results, state-of-the-art representative sampling approaches such as large volume probes seem to enable quantitative Raman analysis of heterogenous food samples. This should encourage the future industrial use of Raman spectroscopy for food analysis.

4. Conclusions

This study demonstrates the potential of Raman, NIRS and fluorescence spectroscopy for rapid and non-destructive determination of collagen in different types of ground meat. Fluorescence spectroscopy is a very sensitive method, but this study shows that the signals are easily distorted by reabsorption by pigments in varying concentrations. These distortions are not easy to correct for and make the method less robust and accurate for determination of collagen. NIR spectroscopy performs well, however, the obtained data suggest that NIR models of collagen are more difficult to use across different species. The regression models for Raman spectroscopy were good with low prediction errors, and the models were easy to interpret, clearly highlighting spectral bands associated with collagen. This shows that the scanning approach presently used for covering a larger part of the sample makes Raman spectroscopy a potential tool for on-line determination of collagen in meat.

Conflict of interest

The authors declare that there is no conflict of interest.

Acknowledgement

This work was partially funded by the Norwegian Research Council through the projects Innovative and Flexible Food Processing Technology in Norway, iProcess (No. 255596/E59) and Novel cascade technology for optimal utilization of animal and marine by-products, Notably (No. 280709/E50); and by the Norwegian Agricultural Food Research Foundation through the projects FoodSMaCK - Spectroscopy, Modelling & Consumer Knowledge (No. 262308/F40) and Smart sensor and optimization systems for future food biorefineries, SmartBio (No. 282466/E50).

Olga Monago Maraña thanks to the Fundación Ramón Areces for a postdoctoral fellowship for studies abroad in the field of Life and Matter Sciences (XXXI edition of grants, 2019/2020) to support her postdoctoral studies at Nofima, Ås, Norway.

References

Alomar, D., Gallo, C., Castañeda, M., & Fuchslocher, R. (2003). Chemical and discriminant analysis of bovine meat by near infrared reflectance spectroscopy (NIRS). *Meat Science*, *63*(4), 441–450.

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

Bailey, A. J., Sims, T. J., Avery, N. C., & Halligan, E. P. (1995). Non-enzymic glycation of fibrous collagen: Reaction produces of glucose and ribose. *Biochemical Journal*, *305*, 385–390.

Bailey, J., & Light, N. D. (1989). *Connective tissue in meat and meat products*. London: Elsevier Science Publishers.

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, *43*(5), 772–777. https://doi.org/10.1366/0003702894202201

Beganovic, A., Hawthorne, L. M., Bach, K., & Huck, C. W. (2019). Critical review on the utilization of handheld and portable Raman spectrometry in meat science. *Foods*, *8*(2).

Bergholt, M. S., Albro, M. B., & Stevens, M. M. (2017). Online quantitative monitoring of live cell engineered cartilage growth using diffuse fiber-optic Raman spectroscopy. *Biomaterials*, *140*, 128–137.

Colgrave, M. L., Allingham, P. G., & Jones, A. (2008). Hydroxyproline quantification for the estimation of collagen in tissue using multiple reaction monitoring mass spectrometry. *Journal of Chromatography A*, *1212*, 150–153.

Dumas, J. B. (1826). Memorie sur quelques points de la Theorie atomistique. *Annales of Chimie*, *33*, 342.

Egelandsdal, B., Dingstad, G., Tøgersen, G., Lundby, F., & Langsrud, Ø. (2005). Autofluorescence quantifies collagen in sausage batters with a large variation in myoglobin content. *Meat Science*, *69*(1), 35–46.

Eskildsen, C., Næs, T., Wold, J., Afseth, N., & Engelsen, S. (2019). Visualizing indirect correlations when predicting fatty acid composition from near infrared spectroscopy measurements. *Proceedings of the 18th International Conference on Near Infrared Spectroscopy*, 39–44.

Gomez-Guillen, M. C., Gimenez, B., Lopez-Caballero, M. E., & Montero, M. P. (2011).

Functional and bioactive properties of collagen and gelatin from alternative sources: A review. *Food Hydrocolloids*, *25*(8), 1813–1827.

González-Martín, M. I., Bermejo, C. F., Hierro, J. M. H., & González, C. I. S. (2009). Determination of hydroxyproline in cured pork sausages and dry cured beef products by NIRS technology employing a fibre-optic probe. *Food Control*, *20*(8), 752–755.

Grønlien, K. G., Pedersen, M. E., Sanden, K. W., Høst, V., Karlsen, J., & Tønnesen, H. H. (2019). Collagen from Turkey (Meleagris gallopavo) tendon: A promising sustainable biomaterial for pharmaceutical use. *Sustainable Chemistry and Pharmacy*, *13*.

Herrero, A. M. (2008). Raman spectroscopy for monitoring protein structure in muscle food systems. *Critical Reviews in Food Science and Nutrition*, *48*(6), 512–523.

Hourant, P., Baeten, V., Morales, M. T., Meurens, M., & Aparicio, R. (2000). Oil and fat classification by selected bands of near-infrared spectroscopy. *Applied Spectroscopy*, *54*, 1168–1174.

Kolar, K. (1990). Colorimetric determination of hydroxyproline as measure of collagen content in meat and meat products: NMKL collaborative study. *Journal of the Association of Official Analytical Chemists*, *73*(1), 54–57.

Lee, J. Y., Park, J. H., Mun, H., Shim, W. B., Lim, S. H., & Kim, M. G. (2018). Quantitative analysis of lard in animal fat mixture using visible Raman spectroscopy. *Food Chemistry*, *254*(August 2017), 109–114.

Lieber, C. A., & Mahadevan-Jansen, A. (2003). Automated method for subtraction of fluorescence from biological Raman spectra. *Applied Spectroscopy*, *57*(11), 1363–1367.

Martens, H., & Naes, T. (1989). *Multivariate Calibration*. New York: Wiley.

Martinez, M. G., Bullock, A. J., MacNeil, S., & Rehman, I. U. (2019). Characterisation of structural changes in collagen with Raman spectroscopy. *Applied Spectroscopy Reviews*, *54*(6), 509–542.

Nguyen, T. T., Gobinet, C., Feru, J., Brassart-Pasco, S., Manfait, M., & Piot, O. (2012). Characterization of type i and IV collagens by Raman microspectroscopy: Identification of spectral markers of the dermo-epidermal junction. *Spectroscopy: An International Journal*, *25*, 421–427.

Nian, Y., Zhao, M., O'Donnell, C. P., Downey, G., Kerry, J. P., & Allen, P. (2017). Assessment of physico-chemical traits related to eating quality of young dairy bull beef at different ageing times using Raman spectroscopy and chemometrics. *Food Research International*, *99*, 778–789.

Perez-Palacios, T., Caballero, D., González-Mohíno, A., Mir-Bel, J., & Antequera, T. (2019). Near Infrared Reflectance spectroscopy to analyse texture related characteristics of sous vide pork loin. *Journal of Food Engineering*, *263*(February), 417–423.

Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A., & Brodsky, B. (2000). Amino acid propensities for the collagen triple-helix. *Biochemistry*, *39*(48), 14960–14967.

Prieto, N., Andrés, S., Giráldez, F. J., Mantecón, A. R., & Lavín, P. (2006). Potential use of near infrared reflectance spectroscopy (NIRS) for the estimation of chemical composition of oxen meat samples. *Meat Science*, *74*(3), 487–496.

Votteler, M., Carvajal Berrio, D. A., Pudlas, M., Walles, H., Stock, U. A., & Schenke-Layland, K. (2012). Raman spectroscopy for the non-contact and non-destructive monitoring of collagen damage within tissues. *Journal of Biophotonics*, *5*(1), 47–56.

Wagnières, G. A., Star, W. M., & Wilson, B. C. (1998). In Vivo Fluorescence Spectroscopy and Imaging for Oncological Applications. *Photochemistry and Photobiology*, *68*, 603–632.

Wold, J. P., Lundby, F., & Egelandsdal, B. (1999). Quantification of connective tissue (hydroxyproline) in ground beef by autofluorescence spectroscopy. *Journal of Food Science*, *64*(3), 377–383.

Wold, Jens Petter, Kvaal, K., & Egelandsdal, B. (1999). Quantification of intramuscular fat content in beef by combining autofluorescence spectra and autofluorescence images. *Applied Spectroscopy*, *53*(4), 448–456.

Wu, B., Dahlberg, K., Gao, X., Smith, J., & Bailin, J. (2019). A rapid method based on fluorescence spectroscopy for meat spoilage detection. In F. Jain, C. Broadbridge, M. Gherasimova, & G. Tang (Eds.), *High performance logic circuits for high-speed electronic systems*. Singapore: World Scientific.

Wubshet, S. G., Wold, J. P., Böcker, U., Sanden, K. W., & Afseth, N. K. (2019). Raman spectroscopy for quantification of residual calcium and total ash in mechanically deboned chicken meat. *Food Control*, *95*(June 2018), 267–273.

Young, O. A., Barker, G. J., & Frost, D. A. (1996). Determination of collagen solubility and concentration in meat by near infrared spectroscopy. *Journal of Muscle Foods*, *7*, 377–387.

Table 1. Composition of samples and correlation coefficients (Pearson's r) among the chemical parameters in ground meat and poultry by-products.

Beef samples					
Parameter (%, w/w)	**Min. value**	**Mean value**	**Max. value**	**Correlation coefficients (r)**	
Protein	6	22	44	**Fat - protein**	-0.55
Fat	1.1	19	72	**Hydroxyproline - protein**	0.65
Hydroxyproline	0.1	0.9	3.3	**Hydroxyproline- fat**	0.26
Poultry by-products samples					
Parameter (%, w/w)	**Min. value**	**Mean value**	**Max. value**	**Correlation coefficients (r)**	
Protein	9	16	25	**Fat - protein**	-0.82
Fat	10	21	42	**Hydroxyproline - protein**	0.72
Hydroxyproline	0.4	0.7	1.5	**Hydroxyproline - fat**	-0.36

The correlations were all significant ($p < 0.05$)

465

Table 2. Summary of PLSR models obtained for predicting hydroxyproline in ground meat and poultry by-products.

	Raman		
	Nº comp.	**R^2 (CV)**	**RMSECV (%)**
Beef samples	4	0.94	0.19
Poultry by-products	3	0.81	0.11
Combined samples	4	0.91	0.17
	NIRS		
	Nº comp.	**R^2 (CV)**	**RMSECV (%)**
Beef samples	6	0.89	0.25
Poultry by-products	4	0.82	0.11
	Fluorescence		
	Nº comp.	**R^2 (CV)**	**RMSECV (%)**
Beef samples	3	0.57	0.50
Poultry by-products	4	0.74	0.13

466

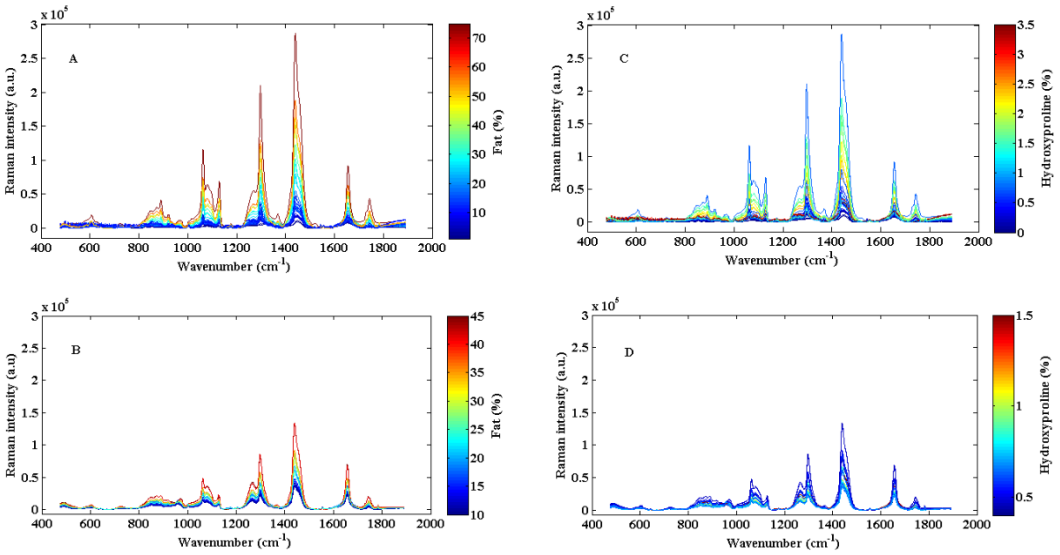
Figure captions

Figure 1. Baseline-corrected Raman spectra for beef samples (upper panel) and poultry by-products (lower panel). The spectra are colored according to percentage of fat (left panel (A, B) and percentage of hydroxyproline (right panel (C, D)).

Figure 2. Normalized NIR spectra of beef samples (upper panel) and poultry by-products (lower panel). The spectra are colored according to percentage of fat (left panel (A, B)) and percentage of hydroxyproline (right panel (C, D)).

Figure 3. Normalized fluorescence spectra of beef (upper panel) and poultry by-products (lower panel). The spectra are colored according to percentage of fat (left panel (A, B)) and percentage of hydroxyproline (right panel (C, D)).

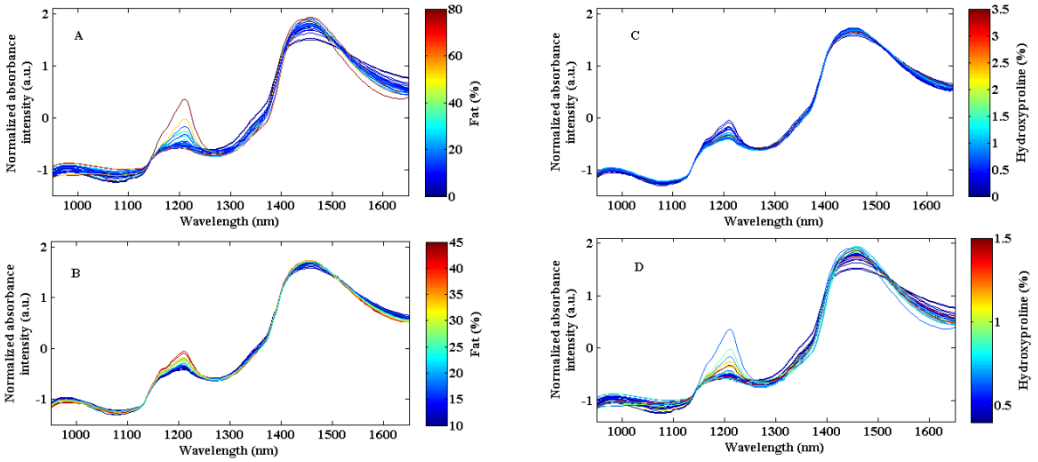
Figure 4. Regression coefficients for the different models obtained: Raman spectroscopy (A), NIRS (C) and fluorescence (D). Raman spectrum obtained from collagen extracted from turkey tendons (B).



480

481 **Figure 1.**

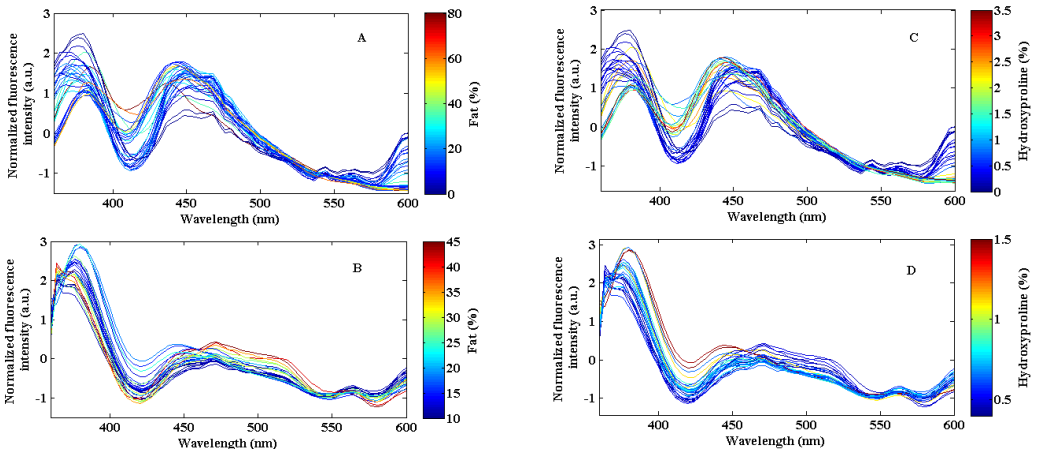
482



483

484 **Figure 2**

485



486

487 **Figure 3**

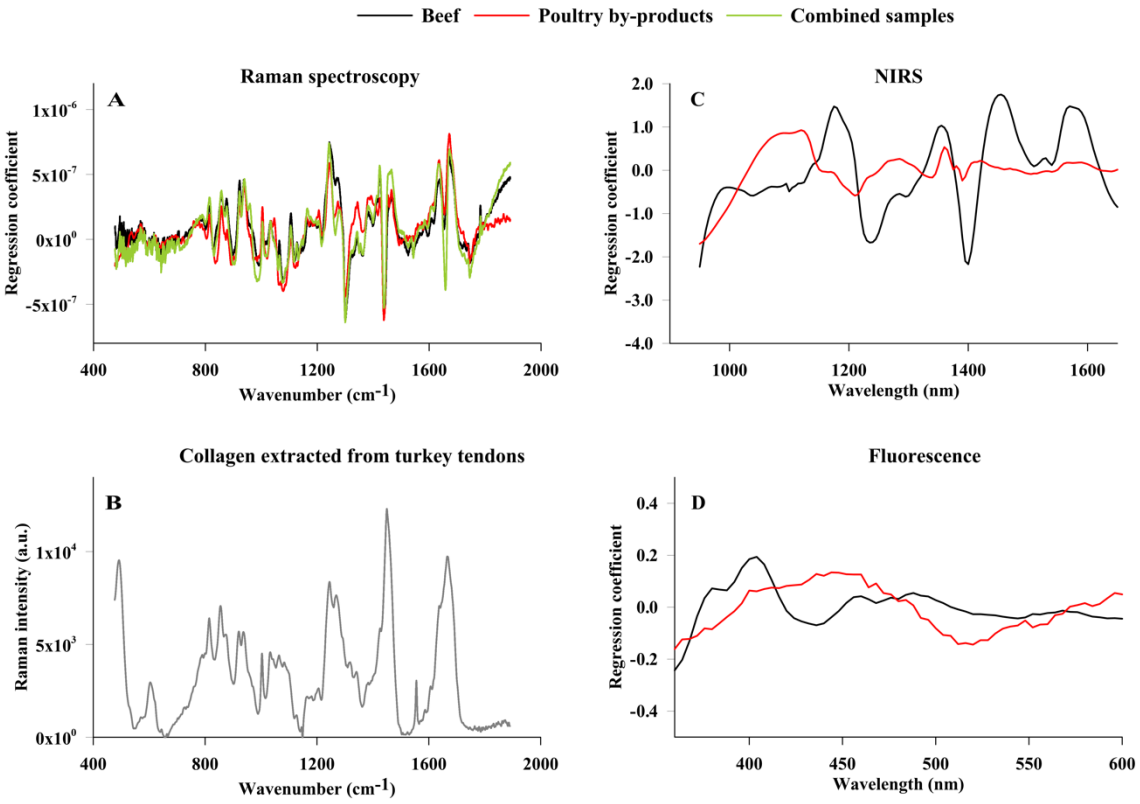

Beef
Poultry by-products
Combined samples
Raman spectroscopy
A
Regression coefficient
1x10-6
5x10-7
0x100
-5x10-7
400
800
1200
1600
2000
Wavenumber (cm-1)
NIRS
C
Regression coefficient
2.0
1.0
0.0
-1.0
-2.0
-3.0
-4.0
1000
1200
1400
1600
Wavelength (nm)
Collagen extracted from turkey tendons
B
Raman intensity (a.u.)
1x104
5x103
0x100
400
800
1200
1600
2000
Wavenumber (cm-1)
Fluorescence
D
Regression coefficient
0.4
0.2
0.0
-0.2
-0.4
400
450
500
550
600
Wavelength (nm)

Figure 4.