Norwegian University of Life Sciences
Faculty of Chemistry, Biotechnology and Food Science

Philosophiae Doctor (PhD)
Thesis 2019:85

# Expanding the analytical toolbox for characterization of proteolytic reactions: Advances in FTIR spectroscopy and classical methods

## Nye analysemetoder for karakterisering av proteolytiske reaksjoner: Utvikling innen FTIR spektroskopi og klassiske metoder

Kenneth Aase Kristoffersen

# Expanding the analytical toolbox for characterization of proteolytic reactions: Advances in FTIR spectroscopy and classical methods

Nye analysemetoder for karakterisering av proteolytiske reaksjoner: Utvikling innen FTIR spektroskopi og klassiske metoder

Philosophiae Doctor (PhD) Thesis

Kenneth Aase Kristoffersen

Norwegian University of Life Sciences
Faculty of Chemistry, Biotechnology and Food Science

Ås 2019

Thesis number 2019:85
ISSN 1894-6402
ISBN 978-82-575-1647-5

# Acknowledgements

This work was carried out as part of Nofima's internally funded strategical research program Peptek, which aims at providing a characterization- and processing-platform for future protein production. The work was performed in the period 2016-2019 at the Department of Raw Materials and Process Optimization, Nofima, Ås, Norway. During this period, I also carried out work for four months at the Wageningen Food and Biobased Research Institute at Wageningen University in the Netherlands. This work was also funded by Nofima. I would like to thank Nofima for both financial support and giving me the opportunity to conduct my PhD studies at their facilities. I would also like to thank the Norwegian University of Life Sciences for enrolling me in their PhD program and the Wageningen Food and Biobased Research Institute for hosting me during my research stay in the Netherlands.

Many thanks to my three supervisors, Prof. Svein Jarle Horn (Faculty of Chemistry, Biotechnology and Food Science, NMBU, Ås, Norway), Dr. Nils Kristian Afseth (Nofima) and Dr. Diana Lindberg (Nofima). All three have supervised me in carrying out research independently, but also as part of an international research effort. Many thanks to Dr. Aart van Amerongen and Dr. Heleen de Vogel-van den Bosch, which were my supervisors during my research stay in the Netherlands. Both Aart and Heleen have contributed extensively to this work and they made sure that my stay in the Netherlands were highly productive.

Many thanks also to the Department of Raw Materials and Process Optimization team (past and present), Peptek members and students. Everyone has contributed to make the working environment both fun and interesting. I also appreciate all the valuable technical support and discussions around the research conducted as part of this thesis and in general. Everyone has supported me in so many ways that it would be impossible to list them all, but a special gratitude and appreciations are given to those contributing as co-authors to the papers presented as part of this thesis.

I would also like to express my deeps gratitude and appreciation to family and friends for putting up with me during the course of this PhD. There has been ups and downs, but they have all supported and helped during this period. Special thanks is given to my friends Dr. Ulrike Böcker, Dr. Ole Kristian Merkesvik Brandtzæg, Ms. Katinka Dankel MSc, Dr. Emily McEwan and Mrs. Rachel McEwan for proofreading this thesis.

Last of all, but certainly not least, I would like to thank Dr. Renaud Laurain for all the support during my PhD studies. He has supported and helped with language and graphics for publications and the thesis. He has also made sure that my dog Otto has had all that he needed during the long hours I spent writing this thesis. Thank you Renaud, this would not be possible without your help and support.

Ås, October 2019

Kenneth Aase Kristoffersen

# Abstract

Enzymatic protein hydrolysis is a well-established industrial process recognized for its potential to improve sustainability in food production by valorization of protein-rich by-products from the food industry. Monitoring such processes is still a significant challenge, as the existing classical analytical methods are not easily applicable to industrial environments. The lack of fast analytical tools to monitor and control enzymatic protein hydrolysis processes results in high variation in product quality, reducing the possible applications of the hydrolysate products in food and other supplements. The aim of this thesis was, therefore, to expand the analytical toolbox for characterization of proteolytic reactions by the development of new applications based on FTIR spectroscopy and classical analytical methods.

This thesis presents methods and approaches that can be adapted to industrial setups as process control and monitoring tools. The four papers that are part of this work also address, in one way or the other, the need for new analytical approaches in this industry. The focus of the three first papers was FTIR-based methods to monitor protein degradation during proteolytic reactions. Dry-film FTIR spectra were used to monitor this development in complex reaction mixtures. Partial least squares regression calibration models were then constructed to predict the two common protein hydrolysate quality parameters, degree of hydrolysis and average molecular weight, by linking the FTIR spectra to classical analysis results. The fourth paper showed how different poultry raw materials and enzyme products affected the protein degradation pattern and the product quality of protein hydrolysates.

The work conducted has resulted in a substantial analytical data library, including analytical data of raw materials, hydrolysate products and degradation patterns by following degree of hydrolysis and average molecular weight during enzymatic protein hydrolysis processes. The data collected will be useful in future EPH studies. Overall, the present work represents a step forward towards the increased use and valorization of

existing protein resources, as it offers new methods and approaches to monitor, control and understand ongoing enzymatic protein hydrolysis processes.

# Sammendrag

Enzymatisk proteinhydrolyse er en veletablert industriell prosess som gir muligheter for økt utnyttelse av proteinrike biprodukter fra matindustrien, og dermed en mer bærekraftig matproduksjon. Overvåking av slike prosesser er fremdeles en betydelig utfordring siden de eksisterende klassiske analysemetodene ikke enkelt kan brukes industrielt. Mangelen på raske analyseverktøy for å overvåke og kontrollere enzymatiske proteinhydrolyseprosesser resulterer ofte i høy variasjon i produktkvalitet. Dette er med på å begrense de mulige bruksområdene for hydrolyseproduktene. Målet med denne avhandlingen var derfor å utvikle nye karakteriseringsteknikker, basert på FTIR-spektroskopi og klassiske analysemetoder, for å kunne følge proteolytiske reaksjoner.

Denne avhandlingen omhandler nye metoder som kan benyttes til prosesstyring og prosessovervåking. De fire artiklene som er en del av dette arbeidet tar på ulike måter for seg behovet for nye analytiske løsninger innen enzymatisk proteinhydrolyse. De tre første artiklene viser hvordan FTIR-spektroskopi kan benyttes for å overvåke proteolytiske nedbrytningsreaksjoner. Tørrfilm FTIR-spektroskopi ble benyttet for å følge nedbrytningen av proteiner i komplekse reaksjonsblandinger, og kvantitative modeller basert på FTIR-spektre og resultater fra klassiske analyser av proteinhydrolyseprøver ble utviklet. Resultatene viser at metodene kan benyttes til å raskt estimere både hydrolysegrad og gjennomsnittligmolekylvekt, som er to vanlige kvalitetsparameterne benyttet i hydrolyseindustrien. Den fjerde artikkelen viser hvordan forskjeller i biproduktsammensetning fra kylling og kalkun, samt ulike enzymtyper, påvirker proteinnedbrytningsprosessen. Resultatet fra analysene viser blant annet at det er forskjeller i nedbrytningsmønsteret og produktkvaliteten til proteinhydrolysatene som ble produsert.

Arbeidet har resultert i et omfattende databibliotek med resultater fra analyser av biprodukter før prosessering, hydrolyseprodukter og prøver tatt ut under hydrolyseprosessen. I kombinasjon med nye og eksisterende måledata, vil dette

databiblioteket være svært nyttig også i fremtidige studier. De nye metodene som er utviklet for å kunne overvåke enzymatiske proteinhydrolyseprosesser representerer et viktig skritt mot økt bruk av eksisterende proteinressurser og dermed økt bærekraft i dagens matproduksjon.

# Abbreviations

| | |
|---|---|
| ACE | Angiotensin-I-converting enzyme |
| BSA | Bovine serum albumin |
| CLN | Chemiluminescence nitrogen |
| CMDR | Chicken mechanical deboning residues |
| DH% | Degree of hydrolysis |
| EMSC | Extended multiplicative signal correction |
| EPH | Enzymatic protein hydrolysis |
| FTIR | Fourier-transform infrared |
| GRAS | Generally recognized as safe |
| HC-PLSR | Hierarchical cluster-based partial least squares regression |
| Hot PLS | Hierarchically ordered taxonomic classification by partial least squares |
| HPLC | High-performance liquid chromatography |
| IR | Infrared |
| LWR | Locally weighted regression |
| MSC | Multiplicative signal correction |
| $M_w$ | Weight-average molecule weight |
| MWD | Molecular weight distribution |
| OPA | *o*-phthaldialdehyde |
| PC1 | First component |
| PCA | Principal component analysis |
| PLSR | Partial least squares regression |
| QPS | Qualified presumption of safety |
| $R^2$ | Coefficient of determination of the cross-validation |
| RMSE | Root mean squared error |
| RMSECV | Root mean squared error of the cross-validation |
| SDS-PAGE | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| SEC | Size exclusion chromatography |

| | |
|---|---|
| TFA | Trifluoroacetic acid |
| TNBS | Trinitrobenzenesulfonic acid |
| UV | Ultraviolet |
| WMP | Whole milk powder |
| WPC80 | Whey protein concentrate powder |
| WPO | Milk whey powder |

# List of papers

This thesis is based on the work contained in the following papers, referred to by roman numbers in the text.

I **FTIR-based hierarchical modeling for prediction of average molecular weights of protein hydrolysates**

Kenneth Aase Kristoffersen*, Kristian Hovde Liland, Ulrike Böcker, Sileshi Gizachew Wubshet, Diana Lindberg, Svein Jarle Horn and Nils Kristian Afseth, *Talanta* 205 (2019) 120084.

II **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 and Sileshi Gizachew Wubshet, revised paper submitted to *Food Chemistry* September 2019.

III **Fourier-transform infrared spectroscopy for protein hydrolysate characterisation using dry-films treated with trifluoroacetic acid**

Kenneth Aase Kristoffersen*, Aart van Amerongen, Ulrike Böcker, Diana Lindberg, Sileshi Gizachew Wubshet, Heleen de Vogel-van den Bosch, Svein Jarle Horn and Nils Kristian Afseth, Submitted to *Scientific Reports* September 2019.

IV **Effects of poultry raw material variation and choice of protease on protein hydrolysate quality**

Diana Lindberg*, Kenneth Aase Kristoffersen, Heleen de Vogel-van den Bosch, Sileshi Gizachew Wubshet, Ulrike Böcker, Anne Rieder, Enno Fricke and Nils Kristian Afseth, prepared for submission to *Process Biochemistry*.

# Additional publication and paper

**Bioanalytical Aspects in Enzymatic Protein Hydrolysis of By-products**

Sileshi Gizachew Wubshet*, Diana Lindberg, Eva Veiseth-Kent, Kenneth Aase Kristoffersen, Ulrike Böcker, Kathryn Elizabeth Washburn and Nils Kristian Afseth, *Academic Press*, Book chapter in: Charis Galanakis (Ed.), Proteins: Sustainable Source, Processing and Applications, 2019, Pages 225-258.

**By-products from the food industry as a substitute for serum in cell cultures**

Randi Christel Andreassen, Mona Elisabeth Pedersen, Kenneth Aase Kristoffersen, Ingrid Måge, Sissel Beate Rønning*, Prepared for submission to *Food & Function*.

# Table of Contents

# 1 General introduction

The intention of Chapter 1 is to provide the general background, motivation and aims for the work conducted as part of this thesis. The main motivation for this research is to increase the utilization of existing protein resources to ensure more sustainable food production in the future. This chapter is therefore devoted to protein as a food source, starting with a global perspective on protein as a limited resource.

## 1.1 Protein as food: A global perspective

Proteins are a large and diverse class of complex biomolecules. These molecules are the building blocks of body tissues, and they also constitute an essential part of our diet. As an energy source, proteins provide 4 kcal (17 kJ) per gram, which is comparable to carbohydrates. However, the most important aspect of proteins from a nutritional standpoint, is the amino acid composition. The amino acids which proteins consist of, can when consumed, be used to construct new molecules important for a variety of biological processes and growth. Some of the amino acids are essential, as they cannot be synthesized by the body itself at the rate which they are used, and therefore must be supplied through food. Most individuals meet these needs through their daily diet by eating foods containing protein from both animal and plant origins, but some groups such as elderly people may need a protein enriched diet to fulfill their bodies needs for protein.

The food that we eat is mainly produced by growing crops, raising livestock and by harvesting natural resources. This is something humans have done for generations without much concerns for how it affects the environment. Now, with population growth, increasing urbanization and prosperity, food production has become one of largest global challenges humanity has ever encountered. There is now robust evidence and high agreement between nations that food production is one of the leading factors contributing to climate change and reduction in biodiversity. This is clearly stated in the many reports addressing these issues, such as the latest reports from the Food and

Agriculture Organization (FAO), the Intergovernmental Panel on Climate Change (IPCC), the United Nations Climate Change Secretariat (UNCCC) and the United Nations Environment Program (UNEP) [1-6]. These reports agree that drastic action is needed, in all steps of the food supply chains, in order to meet these challenges. One of the major concerns is, however, that the population is still increasing. It is estimated that by the year 2050, the global population will be about 9.7 billion people [7]. An increase in population will almost certainly mean a need for increased food production, and based on the current food consumption trends it has been estimated that an increase of 70% is necessary [8].

The world's food production systems are complex, and different crops and livestock contribute differently to climate change and reduction in biodiversity. The present thesis has as its main focus protein derived from animal sources. This is an important protein source in many cultures all over the world despite the fact that its production is highly energy demanding and a large source of pollution. It is, for example, estimated that raising livestock alone accounts for 14.5% of the total human-induced greenhouse gas emissions [9]. Growth of 70% in all food production sectors, including raising livestock, will have a global impact and should be controlled in order to reach the goal in the Paris agreement, keeping a global temperature rise this century well below 2 $^{O}$C [10]. There are numerous ways to reduce or stop the growth in emissions from livestock production. One way, which is central in this study, is to use more of the protein which is already produced, and thereby reduce the need for production increase. This can be achieved through valorization of by-products as large parts of the animals are not used as food, even if it holds a quality suitable for human consumption after the European Union (EU) regulations [8].

## 1.2 Protein structure

Protein molecules play many critical roles in all living organisms. They serve as structural elements, transportation channels, signal receptors, transmitters, enzymes and more. Proteins are therefore an incredible diverse class of biomolecules when it comes to structure. This diversity is made possible by the different amino acids which

are the building blocks of proteins. There are 20 common amino acids and they have a general structure shown in Fig. 1 [11-13]. In the general structure the central carbon is named the alpha carbon and is located between an amino and a carboxyl group. The remaining two positions on the alpha carbon are generally bonded to a hydrogen and variable R-group (also called a side chain or a side group). The variable R-group defines which class the amino acid belongs to. For example, a carboxyl group on the side-chain defines glutamic acid and aspartic acid. These side-groups are under biological pH deprotonated and negatively charged, and the amino acids are therefore classified accordingly, as shown in Fig. 1.

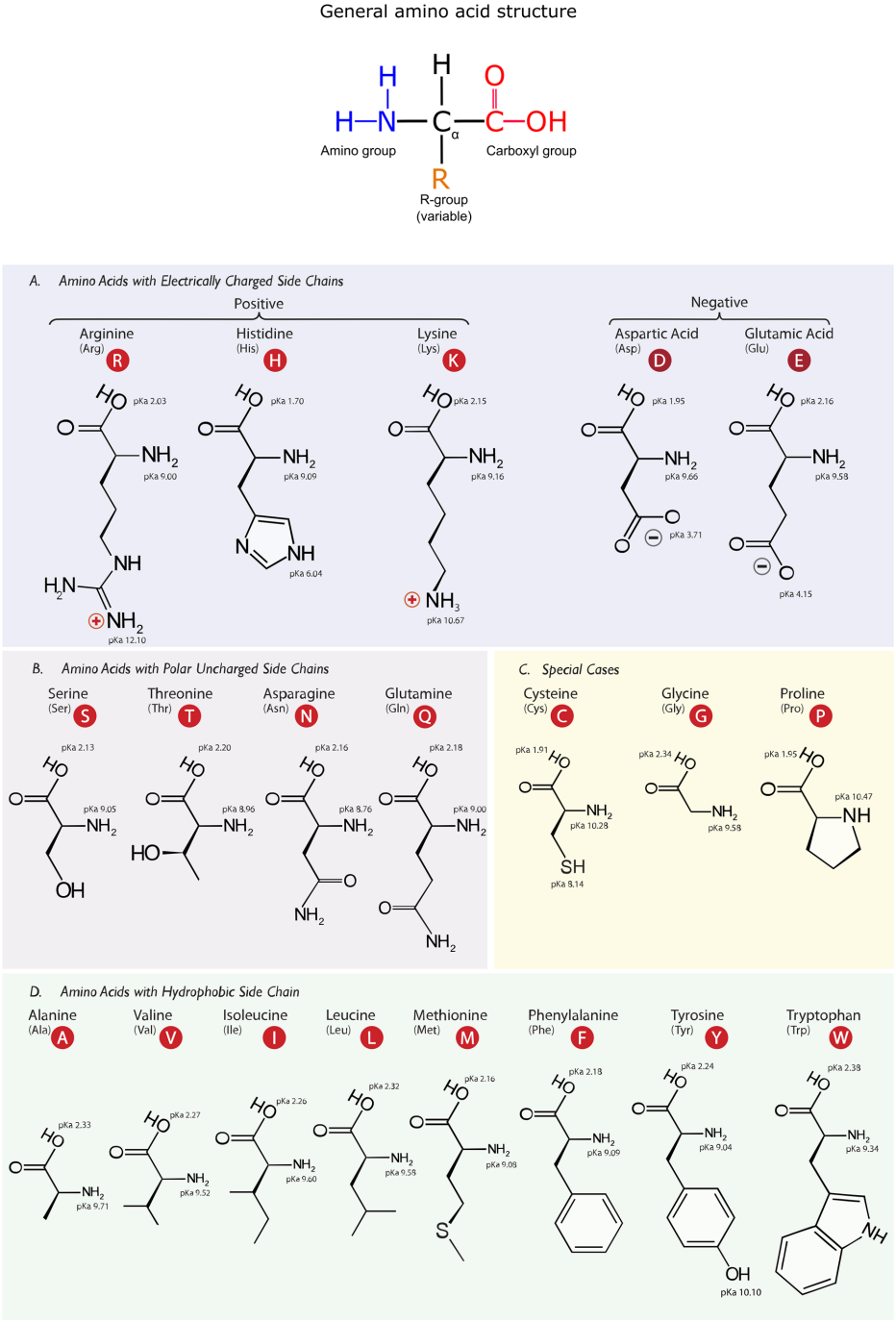



**Fig. 1.** The general structure of amino acids and an overview of 20 common amino acids. Adapted from Wikipedia [14].

The three-dimensional structures of proteins are described by the primary, secondary, tertiary and quaternary structures illustrated in Fig. 2 [11-13]. The primary structure is the sequence of covalently bonded amino acids in a polypeptide chain arranged as $C_\alpha$-C-N-$C_\alpha$ [12]. These three repeating covalent bonds build up the protein backbone and have special features when it comes to rotation. The N–$C_\alpha$ and the $C_\alpha$–C bonds can rotate freely depending on the side-groups and surrounding environment. The C–N bond has a partial double bond nature and is therefore restricted [12]. This degree of freedom gives rise to the secondary structure of proteins by enabling folding of the polypeptide chain into specific structures such as alpha helixes and beta sheets. The secondary structures are further folded into the tertiary structure and if a protein is composed of several polypeptide chains, these tertiary structures are arranged into the quaternary structure of the complete protein as illustrated in Fig. 2.

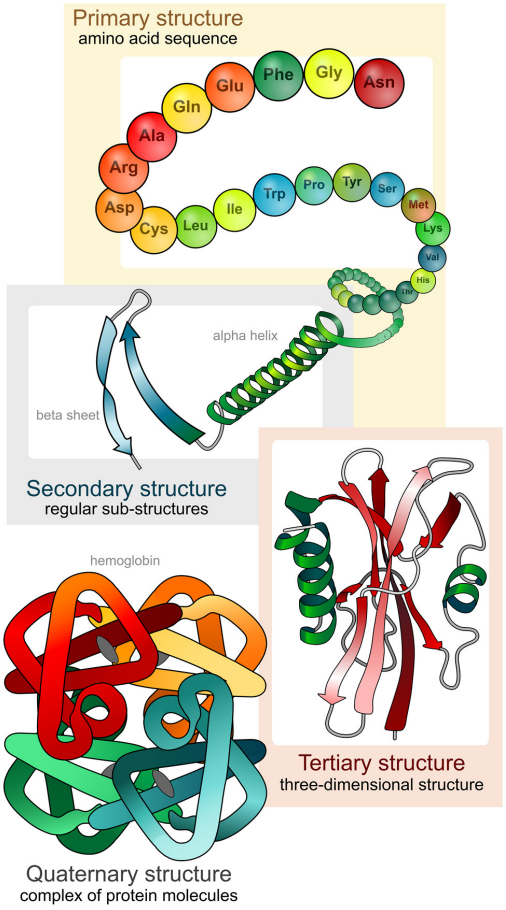



**Fig. 2.** Protein primary, secondary, tertiary and quaternary structures. Image adapted from Wikipedia [15].

The three-dimensional structure of proteins is diverse and dependent on the amino acid sequence of the polypeptide chains. The structural differences give the proteins features which, for example, can be seen in different tissues. Muscle tissue, for instance, is largely composed of the proteins actin and myosin, while connective tissues contains large quantities of collagens [13, 16, 17].

## 1.3 Valorization of protein-rich by-products

Several industrial technologies are currently applied to increase the value of the by-products left over from for example the slaughtering of livestock. This includes rendering, protein hydrolysis and ensilage processes as well as others. The main focus in the present thesis is on proteolytic reactions for liberation and fractionation of proteins from by-products. During this type of process, the proteins are cleaved to produce smaller peptides and free amino acids making them more water soluble and easily extractable. In recent years, this type of processing has gained significant attention, especially enzymatic protein hydrolysis (EPH) processes, as it is a more versatile technology. The advantage of the of EPH reactions compared to other proteolytic reactions in processing, is that the technology increases the possibility of adapting the process parameters to meet specific product qualities. EPH processing is also regarded as a mild biotechnological process as the need for harsh reaction conditions is reduced, resulting in high product yields without affecting the nutritional quality [18, 19]. Several studies and currently applied industrial EPH processes show that it is possible to use a larger portion of the existing protein as human food or for other applications such as pharmaceuticals, cosmetics, sport formulations and more, as illustrated in Fig. 3.

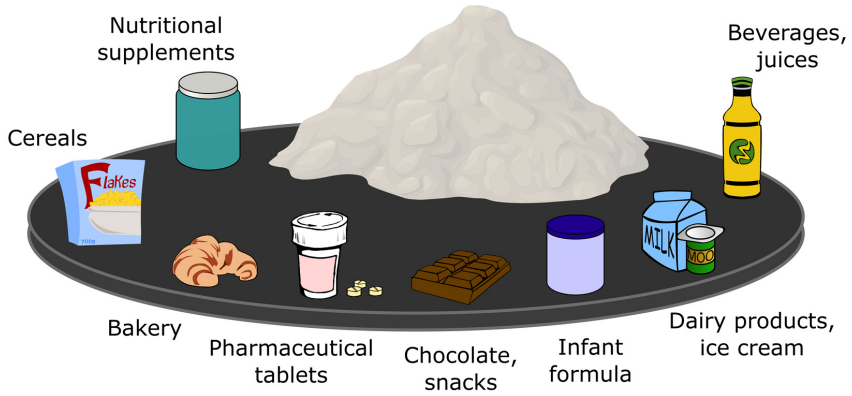



**Fig. 3.** Examples of possible products where protein hydrolysates can be used.

Fig. 3 illustrates an ideal situation with many potential markets. Naturally, this is what the industry is aiming for. Currently, however, there are challenges in achieving this goal related to both market and technological aspects. To introduce protein and peptide fractions to well-paying markets, such as nutrition ingredients, pharmaceuticals, weight control formulas, or special products for the pet food marked, the industry needs to produce EPH products with strictly defined specifications [20-22]. This places extra requirements on the ability for continuous optimization and fine-tuning of the EPH process as it is running, especially as there are large variations in the by-product raw materials going into the process [19, 21]. For instance, the by-product material can vary in composition of biomolecules and minerals, from day to day, and even from minute to minute. By-product materials usually have a high protein and lipid content, but they might also have significant fractions of sugars, depending on the source [22]. Monitoring, controlling and optimizing EPH reactions under such complex reaction conditions is a major challenge and fast tools for process monitoring is close to non-existent in the current EPH industry. A variety of classical analytic methods, some of which are described in Chapter 3, have been applied in order to follow EPH process and measure quality of the hydrolysate products. The challenge with these classical methods is that they are labor intensive, including multiple manual procedures, time-consuming and difficult to implement in an industrial environment. Controlling the outcome of an ongoing reaction using such methods is, therefore, not practical as the results would only be available after the hydrolysis reaction has progressed significantly, or come to completion. Spectroscopic methods such as Fourier-transform infrared (FTIR) spectroscopy linked to these classical measurements, represent a fast alternative which can potentially be utilized in industrial settings [22-24]. Development of FTIR-based methods for following EPH reactions have advanced significantly in recent years, but there are still challenges that need to be addressed before commercial systems are available.

## 1.4 Aims of the thesis

The main objective of the thesis is to expand the analytical toolbox for characterization of proteolytic reactions by the development of new applications based on FTIR spectroscopy and classical analytical methods.

Sub-objectives:

- To evaluate FTIR based approaches for prediction of weight-average molecular weight of protein hydrolysates produced using different raw materials and enzymes (**Paper I**).
- To study the relationship between weight-average molecular weight and degree of hydrolysis and dry-film FTIR fingerprints of milk protein hydrolysates (**Paper II**).
- To assess trifluoroacetic acid treated dry-films of EPH samples as an improved strategy for FTIR based characterization of protein hydrolysates (**Paper III**).
- To study different quality parameters and their changes related to raw material composition and enzymes used in production of EPH products (**Paper IV**).

# 2 Production of protein hydrolysates

Chapter 2 gives an overview of proteolytic reactions and how they are utilized for extraction and valorization of food proteins. In a proteolytic reaction, peptide bonds are cleaved. This is a critical reaction in biological systems allowing protein degradation to occur so that organisms can reorganize the amino acids into new molecules essential for maintaining their life processes and growth. Cleaving peptide bonds, however, demands a relatively high activation energy. Organisms usually solve this by producing proteolytic enzymes that catalyzes the reaction. These enzymes are specific when it comes to which peptide bonds they cleave, resulting in for example differences in molecular weight distributions in EPH products. Protein hydrolysates hydrolyzed by different enzymes can as a result have very different functional and biological properties. This specificity has been exploited by humans for food production purposes long before the sciences and understanding of these mechanisms were established; for example by maintaining ideal growth conditions for specific organisms producing proteolytic enzymes in cheese production [25].

There are multiple factors that determine the outcome of an EPH reaction, and the main factors can be classified into the three following sub-groups: 1) Enzyme-specific factors such as specificity, stability, and sensitivity to inhibitors, 2) Substrate-specific factors such as origin, age, feed regime, and complexity, and 3) Process-specific parameters, such as substrate concentration, enzyme to substrate ratio, time, temperature, and pH [22]. Chapter 2 focus on the factors covered in subgroups 1 and 2, influencing chemical analysis measurements. Other specific factors influencing the process, such as pH, hydrolysis time and temperature, are only briefly covered in this thesis.

## 2.1 Protein hydrolysis

The goal in industrial production of protein hydrolysates is normally to liberate and recover as much protein as possible from a biomass at a low cost. This can be achieved by using different hydrolysis approaches to cut the proteins down to smaller and more

water soluble components [26]. The main principle for these approaches is the same, the peptide bonds are cleaved in a hydrolysis reaction where a water molecule is used to form a new C- and N-terminal as shown in Fig. 4. To cleave the peptide bonds, a reduction in activation energy is needed to increase the reaction rate [27]. Hydrolysis of peptide bonds are therefore normally catalyzed either by the use of enzymes called proteases or by the use of acidic/alkaline conditions at elevated temperatures and high pressures. These different approaches result in very different products. Acids and alkaline hydrolysis, hereby referred to as chemical hydrolysis, is for example not specific to which peptide bonds that are cleaved. In addition, the acid or base used needs to be neutralized when the reaction has ended. Thus, additional extraction steps are needed, as the salts formed may need to be removed from the product before further use. Another issue is that thermal processing of proteins, especially under alkaline conditions, will result in the formation of unwanted toxic compounds such as lysinoalanine, and randomization of the stereochemistry of the amino acids [28, 29]. Acidic hydrolysis, although also being a relatively harsh treatment, is therefore preferred over alkaline and is commonly used for food production purposes [30, 31].

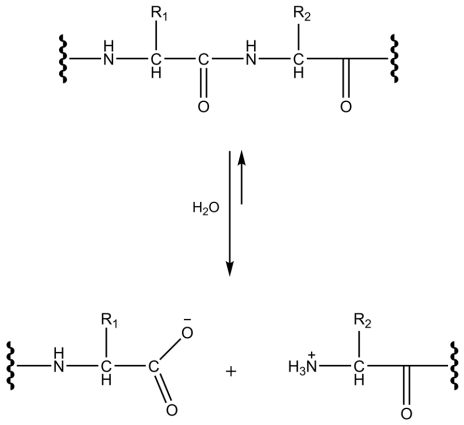



**Fig. 4.** The hydrolysis reaction of a peptide bond.

EPH liberation of peptides and free amino acids is considered a milder approach compared to chemical hydrolysis, as the needs for elevated temperatures and harsh reaction conditions are greatly reduced. This limits the risk for quality loss through side

reactions, preserving more of the nutritional value of the proteins [26, 28, 32]. The specific nature of proteases also opens for production of hydrolysates exhibiting specific functional and biological properties, e.g. water holding, foaming ability, taste, and blood sugar regulating properties [28, 30, 33-36]. A typical setup for industrial production of hydrolysates from processing by-products from meat or fish industry is shown in Fig.5. There is a wide range of setups of unit operations that is used in production of hydrolysates. The unit of operation might involve additional steps in pretreatment and downstream operations.

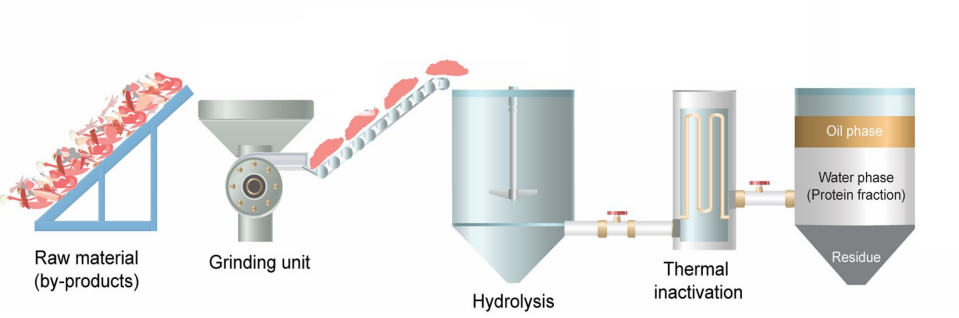



**Fig. 5.** A typical setup for industrial production of hydrolysates processing by-products from meat or fish industry. Image created by Wubshet [37].

## 2.2 Proteolytic enzymes

Proteolytic enzymes, also known as proteases, peptidases or proteinases catalyzes hydrolysis of specific peptide bonds as shown in Fig. 6. They are classified either based on the critical amino acids responsible for the catalytic function, the pH-optimum for activity, the site of where the peptide bond is cleaved, or the requirement of free thiol groups [38, 39]. The proteases can further be divided into endopeptidases and exopeptidases. An endopeptidase cleaves peptide bonds within the peptide chain, while exopeptidases cleaves peptide bonds at the end of the peptide chain, releasing free amino acids, or as small di- or tripeptides [22]. All enzymes are systematically classified by the European Commission (EC). Here, the proteases are numbered according to rules and nomenclature set by the Nomenclature Committee of International Union of

Biochemistry and Molecular Biology (NC-IUBMB). In this system, the enzymes are classified based on the chemistry they perform, and hence, the enzymes are given numbers describing their function and activity. For proteolytic enzymes the first number is 3, which stands for enzymes that form two products from a substrate using hydrolysis (hydrolases). The second number is 4, which stands for hydrolases that act on peptide bonds (proteases). The third number being 11-19 classifies the enzyme as an exopeptidases, and the third number being between 21-99 classifies the enzyme as an endopeptidase.

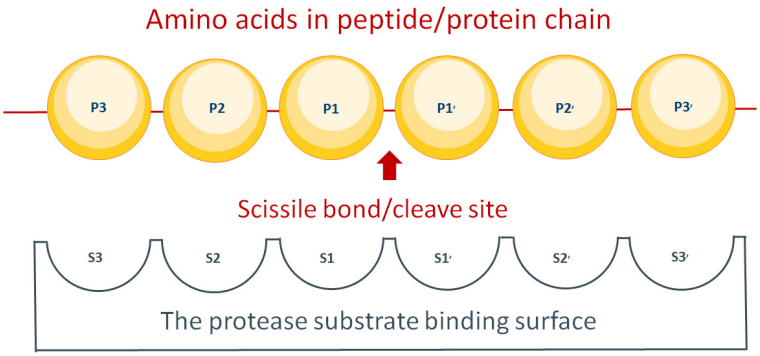



**Fig. 6.** The principle of an enzymatic hydrolysis reaction with nomenclature. Image created by Lindberg [40].

The molecular structures responsible for the specific nature of proteases are today well-known, and a nomenclature describing how the enzyme interact and cleave peptide bonds in the substrate polypeptide chain have been implemented. The nomenclature is shown in Fig. 6, where P1, P2…Pn is the amino acids of the N-terminal side of the substrate peptide chain and P1′, P2′…Pn′ is the amino acids of the C-terminal side. The proteases have substrate-binding pockets corresponding to the amino acids named S1, S2 …Sn and S1′, S2′…Sn′ and the hydrolysis takes place at the scissile bond located between P1 and P1′. [41, 42].

There are a large number of proteases found within organisms and viruses. This shows that proteases are highly specialized for the activities of the organisms from which they originate. Some of these specialized enzymes are utilized in industrial applications, e.g. in food and beverage production. The use of proteases has increased over the years and they are now one of the market-leading class of enzymes used industrially worldwide [19, 43]. Commercially available proteases formulated for food applications usually contain a combination of several proteolytic enzymes, hereby referred to as enzyme or protease products. The specificity of the proteases range from very specific to less specific with regard to which type of proteins and peptide bonds they target [44]. There are also differences in stability, activity and sensitivity to inhibitors [22]. It follows that protein hydrolysates produced using the same substrate can have specific functional and biological properties dependent on the type of protease product used. There are now many food-grade protease products available which may originate from microbes, animals and plants [39, 44]. The protease enzymes produced by microbes are often preferred, as the production offers advantages when it comes to cost and small variations in quality from batch to batch. For an enzyme product to qualify for use in food, the microbial production host and the way an enzyme product is produced need to be generally recognized as safe (GRAS) and/or have a qualified presumption of safety (QPS) status [45, 46].

## 2.3 Raw materials

The quality and complexity of protein-rich raw materials used for protein hydrolysate production usually varies. The raw materials can originate from many sources, both from plants and animals. The latter group of the raw materials is often a by-product from production of other more highly valued parts such as filets. The composition of by-products will inherently rely on the processing settings of the original raw material, which may vary largely from day to day. Variations in fat content and/or cartilage hardness linked to age, feeding regime and cage options for poultry and aquaculture fish species are typically seen [21, 22, 47]. Other differences are more source specific. For example, fish raw materials where it is common to observe large seasonal variations and variation in the endogenous enzyme content [28, 47-49].

The amount of protein-rich by-products that potentially can be used for protein hydrolysate production is huge. In view of the total meat production in the world, Table 1 published by Aspevik et al. illustrates the great potential of protein hydrolysate production. The table includes the amount of by-products and the global production of the most common livestock in 2014 [18]. These numbers are based on the Food and Agriculture Organization Corporate Statistical Database (FAOSTAT) and the percentage of residual raw materials are based on Norwegian data [8]. In 2014 the estimated global annual production of meat and fish was 263 and 128 million tons, respectively [18]. Considering these numbers, it is clear that the potential for improved utilization of protein recourses is substantial.

**Table 1.** Production summary (in 1000 heads) of major types of livestock by continent.

| Continent | Cattle | Pigs | Sheep | Chickens | Turkeys |
|---|---|---|---|---|---|
| **Africa** | 312,327 | 34,332 | 340,749 | 1,809,059 | 23,658 |
| **Americas** | 508,942 | 169,902 | 86,074 | 5,436,151 | 312,477 |
| **Asia** | 491,020 | 590,548 | 536,251 | 11,923,472 | 14,575 |
| **Europe** | 122,011 | 185,546 | 130,118 | 2,114,988 | 110,786 |
| **Oceania** | 40,226 | 5,346 | 102,432 | 126,014 | 1,377 |
| **Summary world** | 1,474,526 | 985,673 | 1,195,624 | 21,409 683 | 462,873 |
| **By-product (%)**[a] | 60 | 37 | 63 | 51 | 45 |

a) Data based on the assumption that everything not sold as meat from the animals can be considered as by-product.

Not all by-products can be used for human consumption. The EU has instated legislation controlling the use of by-product raw materials from fisheries, aquaculture and livestock industries. The quality of the raw material is important for the processing possibilities and the future use of the product. The classification defines the possible applications, according to the hygiene regulations for food of animal origin [50]. Raw materials that do not meet the general regulations for food hygiene are classified as not suitable for human consumption. These by-product materials are classified by the

regulations of animal by-products, which define possible use of the end products not suitable for human consumption [51, 52].

The raw material composition and quality are two of the major factors affecting the outcome of the hydrolysis process [19, 21, 22]. This is due to differences and variations in the accessibility of the substrate proteins and peptides and what type of protein the substrate contains with regard to their amino acid composition and sequence. Variations in composition will therefore, together with the type of enzyme product used, influence and determine the product quality when it comes to nutritional and physicochemical properties of the product [18, 22]. The variability of the product quality could be reduced by optimizing the EPH parameters based on the raw material composition. Continuous optimization and fine-tuning of the EPH process as it is running would, of course, be dependent on fast real-time analytical measurements of the raw material going in to the process, the progression of the ongoing EPH reaction and the protein hydrolysate product. Currently, as these types of analytical tools do not exist, EPH processing are usually preformed using fixed processing parameters, such as temperature, hydrolysis time and protease concentration.

# 3 Analytical methods

Chapter 3 gives an overview of the methods used in the thesis. This chapter is divided in three main parts: Classical methods, FTIR spectroscopy and multivariate data analysis. The first part introduces the methods applied in the papers presented as part of this thesis. The methods are presented in a historic perceptive, and examples of their use in the production of protein hydrolysate are given. The most important classical methods for the current study have been measurements of degree of hydrolysis and molecular weight distribution parameters measured using size exclusion chromatography. The second part, FTIR spectroscopy, introduces the basics of infrared spectroscopy before moving on to dry-film FTIR measurements of protein hydrolysates. The third part presents multivariate data analysis with emphasis on the methods used in the papers. All data collected as part of this thesis is also part of Peptek which is Nofima's strategic research programs aiming to increase knowledge and develop tools to control EPH processes in industrial settings.

## 3.1 Classical methods

Chemical analysis of both raw materials and the hydrolysis products is essential to increase the understanding of the complex EPH reactions and to enable better process control. The normal procedure in raw material characterization of many animal and marine by-products involves analysis to determine the content of the major components. The main components of animal and marine by-products are protein, water, ash and fat. The raw materials can also contain carbohydrates, but the fraction in these materials is normally very low. This chapter provides a general introduction to the methods used in the four papers included in this thesis. The method protocols are presented in the papers. It should be noted that the method protocols may vary from study to study [22, 53-56]. Comparing results from different studies is therefore a challenge and should be done with care. A more detailed review of most of the analytical methods applied in this thesis is found in a co-authored book chapter entitled Proteins: Sustainable Source, Processing and Applications [22].

### 3.1.1 Protein content

Protein content analysis is essential in the study of protein hydrolysates. The two most commonly used analytical methods to determine total protein content are the Dumas and the Kjeldahl methods [57]. Both methods rely on measurements of nitrogen content to calculate the crude protein content with the use of a conversion factor. The factor of 6.25 is the most frequently used, based on the assumption that proteins contain 16% nitrogen and that all nitrogen in food originate from proteins. The nitrogen percentage of proteins and peptides is of course affected by the amino acid composition. Alternate raw material specific conversion factors have therefore been proposed [58, 59]. Conversion factors for both the Dumas and the Kjeldahl methods specified for different raw materials and applications are provided by the International Organization for Standardization [60].

Combustion procedures for measuring protein were first developed during the 19$^{th}$ century based on a description by Dumas [61]. Dumas was credited for developing the most reliable method of his time, although accuracy was relatively low. Therefore, it was not until the 1980's when combustion methods were refined and new reliable combustion nitrogen analyzers were developed, that the use of combustion analysis for protein determination accelerated [54]. The principle of the combustion procedure is simple: The samples are combusted to form water, carbon oxides, sulfur dioxide and nitrogen oxides. The gases are then separated, and the different nitrogen oxides are reduced to nitrogen gas, which is then detected in a thermal conductivity cell measuring the total nitrogen content.

The Kjeldahl method developed in 1883 is, traditionally, the most used method for protein content analysis [62]. The technique can be applied for determination of the nitrogen content in a wide range of samples, including food matrices. The Kjeldahl method is a wet chemistry procedure, divided in four major steps. The first step is to convert the nitrogen to ammonium sulfate. In the next step ammonium sulfate is neutralized forming ammonia, followed by the third step, where the ammonia is distilled into a known volume of a weak acid such as boric acids. In the final step, back titration

of excess boric acid is performed to measure the total nitrogen content. Several adaptions of these basic steps exists [54]. There are several examples where Kjeldahl have been used for determination of total protein content: For example, in salmon by-products, and several different by-products from the poultry industry [21, 63]. The Kjeldahl and Dumas methods are now considered on equal terms, but specific nitrogen to protein conversion factors should be used as the Dumas method generally measures higher nitrogen values [64]. The difference in measured nitrogen between these two methods have been studied and compared for numerus food and feed matrices, from both animal and plant origin. The results show that Dumas is comparable to the Kjeldahl method, while at the same time represent a faster and more environmentally friendly option [54, 59, 65, 66].

### 3.1.2 Amino acid composition

Amino acid composition analysis is important in many research fields, and are critical for protein quantification. The classical method was developed in the 1950's by Moore, Stein and coworkers [67-69]. Using this method, the amino acid composition is determined in a two-step analytical process. In the first step, the substrate is completely hydrolyzed to liberate all the residues. This is followed by a chromatographic analysis and quantification of the liberated amino acids. The hydrolysis is an important step for a successful analysis and studies have shown that errors in the performance of the hydrolysis is a major factor responsible for inaccurate measurements of composition [70]. There are many varieties in amino acid composition analysis methods available today. They can, for example, vary in how the amino acids are detected or separated [67]. Not all amino acids can be quantified using standard methods described above due to instability under the analysis conditions, specialized techniques for these have been developed. Amino acid composition is an essential quality parameter in the study of EPH reactions and their products. The composition can give insights into which proteins/parts of a substrate are digested and released by different proteases. Amino acid analysis has consequently been applied in many studies of EPH reactions and their products [71-74].

### 3.1.3 Ash analysis

Ash analysis is an old and essential method used in the study of organic matter. The method is also an important quality parameter used in food science. The analysis refers to the inorganic residue remaining after either ignition or complete oxidation of organic matter [75]. Ash analysis are commonly applied in the study of raw materials and protein hydrolysate products. The parameter has been applied in study of fish feed containing protein hydrolysates and in studies assessing nutritional properties of EPH products [76, 77]. In the study of EPH reactions and their products, ash analysis can be used to get insight into the raw material composition and which part of the raw materials protein are liberated [75, 78]. Bones found in the by-products used in EPH processes for instance have a high inorganic part. A high ash content in the raw material is, consequently, an indication of high bone content. The ash content in the hydrolysate product, on the other hand, is not only a quality parameter when it comes to inorganic matter; it can also be an indication of the protease activity when it comes to digestion of bone proteins.

### 3.1.4 Fat content

Fat content is one of the key quality control parameters in the production of feed and food. Determination of fat content has also been used in EPH raw material characterization of by-products from various animal origins [21, 63]. To measure the fat content in complex materials such as poultry by-products, the lipids normally need to be extracted. This can be achieved using heat and organic solvents. Examples of such methods are the Soxhlet, Folch and the Bligh & Dyer. These methods both rely on an organic solvent extraction step(s), using solvents such as hexane, petroleum ether and chloroform [56, 79, 80]. For some materials, such as dairy products, a hydrolytic procedure is recommended before the solvent extraction is performed [81]. The benefit of performing the hydrolysis step before fat extraction is that a full digestion of the material enables better extraction of the lipids. The Soxhlet and the Bligh & Dyer methods have been frequently used and studies comparing the methods on complex raw materials such as lean fish muscles, meat and meat products have been conducted. The results show that the methods preformed differently on different materials [82, 83].

### 3.1.5 Degree of hydrolysis

The degree of hydrolysis (DH%) is a well-established parameter for describing the extent of hydrolysis in a peptide product and for monitoring EPH reactions. The parameter is defined as the percentage of cleaved peptide bonds relative to the total amount of peptide bonds available in a protein or peptide sample. The DH% is calculated as shown in Equation 1,

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

where $h$ is the measured amount of cleaved bonds and $h_{tot}$ is the total number of available peptide bonds [84]. Several methods for determination of DH% have been proposed. The most commonly used methods include the pH-stat [85, 86], trinitrobenzenesulfonic acid (TNBS) [87], *o*-phthaldialdehyde (OPA) [84], formol titration [88] and trichloroacetic acid soluble nitrogen [89]. These methods are based on different principles of measurement of the cleaved peptide bonds. It should also be noted that multiple versions of the mentioned methods have been published.

The TNBS and the OPA method are among the most commonly used techniques for monitoring the extent of protein hydrolysis in EPH products. Both methods are based on measurement of the number of free N-terminals (Fig. 7), and are carried out via derivatization of their amino groups to enable ultraviolet-visible or fluorescence detection [90, 91]. For these techniques, however, the derivatization reagents are known to have different reactivity towards some amino acids. This will affect the accuracy of the measurement. For example, the OPA method will not be accurate when applied on proline- and cysteine-rich hydrolysates [92]. There are many examples where these two methods have been used in optimization of process settings in production of hydrolysates for feed, food and functional food applications. For example, the TNBS method has been used in a study aiming at optimizing process parameters for hydrolysis of by-products of skipjack fish for feed applications and to study functional properties of veal bone hydrolysates [93, 94]. The OPA method has been applied to

study the evolution of various processing parameters such as composition, nutritional value, taste and presence of bioactive peptides [36, 95, 96].

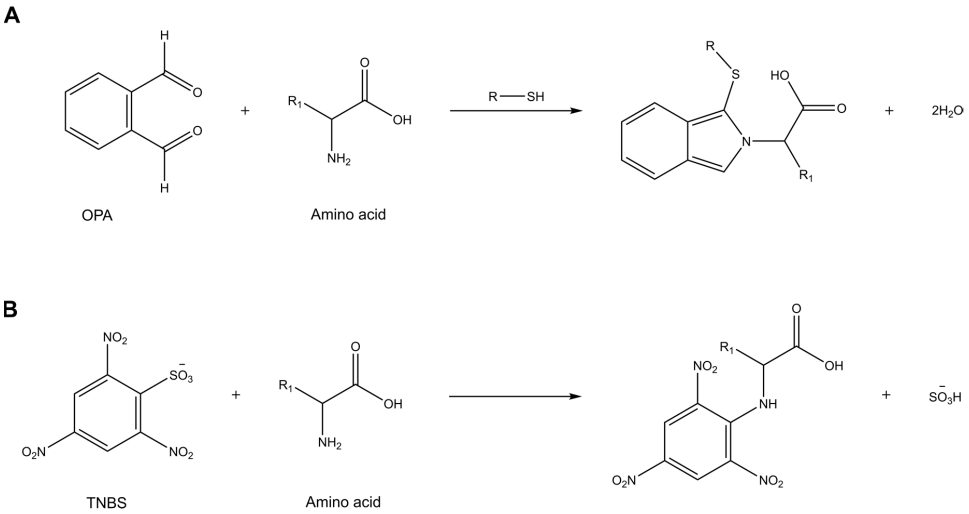



**Fig. 7.** The chemical reactions of (A) OPA with amino acids and (B) TNBS with amino acids.

### 3.1.6 Size exclusion chromatography

Size exclusion chromatography (SEC) is a common method applied in the study of protein and peptide mixtures. The use of this type of chromatography has steadily increased over the years, from the early 1950's when the concept of size-based separation by chromatography was first recognized [97, 98]. SEC is now well-established separation method, also for high-performance liquid chromatography (HPLC) systems. The principle of molecular size separation in a typical column packed with porous particles is illustrated in Fig. 8. The figure shows how smaller molecules are absorbed into the pores of the stationary phase and retained, while larger molecules are not absorbed to the same extent and therefore elute faster. In the column the molecules are separated based on their hydrodynamic volume, that is, how much space a particular molecule occupies when it is in solution. The hydrodynamic volume of a molecule is affected by its surroundings and does not necessarily directly reflect mass of the

molecule [99]. However, several studies where mass spectrometry and ultraviolet (UV) are used as detection methods, show that hydrodynamic volume and the corresponding molecular mass are correlated. Calibration curves of known peptides and proteins can as a result be constructed and used for analysis of protein hydrolysates [100, 101]. It should be noted that proteins and peptides can exhibit various properties e.g. hydrophobicity and possibility for protonation. As a result, molecules having similar molecular mass may have very different hydrodynamic volumes. Some irregularities in calibration curves should therefore be expected.

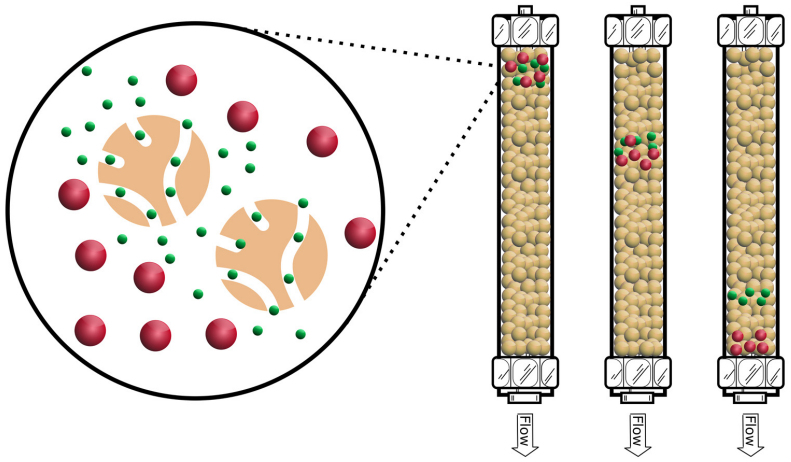



**Fig. 8.** An illustration of the separation principle in size exclusion chromatography.

Another major factor of error when separating and analyzing complex mixtures of proteins and peptides by SEC is the detection method. UV is the predominant detection method. Measurements at different wavelengths give specific information [102, 103]. The wavelengths of 270, 275 and 280 nm, are used to detect the aromatic amino acids (tryptophan, tyrosine and phenylalanine). Lower UV wavelengths (210, 214 and 220 nm) are used for detection of the amide peptide bonds. A major limitation when analyzing protein hydrolysates using UV detection is that only a few free amino acids are detected at these wavelengths, while proteins and peptides are detected by absorption contributions from both peptide bonds and side-groups at lower

wavelengths [103, 104]. Several studies addressing the issue of detecting of proteins, peptides and free amino acids using different detection methods have been published. An example is a comparative study by Petritis et al. where detectors for liquid chromatography are compared based on their ability to detect free amino acids in acid hydrolyzed peptide samples [105]. Another example is a study by Fujinari and Damon Manes which demonstrated the use of UV in combination with chemiluminescence nitrogen (CLN) detection [106]. Both studies concluded that the CLN detector is superior to UV for detection of a complete chromatographic profile of their samples. The specific nature of the CLN detector however, have made it difficult to implement it in standard laboratories. Thus, today it is not used extensively.

Since UV detection is the most commonly used method for SEC, despite of its limitations, it was also the method used in the present study. The molecular weight distribution (MWD) detected at 214 nm can used to calculate the weight-average molecule weight ($M_w$), which is a common quality parameter for EPH products. $M_w$ is calculated by the use of a calibration curve constructed by measurements of known analytical standards. Then, the calculation can be carried out using slicing method similar to those previously used for analysis of protein hydrolysates using Equation 2 [107].

$$M_w = \frac{\sum A_i M_i}{\sum A_i} \quad (2)$$

Here the mass distribution curve is divided into defined integration ranges, each with a detector signal $A_i$ and molar mass $M_i$ [23]. There are several different ways to interpret the MWD, another approach is to divide the chromatograms into regions corresponding to different size ranges [108-110]. The molecular weight distributions has as a result been used to study hydrolysates produced from a wide variety of raw materials and enzyme products and it has been linked to both functional and biological properties [35, 108-111].

### 3.1.7 Electrophoresis

Electrophoresis is comparable to SEC. The difference between the techniques is that migration and separation in electrophoresis is carried out under an electrical field pulling the charged molecules instead of relying on migration by a mobile phase [112]. Electrophoresis is not frequently used to monitor EPH of by-product raw materials, but methods like sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) have been used for characterizations of molecular weight and purity of protein hydrolysates. In a study on the peptide product inhibition in native-state proteolysis of two proteins, SDS-PAGE was used to quantify different size bands in the hydrolysis product [113]. Another study used a combination of capillary electrophoresis and chromatography to study both enzyme kinetics and specificity in hydrolysis of κ-Casein using two different chymosins [114].

### 3.1.8 Rheology

Measurements of rheological properties is essential in food science, and refers to flow phenomena of matter (solids, liquids, and gases) involving time-dependent behavior under the influence of stresses. The study of flow and deformation of matter is an important tool to characterize fundamental material properties of food systems [115]. Fluid mechanics is the foundation of food rheology and the correlation of stress and strain can be used to describe rheological properties of food systems in different models, including properties of protein hydrolysates [116, 117].

## 3.2 Fourier-transform infrared spectroscopy

Fourier-transform infrared (FTIR) spectroscopy is based on molecules' ability to vibrate or rotate when they absorb and are excited by infrared (IR) radiation. For a molecule to absorb IR radiation a change of the net dipole momentum of a chemical bond in the molecule must occur. This absorption is selective as the frequency of the radiation must match the energy gap between a ground state and an excited state. When the frequency matches, the radiation is absorbed, and molecular vibrations are induced [118, 119]. Examples of stretching and bending vibrations are illustrated in Fig. 9. The vibrational

frequency of a vibrating bond is dependent on the bond strength and the probability of the absorption changing the polarity. The most IR active part of a molecule is therefore the more polar groups, such as C=O, N–H and O–H. Other parts of the molecules that are less polar, such as C=C, S–S and aromatic rings, are barely seen in the IR spectra. These groups are better detected using Raman spectroscopy, which is another type of vibrational spectroscopy, complementary to FTIR. The strength of the vibrating bond and the probability of the absorption changing the polarity of a molecular bond are also influenced by intra- and intermolecular interactions. The proximate position of an absorption band in the IR spectrum is determined by the vibrating masses and type of bond. The exact position, on the other hand, is determined by electron withdrawing or donating effects caused by both the surrounding environment and coupling with other vibration systems [120, 121].

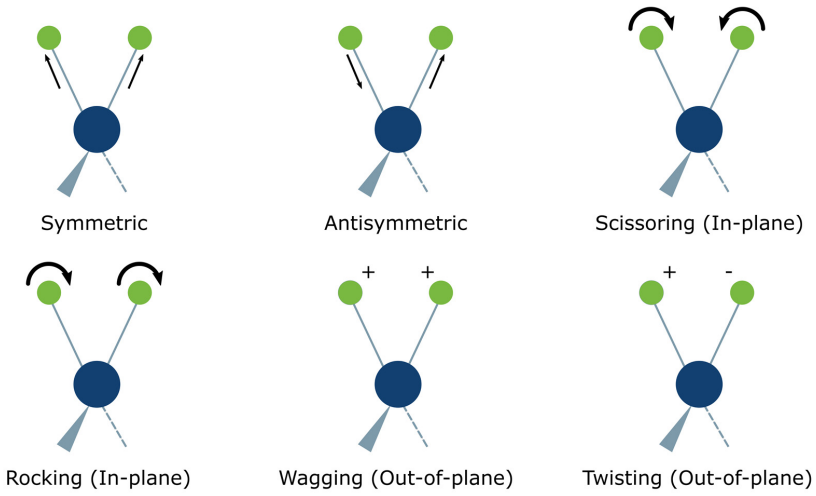



**Fig. 9.** Examples of molecular stretching and bending vibrations.

The development of IR techniques started with the discovery of IR radiation in the beginning of the 19th century. This was followed by the development of first IR detector in 1830 [122, 123]. Various ways to utilize infrared light have since then been established and the use have steadily increased since the first commercial IR

spectrometers became available in the 1940's [124, 125]. A frequently used technique for obtaining IR spectra today is based on the principle of FTIR spectroscopy. The basic idea of an FTIR instrument is the optical design that results in a radiation interference pattern called an interferogram. The interferogram is obtained as a function of change of the pathway between two interfering IR beams. One of beams faces the sample and the other is the reference beam. The detected interferogram is then converted from the time domain to frequency domain by applying Fourier transformation[119]. This technique is fast since it enables collection of absorbance from the whole wavelength range of interest simultaneously.

The IR spectrum is usually described in wavenumber $\bar{v}$, i.e., the inverse of the wavelength $\lambda$. IR radiation is in the wavenumber region between 12500-10 $cm^{-1}$ (0.8 μm to 1000 μm wavelength) of the electromagnetic spectrum, and the mid-infrared range span from approximately 4000 $cm^{-1}$ to 400 $cm^{-1}$ (i.e., 2.5 μm to 25 μm). The mid-infrared region is used to study the fundamental vibrations and associated rotational-vibrational combinations of the molecules [126]. Most organic compounds such as proteins absorb in the mid-infrared region giving rise to spectra containing information of both chemical composition and the structure of compounds. The use of this technique has, therefore, steadily increased over the years as new and improved instrumentation has been developed [127]. FTIR spectroscopy is now an established method for protein and peptide structural characterization. The technique has been used to study a wide variety of protein samples, from simple solutions of pure proteins to complex biological samples [121, 127-129].

### 3.2.1 FTIR spectroscopy for following proteolytic reactions

There are many reasons for the extensive use of FTIR spectroscopy for protein and peptide structural characterization. The method is rapid, non-destructive and little sample preparation is required. In the mid-infrared spectrum, the repeated bonds in the polypeptide chain of proteins and peptides give rise to distinctive IR absorption bands [130, 131]. The secondary amide groups of the protein backbone absorb strongly in the region 3500-3000 $cm^{-1}$ and in the region 1700-1200 $cm^{-1}$, as shown in Fig. 10. The

higher region is where the stretching vibrations of the N–H bond is found. There are two distinctive absorption bands that derive from this movement, i.e., the amide A and amide B. The amide A band (3310-3270 $cm^{-1}$) derives almost exclusively from the N–H group and is therefore in proteins insensitive to the conformation of the protein backbone, as its frequency depends on the strength of the hydrogen bond. The amide B band is usually a part of the amide A band by Fermi resonance doublet with the second component absorbing weakly between 3100 to 3030 $cm^{-1}$. The N–H stretching vibration is also resonant with an overtone of the amide II vibration [121].

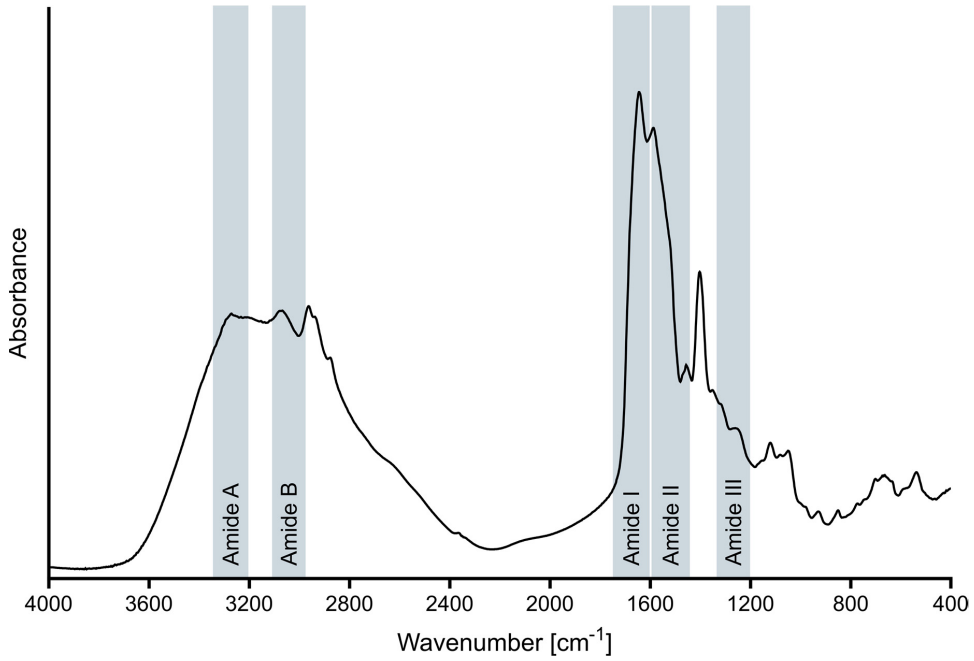



**Fig. 10.** An FTIR spectrum of a protein hydrolysate (4000-400 $cm^{-1}$) with indicated amide absorption bands.

In this thesis, the main spectral region studied has been between 1800 to 400 $cm^{-1}$. This region contains multiple absorption bands from proteins and peptides, but the amide I-III absorption bands are usually dominating in the spectra. The Amide I (~1650 $cm^{-1}$) arise mainly from the C=O stretch vibration, with some contributions from C–N

stretching vibration, the C–C–N deformation and the N–H stretch vibration [132]. The amide I vibration is hardly affected by the nature of the side chains. It depends, however, on the internal interactions of the protein backbone and is therefore commonly used for secondary structure analysis [121]. The amide II (~1550 $cm^{-1}$) arises from the out-of-phase combination of the N–H in plane bend and the C–N stretching vibration, with smaller contributions from the C=O in plane bend and the C–C and N–C stretching vibrations. Similar to the amide I vibration, the amide II vibration is not significantly affected by the side chains, and there is a correlation to the protein secondary structure. The frequency link to secondary structure is, however, less straightforward, as compared to the amide I band. The amide II band still provides valuable structural information, and it has been shown that the band can be used for secondary structure prediction alone [133]. The amide III (1400 to 1200 $cm^{-1}$) is a complex in-phase combination of vibrations such as the N–H bending, the C–N stretching vibration, the C=O in plane bending and the C–C stretching vibration. The amide III band is dependent on the side chain structure since N–H bending contributes to several modes in the 1400 to 1200 $cm^{-1}$ region. This band still contains information that can be used for secondary structure prediction [121].

FTIR spectroscopy has become an established method for protein and peptide structural characterization due to its sensitivity and ability to detect changes in their structures. The amide bands of proteins and peptides give rise to nine distinctive infrared IR absorption bands that can be used to monitor the protein backbone [130, 131]. This inherent ability opens for a range of possibilities to study parameters related to protein secondary structures. These parameters include hydration and solvent effects, pH and peptide size [134-139]. In a proteolytic reaction the peptide bonds are cleaved forming C-terminals ($COO^-$) and N-terminals ($NH_3^+$), consequently changing both the primary and secondary structures of proteins. Several studies on pure model proteins like hemoglobin, β-lactoglobulin, β-casein, and bovine serum albumin have demonstrated that FTIR spectroscopy can be used to monitor proteolytic reactions [140-144]. The amide absorption bands are also similar in different hydrolysate samples, even when comparing the spectral bands of a hydrolysate from a pure protein to a hydrolysate produced from a complex poultry by-product, as shown in Table 2. This makes it

possible to follow the development of the amide, C-terminal and N-terminal bands over time in EPH reactions. In a study published by Poulsen et al. the use of FTIR spectroscopy to monitor EHP reactions was expanded to complex samples where the development in the IR bands was linked to DH% and followed in EPH reactions of dairy proteins [145]. Researchers have recently extended this to even more complex EPH reactions, using salmon and poultry-based raw materials as substrates [24]. Also recently, a FTIR based multivariate approach for monitoring the change in $M_w$ during enzymatic hydrolysis of chicken by-products has been reported [23].

**Table 2:** 2nd derivative bands between 1700-800 $cm^{-1}$ for all sample materials.

| | | **Band positions [$cm^{-1}$]** | |
|---|---|---|---|
| **Annotation** | **Region** | **BSA** | **CMDR** |
| C=O amide I: turns | *i* | 1687 | 1675-1664 |
| C=O amide I: α-helix | *i* | 1656 | 1645 |
| $COO^-$(asym stretch) | *ii* | 1589 | 1583 |
| Amide II: α-helix | *ii* | 1548 | 1547 |
| $-NH_3^+$ (scissor) | *iii* | 1516 | 1516 |
| $CH_2$ (scissor) | | 1454 | 1454 |
| $COO^-$ (sym stretch) | *iv* | 1396 | 1405 |
| Amide III, $CH_2$ (def, rock), OH (def, bend) | | 1315 | 1313 |
| Amide III, C-O (stretch) | | 1240 | 1242 |
| $CNH_3$ (rock), $CH_2$ (wag) | *v* | 1115 | 1118 |
| CO, CC, CN (stretch) | *vi* | 1045 | 1045 |
| CCOO (wagging) | *vii* | 995 | 997 |
| $CH_2$ (twist) | *vii* | 928 | 928 |
| not assigned | | 850 | 851 |

The information in table is adapted from Böcker et al. [24]. Chicken mechanical deboning residues (CMDR). Bovine serum albumin (BSA).

There is a wide variety of sampling techniques and fibre-optic probes available, each with their own benefits and limitations [127]. The studies presented by Böcker et al. and

Wubshet et al. both use dry-film FTIR spectroscopy [23, 24]. This approach has an advantage compared to measuring aqueous samples as important information for following EPH reactions are found in the same spectral region as the water bands. In addition, the protein content in aqueous EPH samples are usually moderately low, and since one is looking at subtle differences in the protein backbone, the sensitivity enhancement achieved by drying is beneficial. The advantage is related to the peak resolution, sensitivity enhancement and information found in a dry-film spectrum over-shadow the increase in measurement time. A new approach also shows that a 30 seconds vacuum treatment to form dry-films prior to FTIR analysis results in better resolved peaks throughout the spectrum, preventing water related O–H stretching vibrations from dominating the spectrum causing the major loss in the protein signals [146].

## 3.3 Multivariate data analysis

In traditional analytical methods concentrations of analytes are typically determined using a calibration curve. An example is the TNBS method. In this method, known concentrations of Leucine react with TNBS and are measured in a spectrophotometer at 340 nm. The measured absorption for each calibration solution is then used to construct a linear calibration curve that again is applied to determine the amount of N-terminals in unknown samples. Information like DH% may also be found in FTIR spectra of the same samples. The challenge here is that molecular vibrations of all the molecules and their interactions contribute to absorption bands, resulting in complex FTIR spectra of multiple and collinear absorption values. Once the spectra are obtained, it is therefore necessary to apply appropriate multivariate analysis tools in order to access the "hidden" information in this complex and collinear data space. A common approach when analyzing multivariate datasets such as FTIR spectra is to apply principal component analysis (PCA). This method is a dimensionality reduction tool and can be used to reveal relations between samples by clustering similar samples in a two or three-dimensional score plot [147]. By applying multivariate regression methods, such as partial least squares regression (PLSR) modeling, the multivariate dataset $X(n \times m)$ can be related to different reference parameters $y(n \times 1)$, e.g. DH% and $M_w$. [148]. The

main principles of these approaches are shown in Fig. 11, where $n$ is the number of samples and $m$ is the number of variables (e.g. wavenumbers).

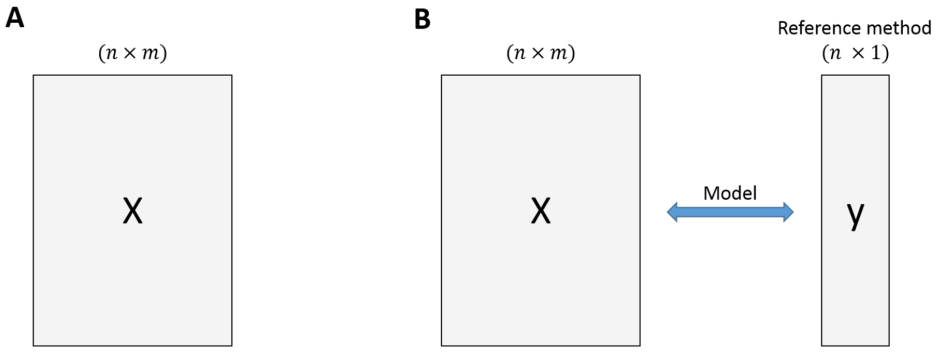



**Fig. 11.** A) Data matrix $X(n \times m)$ for exploratory data analysis PCA. B) The two-way data matrices $X(n \times m)$ and $y(n \times 1)$ to create a PLSR model $y = f(X)$.

### 3.3.1 Preprocessing of data

In vibrational spectroscopy, such as FTIR spectroscopy, the spectra are often affected by many physical phenomena other than the chemical components of interest. When the aim is to extract qualitative or quantitative information from the spectra, these phenomena will affect the results. Examples of such effects include random measurement noise, systematic errors related to, for example, non-linear instrument responses, and interfering effects from undesired chemical and physical variations. These challenges can often be reduced by sample preparation and analytical protocols. However, the optimal way to separate physical and chemical information in the spectra, is through mathematical preprocessing methods [149]. The calculation of derivatives of spectra has traditionally been used for removing baseline features in vibrational spectra. Multiplicative signal correction (MSC), on the other hand, was originally developed as a model-based preprocessing technique to separate physical and chemical features in near-infrared spectra [150, 151]. After the development of MSC, the extended version of MSC, namely EMSC, was developed. Basically, EMSC provides improved means for model-based preprocessing as it adds mathematical terms to the basic

preprocessing model used. It should be noted, however, that when proper baseline correction has been performed, the main effect of EMSC is normalization [149]. The same approach can be applied for other types of multivariate data, such as SEC chromatograms. For instance, in SEC chromatograms of protein hydrolysates, a reduction in signal due to reduced number of peptide bonds will affect the total absorption of the chromatogram, as mainly the peptide bonds absorb at 214 nm. Normalization can be used to account for this effect.

### 3.3.2 Principal component analysis

Principal component analysis (PCA) was introduced by Pearson in 1901 and is one of the most frequently used exploratory analysis techniques for multivariate data [147, 152]. PCA may be regarded as a dimensionality reduction tool, and by applying this method to the preprocessed FTIR spectra, the data set is reduced to single vectors in a low dimensional space, as shown in Fig. 12. The first component (PC1) describes the direction in the data set with the largest variance, and each succeeding (PC2, PC3 ...PCn) component accounts for as much of the remaining variability as possible, while being orthogonally to the previous components. Mathematically, PCA is described as shown in Equation 3.

$$X = TP^T + E \tag{3}$$

Here, the $X(n \times m)$ data matrix is transformed into score vectors $T(n \times R)$ where $R$ is the number of components, loading vectors $P(m \times R)$ and a residual matrix $E(n \times m)$ containing the unexplained variance, for example the random measurement noise.

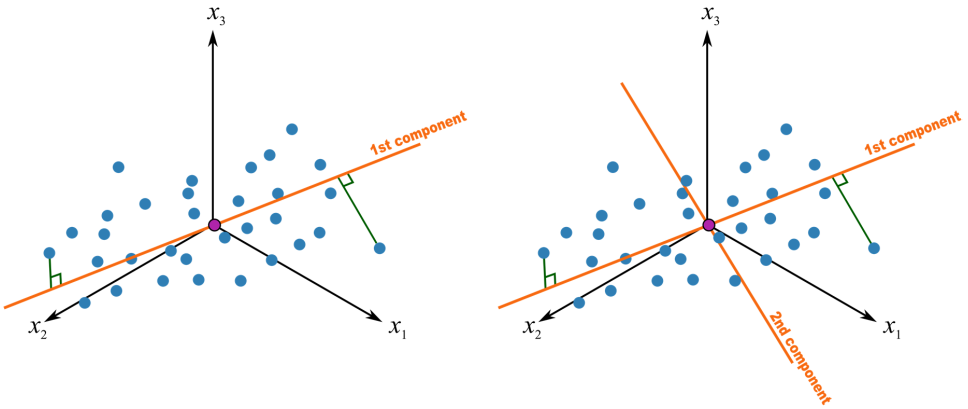



**Fig. 12.** A graphical presentation of PCA reducing multiple variables to a lower dimensional space. Image adapted from Dunn [153].

### 3.3.3 Partial least squares regression

Partial least squares regression (PLSR) is one of the most frequently used multivariate regression methods in spectroscopic analysis [148]. The method is similar to PCA in structure. The main difference is that instead of finding hyperplanes explaining maximum variance in $X$, PLSR finds the hyperplanes explaining the maximum variance between the response $y$ and the variables $X$, for example the spectral variables. In other words, the method relates two or several data matrices in a linear multivariate model. The method can be very useful due to the ability to analyze data with many, noisy, collinear, and even incomplete variables in both $X$ and $y$ [148]. This approach has consequently been widely utilized and is now a standard tool in chemometrics. Multiple versions of the PLSR approach have also been developed for handling different challenges. In this thesis, a two-level PLSR modeling approach was used to improve the accuracy of the predicted $M_w$ from the FTIR spectra from hydrolysates produced using different raw materials and enzyme products. This hierarchical modeling strategy has been exploited in methods such as hierarchical cluster-based partial least squares regression (HC-PLSR) and hierarchically ordered taxonomic classification by partial least squares (Hot PLS), in applications including nonlinear dynamic models and taxonomic classification, respectively [154, 155]. Successful applications of FTIR-based two-level partial least squares modeling have been demonstrated in the determination

of clinical parameters such as urea and glucose as well as complex protein structures [156, 157].

Mathematically, PLSR is described similarly to PCA. The difference is that in PCA the explained variation of $X$ is maximized, whereas in PLSR the covariation between $X$ and $y$ is maximized in a regression model. In the PLSR model the $X$ matrix (the predictors) is linked to the $y$ matrix (the responses). The columns in the matrix of scores $T$ are linear combinations of the data matrix $X$ that try to maximize the covariance between $T$ and the response $y$. This is achieved by finding the directions of highest variation in $X'y$ in the same way as PCA uses only $X$. This is shown by Equation 4 and Equation 5 describing PLSR [148].

$$X = TP^T + E_x \tag{4}$$

$$y = Tq + e_y \tag{5}$$

Here, $T(n \times R)$ is a matrix of scores, $P(m \times R)$ is a matrix of $X$-loadings, $q(R \times 1)$ is $y$-loadings, $E_x(n \times m)$ and $e_y(n \times 1)$ are the error residuals of $X$ and $y$, respectively. R is the number of latent variables included in the PLSR model. The PLSR model is fitted using a calibration sample set (a set of samples with known values of both $X$ and $y$). The PLSR model can then be used to obtain estimates of $y$ in future samples with known $X$.

### 3.3.4 Validation

Any calibration model needs validation. In this respect, validation means checking how well the model works when applied on other samples not directly involved in the development of the calibration. A frequently used approach to estimate the uncertainty in a calibration model, is the estimation of the root mean squared error (RMSE). The RMSE can be calculated by the measured reference values and the model estimated values for the calibration samples using Equation 6.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (6)$$

Here, $\hat{y}_i$ is the model estimated value of sample $i$, $y_i$is the reference value of sample $i$ and $n$ is the number of samples. The value of RMSE is in the same units and scale as the reference values, and robust calibration models have small RMSE values.

A cross-validation is in principle a loop where one or several samples are left out from the modeling at a time. A model is then created based on the remaining data and this model is then applied to the excluded samples. All sample segments are in turn left out from the modeling stage and each cross-validation loop predictions of the excluded samples are collected. Full cross-validation means that every sample in the data set is left out once. The predictions are then used to estimate the error term. In this way the contribution from individual segments to the total error originates from models not impacted by the individual segments. For the final model all samples are used, but the RMSE of the cross-validation (RMSECV) is reported. There are numerous different types of validation methods and only the principle of the methods used in this thesis is described here [158]. Usually, to thoroughly test the feasibility of a calibration model, independent test set validation is preferred.

# 4 Results and discussion

This chapter is divided into two parts. The first part starts with a short summary of the papers included in this thesis before presenting the main findings of each paper. The second part is based on the combined results of the papers placing them in an industrially relevant perspective.

## 4.1 Summary of papers

All papers included in this thesis address, in one way or the other, the need for new analytical approaches to understand and monitor proteolytic reactions. **Papers I** to **III** present new methods and approaches for characterization of proteolytic reactions using FTIR spectroscopy and classical analytical methods. These three studies address challenges in building reliable prediction models for monitoring EPH reactions and present new approaches and methods based on dry-film FTIR spectroscopy for following protein degradation in complex reaction mixtures. The hydrolyzed raw materials originated from dairy, poultry and marine sources. In **Paper I** two different PLSR approaches were compared for the prediction of $M_w$ in protein hydrolysates from different raw material and enzyme product combinations using FTIR. The results showed that the accuracy of prediction can be improved using a two-step PLSR modeling approach, where the spectra are assigned to predefined groups consisting of known enzyme-raw material combinations. In **Paper II**, FTIR spectroscopy was applied to predict DH% and $M_w$ in milk protein hydrolysates. The complimentary nature of DH% and $M_w$ was investigated and a link between the parameters was established by a PLSR model predicting DH% from molecular weight distributions. **Paper III** is a continuation of **Paper I**, where hierarchical modeling was suggested as an approach to increase the accuracy of the $M_w$ predictions. For the hierarchical modeling approach to be generic and industrially relevant as a tool to follow protein degradation, larger spectral libraries will be needed. This challenge is addressed in **Paper III** by introducing a new method capable of changing the protein structures and reducing spectral differences between hydrolysates derived from different enzyme-raw material combinations. At the same

time this method preserves important information contained within the spectra for following the proteolytic reactions.

The work presented in **Paper IV** was carried out simultaneously with **Papers I-III**. **Paper IV** is tightly linked to the three other papers by the aim and methods applied. It also underlines the need for online or at-line process monitoring tools in EPH processing of complex raw materials. Unlike **Papers I-III**, the focus of **Paper IV** was the effects of raw materials and proteolytic enzymes on the hydrolysate products. This study was, therefore, more explorative in the sense that additional analytical methods were applied in order to show differences between hydrolysates which is not necessarily detected by the quality parameters considered in **Papers I** and **II** (i.e., DH% and $M_w$). The study also included Angiotensin-I-converting enzyme (ACE) inhibition ability analysis. Inhibition of ACE is an essential strategy to reduce elevated blood pressure and hydrolysates are known do exhibit this ability [159, 160].

### 4.1.1 Paper I: FTIR-based hierarchical modeling for prediction of average molecular weights of protein hydrolysates

In this study, as illustrated in Fig. 13, dry-film FTIR spectroscopy was used to predict the $M_w$ of protein hydrolysates produced from protein-rich food-processing by-products using commercial enzyme products. Monitoring such processes is a significant challenge in industrial settings, as the classical analytical methods are not easily applicable. In the study, a generic FTIR-based approach for monitoring the $M_w$ of the protein hydrolysates during enzymatic hydrolysis of by-products from food industry was reported. A total of 885 hydrolysate samples based on 28 different combinations of enzyme products and by-products were studied. The FTIR spectra acquired from dry-films of the hydrolysates and their measured $M_w$ values were used to compare two PLSR approaches, a standard and a hierarchical approach. The hierarchical PLSR approach involved supervised classification of the FTIR spectra according to raw material quality and the enzyme products used in the hydrolysis process, and subsequent local regression models tuned to specific enzyme-raw material combinations. The results revealed that the most accurate predictions were obtained with the hierarchical PLSR approach. The RMSECV

and coefficient of determination of the cross-validation ($R^2$) were improved from 446 to 260 g/mol and from 0.83 to 0.94, respectively, using hierarchical PLSR. Overall, the results of the paper clearly underline the potential of using dry-film FTIR for monitoring protein sizes during enzymatic protein hydrolysis in industrial settings. This was also the first time a link between $M_w$ of proteins and FTIR spectra has been studied and established for an extensive set of samples.

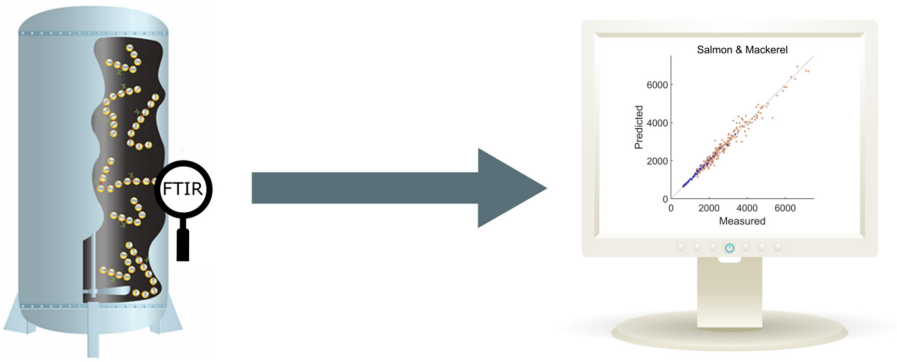



**Fig. 13.** A graphical abstract of **Paper I** [161].

The hierarchical PLSR approach for monitoring $M_w$, performed better than the standard PLSR approach for all combinations of raw materials and enzyme products. This showed that the predictive ability of the FTIR model towards protein size distribution is raw material and enzyme product specific. Differences can be observed in the optimized regression coefficients in the local models, two of which are presented in Fig. 14. The regression coefficients indicate which bands are important in the prediction model. These differences are linked to structure and composition differences seen in different FTIR bands. This demonstrates that the PLSR model is linked to raw material and enzyme product specific protein degradation patterns. The hierarchical PLSR approach for monitoring the $M_w$ can be regarded as an improved and more generic method and a follow up to the study by Wubshet et al. [23]. There are, however, several practical implications when the aim is to establish industrially feasible generic prediction models. For instance, if this approach is to be used in practical industrial environments,

extensive spectral libraries will be needed to predict $M_w$. In **Paper III** a method to reduce the raw material and enzyme-specific information in the FTIR spectra was therefore proposed.

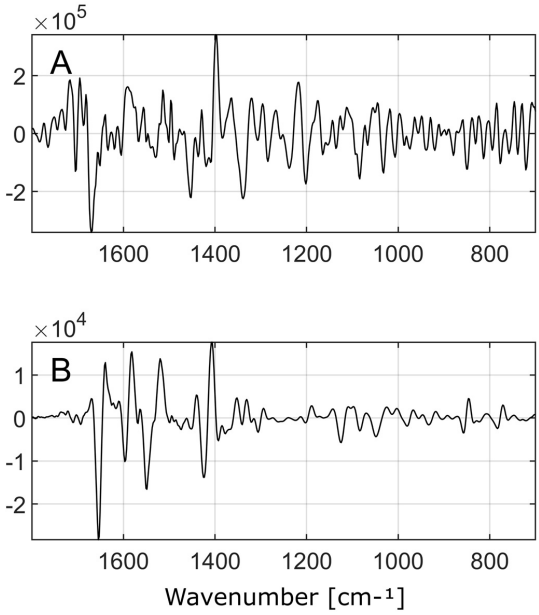



**Fig. 14.** Regression coefficients of the PLSR models of all hydrolysate time-series produced using: A) Chicken muscle and Alcalase. B) Turkey mechanical debone residue and Flavourzyme [161].

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

In this paper dry-film FTIR spectroscopy was used to predict DH% and $M_w$ in milk protein hydrolysates, as illustrated in Fig. 15. The hydrolysates were produced using three different milk raw materials and two different commercial enzyme products. DH% and $M_w$ are both important quality parameters of protein hydrolysates with a complimentary nature. Measuring these parameters and following their development

during proteolytic reactions could therefore be essential for process understanding and optimization in industry. In this study, the intercorrelation and the complementarity of these parameters were investigated and linked by a PLSR model predicting DH% from molecular weight distributions of the SEC chromatograms (RMSECV of 0.86% (DH%) and $R^2$ of 0.97). Finally, PLSR models based on dry-film FTIR spectroscopy for the prediction of both DH% and $M_w$ were developed. The results showed 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 ($M_w$), 1.3% (DH%), and $R^2$ of 0.91 ($M_w$), 0.93 (DH%)). The spectral changes observed in the amide region also showed the benefit of using dry-film spectra, as the amide I band was found to be important for the both calibration models. This shows that dry-film FTIR spectroscopy is a promising tool for dual prediction of DH% and $M_w$.

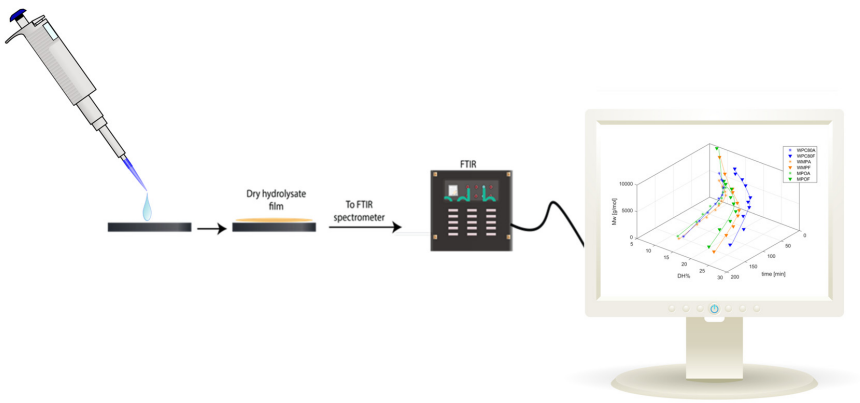



**Fig. 15.** A graphical abstract of **Paper II**.

The study also illustrated how simultaneous prediction of the two parameters may be used as a pragmatic indicator of the enzyme product's mode of action in a given reaction system. A general trend of an inverse correlation was observed when plotting $M_w$ against DH%, and each of the six hydrolysis reactions followed a different exponential decay trend. This can be explained by the use of different raw materials and enzyme

products. One of the two enzyme products used was Alcalase. This enzyme product mainly consists of endopeptidases, consistent with the observed relatively fast drop in $M_w$ with increasing DH% in the start of the reactions [44]. The other enzyme product used, Flavourzyme, consists mostly of exopeptidases, which digest peptides from the peptide ends, resulting in a slow reduction in $M_w$ with increasing DH% [44]. This difference can easily be observed at any specific DH% value in Fig. 16. Here, Flavourzyme samples always have higher $M_w$ as compared to the Alcalase samples from the same raw material: For example, WMP samples at DH% 10. At this DH% value WMP hydrolyzed with Flavourzyme has an $M_w$ of about 5500 g/mol while WMP hydrolyzed with Alcalase has an $M_w$ of approximately 2700 g/mol. This, along with 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. In other words, this study showed that dry-film FTIR spectroscopy is a promising tool for dual prediction of the parameters. It also showed that this information can give important insights related the enzymatic activity that may be used as a more powerful tool to control the EPH process if measured online or at-line.

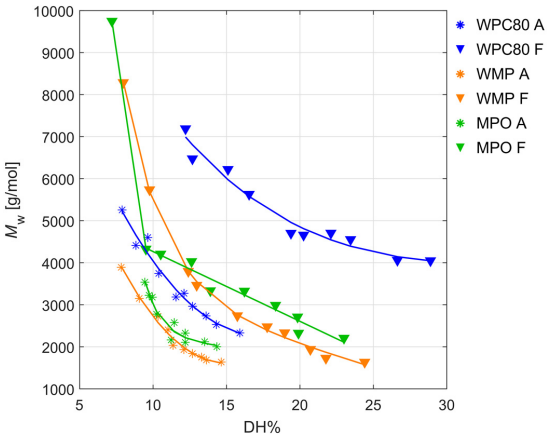



**Fig. 16.** $M_w$ vs. DH% for six hydrolysis reactions produced using three different raw materials (whey protein concentrate powder (WPC80), whole milk powder (WMP) and milk whey powder (WPO)) and two different enzyme products (Alcalase (A) and Flavourzyme (F)).

### 4.1.3 Paper III: Fourier-transform infrared spectroscopy for protein hydrolysate characterisation using dry-films treated with trifluoroacetic acid

In this study, as illustrated in Fig.17, the potential of using FTIR spectra of trifluoroacetate-protein complexes for gaining information related to enzymatic protein degradation and protein characterization was explored for the first time. Protein hydrolysates originating from different raw materials and even different enzyme products display unique FTIR fingerprints, as discussed in **Paper I**. This is a challenge when aiming to establish generic prediction models for protein degradation, as large spectral libraries will be needed. Trifluoroacetic acid (TFA) is known for its ability to interact with and modify the structure of proteins and peptides [162-164]. The idea of treating dry-films of proteins with TFA prior to FTIR analysis was based on the unique properties of TFA. By adding a large excess of TFA to protein hydrolysate samples, the possible protonation sites of the proteins and peptides will be saturated. This includes the oxygens of the secondary amides in the protein backbone [165, 166]. The resulting protein denaturation and protonation will occur using any strong acid. However, TFA has a low boiling point when protonated, and it also has complex forming abilities. When forming TFA-treated dry-films of hydrolysates, the excess TFA acid will evaporate and the deprotonated acid ($CF_3COO^-$) will act as a counter ion with the positive charges of the sample material [167, 168]. During an EPH reaction the ratio of N-terminals relative to peptide bonds (secondary amides) is increasing. This ratio will be proportional to the $CF_3COO^-$ complexes formed as a film after the excess TFA evaporation. Therefore, as the proteolytic reaction proceeds, systematic changes are expected in both the $CF_3COO^-$ and amide absorption bands of the denatured proteins and peptides in the corresponding FTIR spectra. In the study, spectral changes in TFA-treated dry-films of a pure protein as well as complex protein mixtures were compared to the FTIR fingerprints of untreated dry-films. The results showed that time-dependent information on EPH and, consequently, on the characteristics of a protein hydrolysate can be obtained.

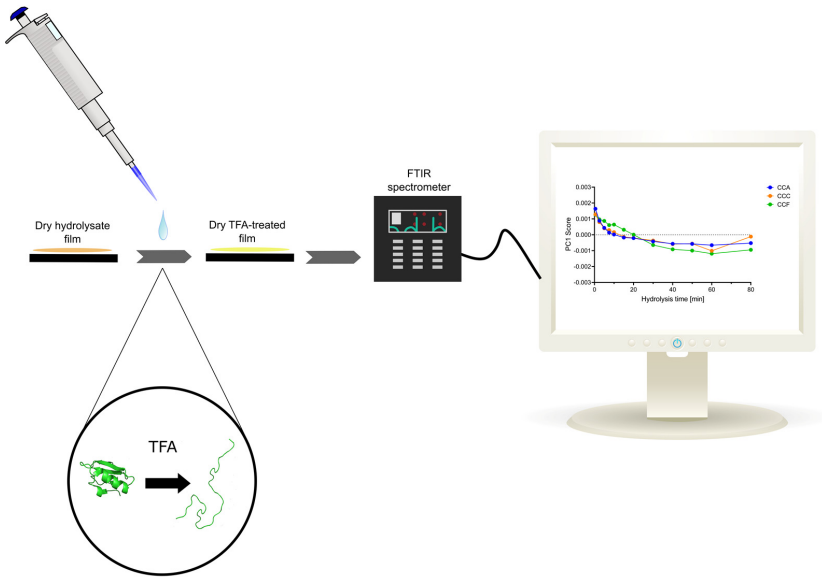



**Fig. 17.** A graphical abstract of **Paper III**.

The effects of $CF_3COO^-$-interactions with protonated proteins and peptides can be observed in the FTIR spectra, as shown in Fig. 18. The figure (Fig. 18A and 18C) shows the second derivative FTIR spectra of two EPH time-series from poultry. The spectra from the two time-series are visually different, and large differences are seen in the *i-iii* regions where important IR bands for EPH monitoring, such as the. amide I and II bands, the N-terminal ($NH_3^+$) and the C-terminal ($COO^-$), are located. The amide band region also contains secondary structure attributes. The secondary structure differences result in complexity difference observed in the spectra derived from the two EPH time-series. The turkey raw material contains more collagen than the chicken, and Flavourzyme (exopeptidase) is less efficient when it comes to degradation of these structures as compared to Alcalase (endopeptidase) [44]. In summary, differences like the ones shown in Fig 18A and 18C are typical for hydrolysates from different raw materials and enzyme products. The spectra of the same samples treated with TFA are shown in Fig. 18B and 18D. By comparing the untreated spectra with the TFA-treated spectra, the differences between the spectra are dramatically reduced (i.e., Fig. 18A vs. 18C as compared to Fig. 18B vs. 18D) and the time-series become more similar. The new

$CF_3COO^-$ bands observed in the treated spectra are also directly comparable to the bands described for the TFA-treated BSA spectra (see **Paper III**). Compared to untreated dry-films, where protein degradation patterns are related to raw-material specific changes in secondary structures, TFA-treated samples reveal protein degradation patterns related to a more generic $CF_3COO^-$ counter ion effect.

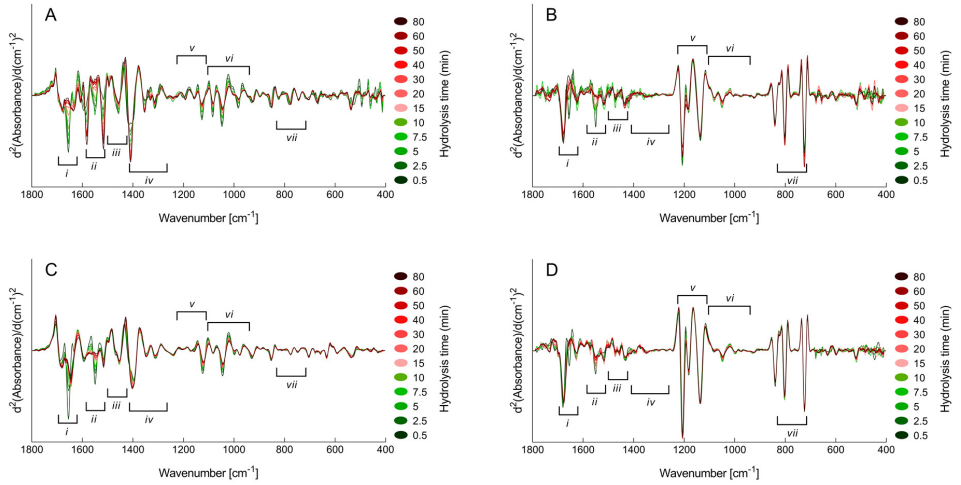



**Fig. 18.** Second derivative FTIR spectra of two different hydrolysis time-series, treated with TFA and untreated. A and B) Turkey mechanically deboned residue hydrolyzed with Flavourzyme. C and D) Chicken carcass hydrolyzed with Alcalase. A and C) Untreated dry-film spectra. B and D) TFA-treated dry-film spectra.

### 4.1.4 Paper IV: Effects of poultry raw material variation and choice of protease on protein hydrolysate quality

**Paper IV** presents a process-related study where analytical data was collected from poultry raw materials, the enzymatic hydrolysis process and product, as illustrated in Fig. 19. The hydrolysates were prepared using four different raw materials from poultry production and three different commercially available proteolytic enzymes. The raw materials and products were subjected to fat, ash, protein, and amino acid composition analysis, and the EPH processes were monitored by measuring $M_w$. The product was also

subjected to DH%, rheological and ACE inhibition ability analysis. The results showed that both the enzyme product and raw material highly influenced the outcome of the EPH processes. It was also shown that there are differences in the products that are not easily linked to gross quality parameters monitored during the process, such as $M_w$ and DH%.

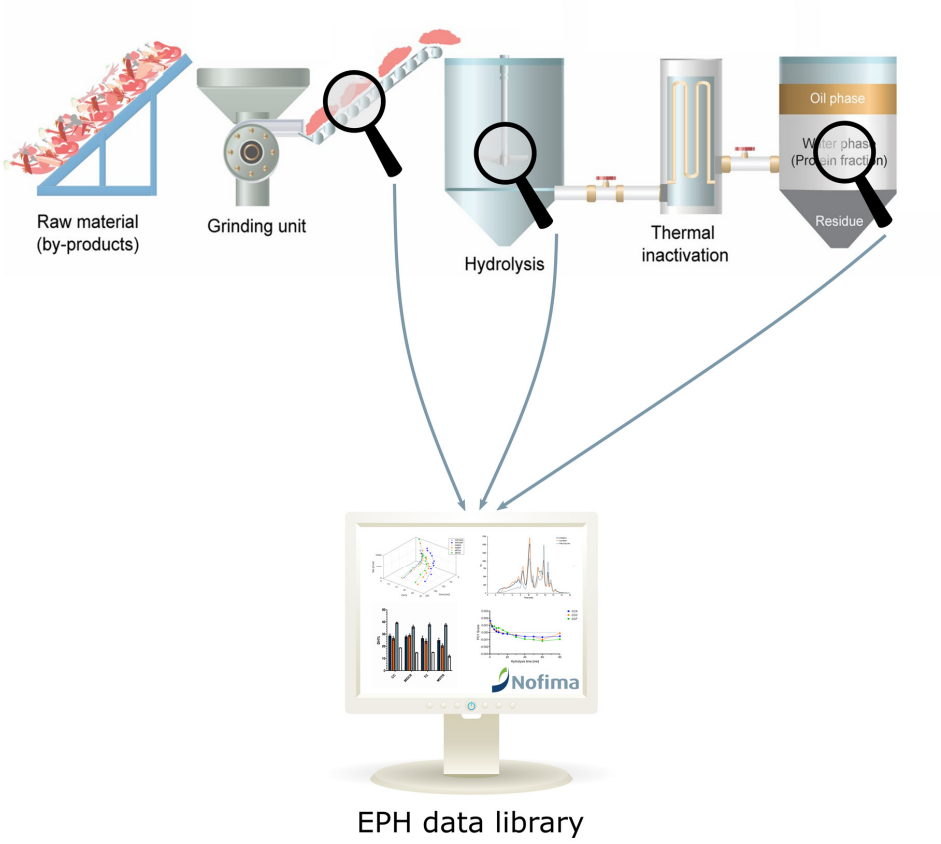



**Fig. 19.** A graphical abstract of **Paper IV**.

The MWD of a protein hydrolysate is known to be affected by both the raw material composition and the enzyme product. By comparing the MWD profiles it is possible to see differences that are linked to both raw material and the enzyme product used, as shown in Fig. 20. In **Papers I-III**, $M_w$ was used as a parameter to follow the EPH

reactions. This parameter is derived from the whole MWD profile and it is not sensitive to small differences in the MWD profile. In this study, the chromatograms were therefore divided into four size ranges as well. Interpretation of the MWD dividing the chromatograms into size ranges has been used in several studies of protein hydrolysates [108-110]. In Fig. 20, the MWD are divided according to raw material and the enzyme products (specified by color codes). By comparing the relative area of the size ranges between samples, it is likely that smaller differences, which do not significantly affect the calculated $M_w$ between the hydrolysate products, will be detected.

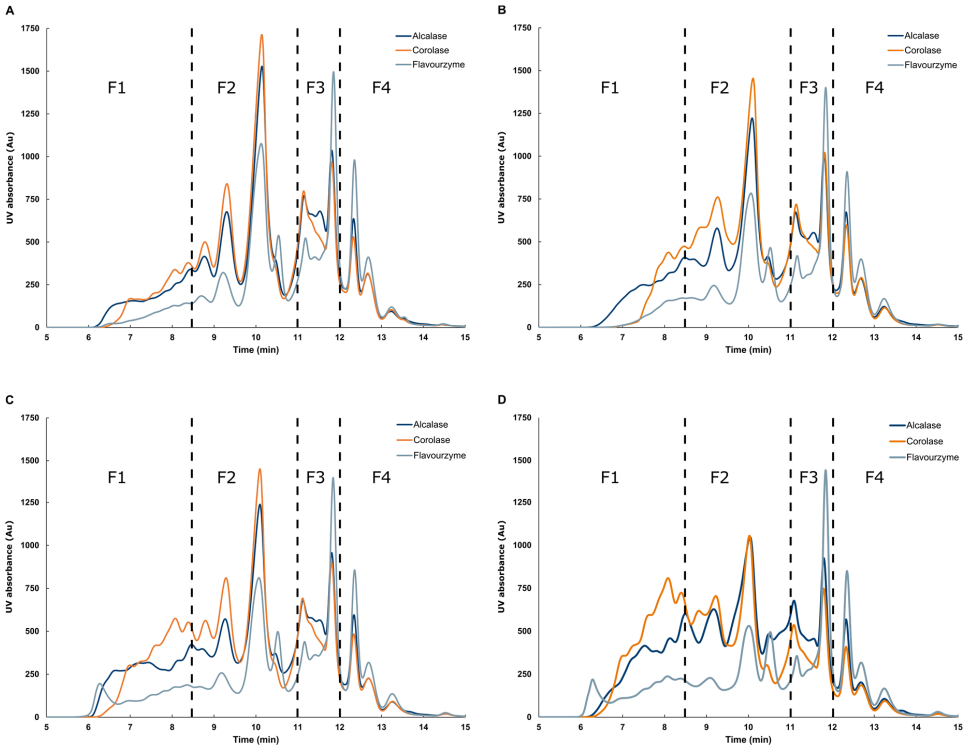



**Fig. 20.** The resulting size exclusion chromatograms for the hydrolysates produced using three enzyme products (i.e., Alcalase, Corolase 2TS and Flavourzyme) and four poultry raw materials, hydrolyzed for 80 minutes. The MWD profiles were obtained using 2.5% w/v hydrolysate samples from: A) Chicken carcass. B) Chicken mechanical deboning residue. C) Turkey carcass, and D) Turkey mechanical deboning residue.

## 4.2 General discussion

The aim of the thesis was to expand the analytical toolbox for proteolytic reactions for both lab-scale and industrial scale reactions. All hydrolysis reactions studied in the thesis were carried out in lab-scale. The results of this thesis illustrate the potential of FTIR spectroscopy for fast prediction of quality parameters. We have found that FTIR can be used for prediction of the common quality parameters of EPH, i.e., DH% and $M_w$, with high accuracy for complex samples. There are, however, some elements that have not been given a lot of attention in the presented papers, such as errors in the reference measurements and further method development and improvements. This general discussion is, as a result, focused around method errors, method development and improvements and how the results of the papers can be applied industrial settings.

In **Paper I** it was suggested that the some of the errors (RMSECV) in prediction of higher $M_w$ may be related to inaccuracies of the reference analysis. The error in the $M_w$ value can be a result of multiple factors, such as errors related to the SEC column exclusion range and the detection method used. In this thesis, UV detection was used to measure peptide bonds at 214 nm. This is the most common wavelength used for this purpose. The limitations of the method are as previously described linked to how free amino acids are detected compared to proteins and peptides at the wavelength used [98, 104]. This, together with poor retention of larger peptides and protein fragments, will result in scaling errors, which in turn will affect how well the calculated $M_w$ reflects the actual MWD in the sample. An improvement in RMSECV could perhaps be possible using a CLN detector in the SEC setup, as it provides a total MWD profile detecting free amino acids as well as peptides and proteins [105, 106]. Likewise for DH%, the measurements may contain errors related to the methods used and technical errors [92]. The TNBS method used in **Paper II** and **IV** was selected based on the study by Spellman et al. who reported the TNBS method to be an excellent method for quantifying DH% in a EPH study where whey protein was hydrolyzed using the enzyme products Alcalase and Debitrase HYW20 [169]. In **Paper II**, it was also shown that the reproducibility is higher for $M_w$ then DH%. This illustrates that there are differences in reproducibility and errors that should be taken into account when using this methodology as reference measurements for FTIR spectroscopy.

The hierarchical PLSR approach presented in **Paper I** confirmed that the predictive ability of FTIR towards protein $M_w$ is raw material and enzyme specific. This was seen in the regression coefficients of the local PLSR models. The raw material and enzyme specificity have several practical implications, and larger spectral libraries will be needed to account for the specific protein degradation patterns. An alternative to the hierarchical approach proposed could be the locally weighted regression (LWR) method. The LWR method searches for similar spectra in order to choose samples for a calibration set, but the method still requires a substantial database [170, 171]. A challenge that arose from **Paper I** was how to reduce the raw material and enzyme-specific information in the FTIR spectra, while keeping the information related to, e.g. DH%. In **Paper III** we, therefore, proposed a new method that can reduce protein secondary structure attributes in the FTIR spectra by denaturation.

In **Paper III** PCA was applied in a similar manner to Böcker et al., it was found that FTIR bands of TFA-treated EPH samples contain information related to the state of the EPH reaction [24]. The study showed that the spectra of trifluoroacetate-protein complexes contain important information for monitoring EPH reactions. Compared to untreated dry-films, where protein degradation patterns are related to raw-material specific changes in secondary structures, TFA-treated dry-films reveal protein degradation patterns related to a more generic $CF_3COO^-$ counter ion effect. Protonation of a secondary amides are known to affect the amide II band, as it changes the C-N and N-H bond length of the corresponding amide group [172]. The C-N bond becomes shorter when the carbonyl oxygen is protonated [165]. This gives the C-N bond a stronger double bond nature and the nitrogen will have a net positive charge which are stabilised by the $CF_3COO^-$ counter ion. The $CF_3COO^-$ interaction with secondary amide has been assigned to a band around the 1620 $cm^{-1}$ region [168]. When the spectrum of a TFA-treated sample was compared to the untreated, it was concluded that original elements from amide I and II were preserved. New bands from the $CF_3COO^-$ of the trifluoroacetate-protein and peptide complexes were also clearly visible, and the most dominant were the C=O (~1677 $cm^{-1}$), C-F (~1250-1100 $cm^{-1}$) and OCO (~950-700 $cm^{-1}$) stretching bands. The equilibrium between $CF_3COO^-$ and the many possible resonance structures stabilizing the positive charges of proteins and peptides are as

described above a well-known phenomenon. Yet, the vibrating nature of $CF_3COO^-$ when interacting with different parts of the protonated proteins and peptides are not fully described nor understood, as it is for the amide absorption bands. An investigation and a full assignment of trifluoroacetate-protein complex bands should therefore be carried out. The FTIR spectra of TFA-treated samples should also be linked to parameters such as $M_w$ and DH%. This is to investigate if a standard PLSR model based on dry-film spectra of trifluoroacetate-protein complexes is, in fact, a more generic approach as compared to the hierarchical approach presented in **Paper I**.

Another challenge that was encountered when working with TFA treatments of protein hydrolysate dry-films was that the surface of the treated films was uneven. There was also issues with the sample moving over the sample well lines on the Si-plate during the drying process. An uneven film surface will affect the FTIR spectrum as it will scatter the IR radiation differently compared to a more even film. Several TFA concentrations and solvent compositions were tried in order to reduce this issue. From this investigation it was found that treating dry-films of the protein hydrolysates with TFA in water gave satisfactory results. The ion effects caused by the TFA treatment changes the surface tension of the sample. These effects are constantly changing as the solvent evaporates during the film forming process. The effects of ions on surface tension is well-known and methods to reduce the effect when forming films have been developed [173, 174]. An example where surface tension effects are reduced, is in the print process of biomolecules on surfaces by extremely small droplets for diagnostic applications [175, 176]. Another approach that can resolve this issue is to change the hydrophobicity of the Si-plate or the relative humidity upon drying of the samples on the plate [177, 178]. This shows that there are possible solutions to the issues encountered in **Paper III**. It also shows that several studies are needed before a method based on the TFA interaction principle can be established for following quality parameters in proteolytic reactions.

**Paper IV** is a paper where several quality parameters were obtained in order to study the effects of the raw materials and enzyme products used in EPH processes. The paper

showed that there are differences in compositional, functional, and biological properties of protein hydrolysates that are not easily linked to the corresponding DH% of the samples. An example is the link between DH% and ACE inhibition ability. Several studies have shown that ACE inhibition (which also was measured in **Paper IV**) can be linked to DH%, which then can be used to optimize EPH processes [179, 180]. The results of the study showed that hydrolysates produced using the enzyme product Corolase 2TS exhibited the highest ACE inhibition activity. The hydrolysates products using Alcalase exhibited some activity and hydrolysates produced using Flavourzyme had little or no ACE activity. Comparing the DH% values, only minor differences between hydrolysates produced using the same enzyme product were observed. For example, all hydrolysates prepared using Corolase 2TS showed that differences in the ACE activity were relatively large while DH% differences were minor. Taking into account what was shown in **Paper II**, where DH% measurements had a relatively low reproducibility as compared to $M_w$, it is clear that optimizing an EPH process only using DH% can be a challenge. This is especially true when complex raw materials are used for production of protein hydrolysates with potential ACE inhibition ability. A combination of $M_w$ and DH%, on the other hand, could serve as a more powerful optimisation tool. However, as the mechanisms for ACE inhibition are much more complex than peptide length and DH%, it will have limitations, especially when comparing hydrolysates produced using complex raw materials.

The main finding of **Paper IV**, in the context of **Papers I-III**, is that it underlines the need for process control in EPH processes due to the inherent variations in raw material. It has been postulated that there is significant correlation between raw material properties and process parameters, as well as within the process parameters themselves [22]. Several of the analytical techniques described in Chapter 3 and the data collected using them could consequently be used in industrial environments if they fit into a process optimization and control scheme. EPH processes can be controlled using feed-forward or feed-backward approaches, or as a combination of the two [21, 22]. The feed-backward scheme is today the most commonly used strategy. In this approach control action is based on product quality measures. The disadvantage of the feed-backward approach for enzymatic hydrolysis is that the control action is taken after the

product already has been produced. Since **Paper IV** showed that raw material variation highly affects the product, a feed-forward scheme would be a more appropriate approach. In a feed-forward approach, measured raw material composition can be used as input. This together with the processing settings can be used to make statistical models for predicting process outcomes. Fast raw material composition analysis can be achieved using vibrational spectroscopy, linking the spectra to parameters such as fat, protein and water content [181-184]. Linking the raw material input to the development in quality parameters such as $M_w$ and DH% during EPH processes could be used to adjust process parameters to reduce variations in product quality. The datasets presented in this thesis are, however, relatively small and focus on raw material variation, enzyme product type and hydrolysis time. To use the feed-forward approach in an industrial setting for fine tuning the EPH process, more knowledge about how other process parameters such as temperature, pH and enzyme (type and amount), stirring effects, and others affect the yield and product quality is needed.

**Papers I**, **II** and **IV** show that there is a need for better optimization tools in EPH processes, and that there are spectroscopic tools available that can be used to monitor and control raw material variations, products and the ongoing EPH process. Industrial EPH processes are very complex, and the complexity of bioprocesses and their raw materials will call for extensive data sets and multivariate data handling approaches. It has previously been shown that models of this type can be obtained using multiblock approaches [185, 186]. The models are empirical and depend on reliable data that span over a relevant variation range. **Paper I** showed how FTIR can be used to follow the kinetics of protein degradation. This is achieved in lab-scale through the use of several different enzyme products and raw materials. This approach is often the best way to collect such data, as statistical design of experiments can spread out a large variation. In industrial scale EPH production this not practical, as spreading out a large variation will be extremely resource demanding. In addition, there are multiple phenomena appearing in industry which do not occur on a lab-scale. Thus, obtaining mathematical models that are representative at an industrial scale from lab-scale data, are of course challenging. One way to solve this issue is to use a combined strategy of datasets that are constructed based on lab-scale and data from an industrial scale EPH processes. The lab-scale data

can be used to identify the critical process parameters and their correlation to raw material characteristics. In the next step, the data can be compared to data of samples from industrial scale EPH production collected over a period of time. The samples from the industrial scale EPH process will not have the same variation range and statistical properties as the lab-scale results, but they can be used to validate the lab-scale findings and adjust the model to better fit an industrial process.

Valorization of protein-rich by-products from for example meat and fishery industries using EPH processes is a growing industrial segment. For these sectors, there is a need for robust and fast industrial tools for following product quality development during processing to ensure quality stability. The present thesis clearly shows that dry-film FTIR in combination with multivariate analysis and classical methods have the potential to be used as a tool in EPH process control schemes. The data collected can also be used in combination with previously collected and new data to gain more insights to EPH processes by applying the appropriate multivariate data analysis algorithms.

# 5 Conclusion and future aspects

The main objective of this thesis was to expand the analytical toolbox for characterization of proteolytic reactions. The studies have presented new approaches and methods for the characterization of proteolytic reactions based on the combination of FTIR spectroscopy and classical analytical methods. The work expands upon the existing analytical toolbox as it offers fast prediction of classical parameters using FTIR spectroscopy. The results also represent a substantial collection of data that may be used in future projects for improving the process control in the production of protein hydrolysates. There is, however, still work to be done and challenges that need to be overcome in order to take full advantage of all the presented results.

**Papers I** and **II** demonstrated that dry-film FTIR spectroscopy, in combination with multivariate analysis and classical analytical methods, can be used to monitor the development of common quality parameters in EPH processes, i.e., $M_w$ and DH%. In the context of these papers we will, in the future, need to employ a proper test set validation of the regression models when larger datasets are collected. DH% prediction will also need to be extended to larger and more complex sample sets which already exist for $M_w$. To transfer the methods described in **Papers I** and **II** to industrial settings, an online or at-line system for sampling of the aqueous phase and dry-film formation needs to be developed. There are examples of FTIR setups where sample preparation for dry-film analysis is automated [187, 188]. Other developments, such as the novel quantum cascade lasers system, could also potentially be applied in order to reduce the time from when the sample is collected to when the prediction result is obtained [189]. These examples show that there are possibilities for adapting existing and new technologies for monitoring EPH processes online or at-line.

In **Paper III**, the use of TFA-treated dry-films of EPH samples as a strategy for FTIR based characterization of protein hydrolysates was assessed. This was achieved by comparing the spectra of untreated to treated samples. The results showed that the typical protein degradation patterns related to raw-material specific changes in

secondary structures were reduced in the spectra of the TFA-treated films. The protein degradation patterns observed in the TFA-treated spectra were also strongly based on $CF_3COO^-$ counter ion effect which were shown to be more generic. The results were promising but further development and improvements are needed. This includes solving the issue with dry-film quality, assignment of absorption bands and construction of calibrations for prediction of e.g. DH%. There are several possible methods that might be applied in order to achieve uniform surfaces of TFA-treated samples. This includes methods such as sample printing, changing the hydrophobicity of the sample plates or changing the relative humidity upon drying of the samples [175-178]. A more thorough study of absorption bands and assignment of the bands for a trifluoroacetate-protein complex could give important information for future studies and development of this method. This could be achieved by studying a simple and short polypeptide such as a Gly-Gly-Gly chain with and without TFA treatment. A link between the spectra of trifluoroacetate-protein complexes and a quality parameter will also be needed. From the nature of interaction described between $CF_3COO^-$ and the protonated proteins and peptides, we believe that DH% is the best option since this parameter is a measure of N-terminals relative to peptide bonds.

**Paper IV** demonstrated that there are differences in amino acid composition, as well as functional and biological properties in the hydrolysates prepared using different poultry raw materials and enzyme products. In this study, compositional data of the raw materials and products were collected. The EPH processes was also monitored by $M_w$, and DH% and MWD profiles of the hydrolysate products were obtained. The different raw material and enzyme combinations followed different $M_w$ development trends which showed that the protein degradation is highly raw material and enzyme dependent in complex EPH processes. The data collected for **Paper IV** and the other papers of this thesis will clearly be useful in future EPH studies.

Overall, the present thesis represents a step forward towards the increased use and valorization of existing protein resources, as it offers new methods and approaches to monitor, control and understand ongoing EPH processes in industrial settings.

# 6 References

[1] The Intergovernmental Panel on Climate Change, Climate Change and Land—An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems, 2019.
[2] Food and Agriculture Organization of the United Nations (FAO), Food Outlook—Biannual Report on Global Food Markets, 2019.
[3] Food and Agriculture Organization of the United Nations (FAO), The State of Food and Agriculture: Climate Change, Agriculture and Food Security, 2016.
[4] United Nations Environment, Global Environment Outlook (GEO-6), 2019.
[5] United Nations Climate Change Secretariat, Climate action and support trends, 2019.
[6] P. Smith, M. Bustamante, H. Ahammad, H. Clark, H. Dong, E.A. Elsiddig, H. Haberl, R. Harper, J. House, M. Jafari, O. Masera, C. Mbow, N.H. Ravindranath, C.W. Rice, C. Robledo Abad, A. Romanovskaya, F. Sperling, F. Tubiello, Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 2014.
[7] Department of Economic and Social Affairs Population Division, World Population Prospects United Nations, 2019.
[8] D. Lindberg, K. Aaby, G.I.A. Borge, J.-E. Haugen, A. Nilsson, R. Rødbotten, S. Sahlstrøm, Kartlegging av restråstoff fra jordbruket, Nofima, 2016.
[9] P.J. Gerber, H. Steinfeld, B. Henderson, A. Mottet, C. Opio, J. Dijkman, A. Falcucci, G. Tempio, Tackling climate change through livestock—A global assessment of emissions and mitigation opportunities, Food and Agriculture Organization of the United Nations (FAO), 2013.
[10] The Paris Agreement in: United Nations Framework Convention on Climate Change (UNFCCC) (Ed.) 2015.
[11] R.J. Brooker, E.P. Widmaier, L.E. Graham, P.D. Stiling, Biology 1ed., McGraw-Hill Higher Education 2008.
[12] D.L. Nelson, M.M. Cox, Principles of biochemistry, 5 ed., W. H. Freeman and Company 2008.
[13] D. Voet, J.G. Voet, Biochemistry, 4 ed., John Wiley & Sons 2008.
[14] Wikipedia, Amino Acids 2010. https://commons.wikimedia.org/wiki/File:Amino_Acids.svg. (Accessed 31 Agust 2019).
[15] Wikipedia, Main protein structure levels, 2008. https://en.wikipedia.org/wiki/File:Main_protein_structure_levels_en.svg. (Accessed 31 August 2019).
[16] M.A.M. Ahmed, V.V. Bamm, L. Shi, M. Steiner-Mosonyi, J.F. Dawson, L. Brown, G. Harauz, V. Ladizhansky, Induced Secondary Structure and Polymorphism in an Intrinsically Disordered Structural Linker of the CNS: Solid-State NMR and FTIR Spectroscopy of Myelin Basic Protein Bound to Actin, Biophysical Journal 96(1) (2009) 180-191.
[17] A. Hagenau, P. Papadopoulos, F. Kremer, T. Scheibel, Mussel collagen molecules with silk-like domains as load-bearing elements in distal byssal threads, Journal of Structural Biology 175(3) (2011) 339-347.
[18] T. Aspevik, Å. Oterhals, S.B. Rønning, T. Altintzoglou, S.G. Wubshet, A. Gildberg, N.K. Afseth, R.D. Whitaker, D. Lindberg, Valorization of Proteins from Co- and By-Products from the Fish and Meat Industry, Topics in Current Chemistry 375(3) (2017) 53.
[19] O.L. Tavano, Protein hydrolysis using proteases: An important tool for food biotechnology, Journal of Molecular Catalysis B: Enzymatic 90 (2013) 1-11.

[20] T. Lafarga, M. Hayes, Bioactive protein hydrolysates in the functional food ingredient industry: Overcoming current challenges, Food Reviews International 33(3) (2017) 217-246.
[21] S.G. Wubshet, J.P. Wold, N.K. Afseth, U. Böcker, D. Lindberg, F.N. Ihunegbo, I. Måge, Feed-forward prediction of product qualities in enzymatic protein hydrolysis of poultry by-products: a spectroscopic approach, Food and Bioprocess Technology 11(11) (2018) 2032-2043.
[22] S.G. Wubshet, D. Lindberg, E. Veiseth-Kent, K.A. Kristoffersen, U. Böcker, K.E. Washburn, N.K. Afseth, Chapter 8 - Bioanalytical Aspects in Enzymatic Protein Hydrolysis of By-Products, in: C.M. Galanakis (Ed.), Proteins: Sustainable Source, Processing and Applications, Academic Press 2019, pp. 225-258.
[23] S.G. Wubshet, I. Mage, U. Bocker, D. Lindberg, S.H. Knutsen, A. Rieder, D.A. Rodriguez, N.K. Afseth, FTIR as a rapid tool for monitoring molecular weight distribution during enzymatic protein hydrolysis of food processing by-products, Analytical Methods 9(29) (2017) 4247-4254.
[24] U. Böcker, S.G. Wubshet, D. Lindberg, N.K. Afseth, Fourier-transform infrared spectroscopy for characterization of protein chain reductions in enzymatic reactions, Analyst 142(15) (2017) 2812-2818.
[25] A.S. Bommarius, B.R. Riebel, Biocatalysis: fundamentals and applications, Wiley-VCH 2004.
[26] A.G.B. Wouters, I. Rombouts, E. Fierens, K. Brijs, J.A. Delcour, Relevance of the Functional Properties of Enzymatic Plant Protein Hydrolysates in Food Systems, Comprehensive Reviews in Food Science and Food Safety 15(4) (2016) 786-800.
[27] A. Radzicka, R. Wolfenden, Rates of Uncatalyzed Peptide Bond Hydrolysis in Neutral Solution and the Transition State Affinities of Proteases, Journal of the American Chemical Society 118(26) (1996) 6105-6109.
[28] H.G. Kristinsson, B.A. Rasco, Fish protein hydrolysates: Production, biochemical, and functional properties, Critical reviews in food science and nutrition 40(1) (2000) 43-81.
[29] M. Friedman, R. Liardon, Racemization kinetics of amino acid residues in alkali-treated soybean proteins, Journal of Agricultural and Food Chemistry 33(4) (1985) 666-672.
[30] M.D. Aaslyng, M. Martens, L. Poll, P.M. Nielsen, H. Flyge, L.M. Larsen, Chemical and Sensory Characterization of Hydrolyzed Vegetable Protein, a Savory Flavoring, Journal of Agricultural and Food Chemistry 46(2) (1998) 481-489.
[31] A. Jarunrattanasri, C. Theerakulkait, K.R. Cadwallader, Aroma Components of Acid-Hydrolyzed Vegetable Protein Made by Partial Hydrolysis of Rice Bran Protein, Journal of Agricultural and Food Chemistry 55(8) (2007) 3044-3050.
[32] A. Tapal, P.K. Tiku, Nutritional and Nutraceutical Improvement by Enzymatic Modification of Food Proteins, in: M. Kuddus (Ed.), Enzymes in Food Biotechnology, Academic Press 2019, pp. 471-481.
[33] R.d.C.L. Lima, R.S. Berg, S.B. Rønning, N.K. Afseth, S.H. Knutsen, D. Staerk, S.G. Wubshet, Peptides from chicken processing by-product inhibit DPP-IV and promote cellular glucose uptake: potential ingredients for T2D management, Food & Function 10(3) (2019) 1619-1628.
[34] J. Wasswa, J. Tang, X.-h. Gu, X.-q. Yuan, Influence of the extent of enzymatic hydrolysis on the functional properties of protein hydrolysate from grass carp (*Ctenopharyngodon idella*) skin, Food Chemistry 104(4) (2007) 1698-1704.
[35] R. Šližytė, R. Mozuraitytė, O. Martínez-Alvarez, E. Falch, M. Fouchereau-Peron, T. Rustad, Functional, bioactive and antioxidative properties of hydrolysates obtained from cod (*Gadus morhua*) backbones, Process Biochemistry 44(6) (2009) 668-677.
[36] J. Newman, T. Egan, N. Harbourne, D. O′Riordan, J.C. Jacquier, M. O′Sullivan, Correlation of sensory bitterness in dairy protein hydrolysates: Comparison of prediction

models built using sensory, chromatographic and electronic tongue data, Talanta 126 (2014) 46-53.
[37] S.G. Wubshet, Industrial production of hydrolysates, 2017.
[38] F.L. Garcia-Carreno, M.A. Navarrete del Toro, Classification of proteases without tears, Biochemical Education 25(3) (1997) 161-167.
[39] M.B. Rao, A.M. Tanksale, M.S. Ghatge, V.V. Deshpande, Molecular and biotechnological aspects of microbial proteases, Microbiology and Molecular Biology Reviews 62(3) (1998) 597-635.
[40] D. Lindberg, A deep dive into enzymatic protein hydrolysis, 2019.
[41] J.J. Perona, C.S. Craik, Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold, Journal of Biological Chemistry 272(48) (1997) 29987-29990.
[42] W. Bode, C. Fernandez-Catalan, H. Tschesche, F. Grams, H. Nagase, K. Maskos, Structural properties of matrix metalloproteinases, Cellular and Molecular Life Sciences 55(4) (1999) 639-652.
[43] Q. Li, L. Yi, P. Marek, B.L. Iverson, Commercial proteases: Present and future, FEBS Letters 587(8) (2013) 1155-1163.
[44] M. Merz, W. Claaßen, D. Appel, P. Berends, S. Rabe, I. Blank, T. Stressler, L. Fischer, Characterization of commercially available peptidases in respect of the production of protein hydrolysates with defined compositions using a three-step methodology, Journal of Molecular Catalysis B: Enzymatic 127 (2016) 1-10.
[45] FDA, Enzyme Preparations Used in Food (Partial List), 2019. https://www.fda.gov/food/generally-recognized-safe-gras/enzyme-preparations-used-food-partial-list. (Accessed 23 July 2019).
[46] EFSA, Food enzymes, 2019. https://www.efsa.europa.eu/en/topics/topic/food-enzymes. (Accessed 23 July 2019).
[47] S.L. Sovik, T. Rustad, Proteolytic Activity in Byproducts from Cod Species Caught at Three Different Fishing Grounds, Journal of Agricultural and Food Chemistry 53(2) (2005) 452-458.
[48] R. Hardy, J.N. Keay, Seasonal variations in the chemical composition of Cornish mackerel, *Scomber scombrus* (L), with detailed reference to the lipids, International Journal of Food Science & Technology 7(2) (1972) 125-137.
[49] M. Díaz-López, F.L. García-Carreño, Applications of fish and shellfish enzymes in food and feed products, in: N.F. Haard, B.K. Simpson (Eds.), Seafood Enzymes, Marcel Dekker 2000, pp. 571-613.
[50] EU, Regulation (EC) No 853/2004 of the European Parliament and of the Council of 29 April 2004: Laying down specific hygiene rules for on the hygiene of foodstuffs, 2004.
[51] EU, Regulation (EC) No 1069/2009 of the European Parliament and of the Council of 21 October 2009: Laying down health rules as regards animal by-products and derived products not intended for human consumption and repealing Regulation (EC) No 1774/2002 (Animal by-products Regulation), 2009.
[52] EU, Commission regulation (EU) No 142/2011 of 25 February 2011: Implementing regulation (EC) No 1069/2009 of the European Parliament and of the Council laying down health rules as regards animal by-products and derived products not intended for human consumption and implementing Council Directive 97/78/EC as regards certain samples and items exempt from veterinary checks at the border under that directive, 2011.
[53] S. Chutipongtanate, K. Watcharatanyatip, T. Homvises, K. Jaturongkakul, V. Thongboonkerd, Systematic comparisons of various spectrophotometric and colorimetric methods to measure concentrations of protein, peptide and amino acid: Detectable limits, linear dynamic ranges, interferences, practicality and unit costs, Talanta 98 (2012) 123-129.

[54] J.C. Moore, J.W. DeVries, M. Lipp, J.C. Griffiths, D.R. Abernethy, Total protein methods and their potential utility to reduce the risk of food protein adulteration, Comprehensive Reviews in Food Science and Food Safety 9(4) (2010) 330-357.
[55] M.P.C. Silvestre, Review of methods for the analysis of protein hydrolysates, Food Chemistry 60(2) (1997) 263-271.
[56] C.T. Srigley, M.M. Mossoba, Current analytical techniques for food lipids, in: U.G. Spizzirri, G. Cirillo (Eds.), Food Safety, Scrivener Publishing LLC. 2016, pp. 33–64.
[57] R.D. Etheridge, G.M. Pesti, E.H. Foster, A comparison of nitrogen values obtained utilizing the Kjeldahl nitrogen and Dumas combustion methodologies (Leco CNS 2000) on samples typical of an animal nutrition analytical laboratory, Animal Feed Science and Technology 73(1) (1998) 21-28.
[58] F. Mariotti, D. Tomé, P.P. Mirand, Converting nitrogen into protein—Beyond 6.25 and Jones' factors, Critical reviews in food science and nutrition 48(2) (2008) 177-184.
[59] A.H. Simonne, E.H. Simonne, R.R. Eitenmiller, H.A. Mills, C.P. Cresman, 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) (1997) 39-45.
[60] International Organization for Standardization (ISO), 2019. https://www.iso.org/home.html. (Accessed 31 August 2019).
[61] J.R. Partington, A history of chemistry, St Martin`s Press 1964.
[62] P. Saez-Plaza, T. Michalowski, M.J. Navas, A.G. Asuero, S. Wybraniec, An overview of the Kjeldahl method of nitrogen determination. Part I. Early history, chemistry of the procedure, and titrimetric finish, Critical Reviews in Analytical Chemistry 43(4) (2013) 178-223.
[63] T. Aspevik, H. Egede-Nissen, Å. Oterhals, A systematic approach to the comparison of cost efficiency of endopeptidases for the hydrolysis of Atlantic salmon (*Salmo salar*) by-products, Food Technology and Biotechnology 54(4) (2016) 421-431.
[64] R.K. Owusu-Apenten, Food protein analysis: quantitative effects on processing. , Marcel Dekker 2002.
[65] E. Chiacchierini, F. D'Ascenzo, D. Restuccia, G. Vinci, Milk soluble whey proteins: Fast and precise determination with Dumas method, Analytical Letters 36(11) (2003) 2473-2484.
[66] E.L. Miller, A.P. Bimbo, S.M. Barlow, B. Sheridan, L.B.W. Burks, Repeatability and Reproducibility of Determination of the Nitrogen Content of Fishmeal by the Combustion (Dumas) Method and Comparison with the Kjeldahl Method: Interlaboratory Study, Journal of AOAC International 90(1) (2007) 6-20.
[67] S.M. Rutherfurd, G.S. Gilani, Amino Acid Analysis, Current Protocols in Protein Science 58(1) (2009) 11.9.1-11.9.37.
[68] W.H. Stein, S. Moore, The free amino acids of human blood plasma, Journal of Biological Chemistry 211(2) (1954) 915-926.
[69] S. Moore, W.H. Stein, Chromatography of amino acids on sulfonated polystyrene resins, Journal of Biological Chemistry 192(2) (1951) 663-681.
[70] M. Fountoulakis, H.-W. Lahm, Hydrolysis and amino acid composition analysis of proteins, Journal of Chromatography A 826(2) (1998) 109-134.
[71] S. Benjakul, M.T. Morrissey, Protein Hydrolysates from Pacific Whiting Solid Wastes, Journal of Agricultural and Food Chemistry 45(9) (1997) 3423-3430.
[72] M. Chalamaiah, B.D. Kumar, R. Hemalatha, T. Jyothirmayi, Fish protein hydrolysates: Proximate composition, amino acid composition, antioxidant activities and applications: A review, Food Chemistry 135(4) (2012) 3020-3038.
[73] A. Grazziotin, F.A. Pimentel, E.V. de Jong, A. Brandelli, Nutritional improvement of feather protein by treatment with microbial keratinase, Animal Feed Science and Technology 126(1-2) (2006) 135-144.

[74] J.O. Onuh, A.T. Girgih, R.E. Aluko, M. Aliani, *In vitro* antioxidant properties of chicken skin enzymatic protein hydrolysates and membrane fractions, Food Chemistry 150 (2014) 366-373.
[75] M.R. Marshall, Ash Analysis, in: S.S. Nielsen (Ed.), Food Analysis, Springer 2010, pp. 107-115.
[76] B. Liaset, E. Lied, M. Espe, Enzymatic hydrolysis of by-products from the fish-filleting industry; chemical characterisation and nutritional evaluation, Journal of the Science of Food and Agriculture 80(5) (2000) 581-589.
[77] K. Kousoulaki, I. Rønnestad, H.J. Olsen, R. Rathore, P. Campbell, S. Nordrum, R.K. Berge, S.A. Mjøs, T. Kalananthan, S. Albrektsen, Krill hydrolysate free amino acids responsible for feed intake stimulation in Atlantic salmon (*Salmo salar*), Aquaculture Nutrition 19(s1) (2013) 47-61.
[78] P.J. Bechtel, M.A. Watson, J.M. Lea, K.L. Bett-Garber, J.M. Bland, Properties of bone from Catfish heads and frames, Food Science & Nutrition 7(4) (2019) 1396-1405.
[79] F. Soxhlet, Soxhlet, über gewichtsanalytische bestimmung des milchfettes, Polytechnisches Journal 232 (1879) 461-465.
[80] E.G. Bligh, W.J. Dyer, A rapid method of total lipid extraction and purification, Canadian Journal of Biochemistry and Physiology 37(8) (1959) 911-917.
[81] M.N. Vaghela, A. Kilara, A rapid method for extraction of total lipids from whey protein concentrates and separation of lipid classes with solid phase extraction, Journal of the American Oil Chemists' Society 72(10) (1995) 1117-1121.
[82] T. Pérez-Palacios, J. Ruiz, D. Martín, E. Muriel, T. Antequera, Comparison of different methods for total lipid quantification in meat and meat products, Food Chemistry 110(4) (2008) 1025-1029.
[83] G. Ewald, G. Bremle, A. Karlsson, Differences between Bligh and Dyer and Soxhlet extractions of PCBs and lipids from fat and lean fish muscle: Implications for data evaluation, Marine Pollution Bulletin 36(3) (1998) 222-230.
[84] P.M. Nielsen, D. Petersen, C. Dambmann, Improved method for determining food protein degree of hydrolysis, J. Food Sci. 66(5) (2001) 642-646.
[85] E. Ficara, A. Rozzi, P. Cortelezzi, Theory of pH-stat titration, Biotechnology and Bioengineering 82(1) (2003) 28-37.
[86] C.F. Jacobsen, J. Leonis, K. Linderstrøm-Lang, M. Ottesen, The pH-stat and its use in biochemistry, Methods of biochemical analysis, Interscience Publishers1957, pp. 171-210.
[87] J. Adler-Nissen, Determination of the degree of hydrolysis of food protein hydrolysates by trinitrobenzenesulfonic acid, Journal of Agricultural and Food Chemistry 27(6) (1979) 1256-1262.
[88] M. Levy, The acidity of formaldehyde and the endpoint in the formol titration, Journal of Biological Chemistry 105(1) (1934) 157-165.
[89] A. Margot, E. Flaschel, A. Renken, Continuous monitoring of enzymatic whey protein hydrolysis. Correlation of base consumption with soluble nitrogen content, Process Biochemistry 29(4) (1994) 257-262.
[90] M.J.M. Hernández, E.B. Domingo, R.M.V. Camañas, M.C.G. Alvarez-Coque, Evaluation of the proteolysis degree with the o-phthalaldehyde/N-acetyl-L-cysteine reagent, Fresenius' Journal of Analytical Chemistry 338(1) (1990) 62-65.
[91] K. Satake, T. Okuyama, M. Ohashi, T. Shinoda, The spectrophotometric determination of amine, amino acid and peptide with 2,4,6-trinitrobenzene 1-sulfonic acid, The Journal of Biochemistry 47(5) (1960) 654-660.
[92] S.M. Rutherfurd, Methodology for determining degree of hydrolysis of proteins in hydrolysates: A review, Journal of AOAC International 93(5) (2010) 1515-1522.

[93] R. Chotikachinda, C. Tantikitti, S. Benjakul, T. Rustad, E. Kumarnsit, Production of protein hydrolysates from skipjack tuna (*Katsuwonus pelamis*) viscera as feeding attractants for Asian seabass (*Lates calcarifer*), Aquaculture Nutrition 19(5) (2013) 773-784.
[94] M. Linder, J. Fanni, M. Parmentier, Functional properties of veal bone hydrolysates, Journal of Food Science 61(4) (1996) 712-716.
[95] M. Opheim, R. Šližytė, H. Sterten, F. Provan, E. Larssen, N.P. Kjos, Hydrolysis of Atlantic salmon (*Salmo salar*) rest raw materials—Effect of raw material and processing on composition, nutritional value, and potential bioactive peptides in the hydrolysates, Process Biochemistry 50(8) (2015) 1247-1257.
[96] Y. Fu, J. Liu, E.T. Hansen, W.L.P. Bredie, R. Lametsch, Structural characteristics of low bitter and high umami protein hydrolysates prepared from bovine muscle and porcine plasma, Food Chemistry 257 (2018) 163-171.
[97] R. Synge, A. Tiselius, Fractionation of hydrolysis products of amylose by electrokinetic ultrafiltration in an agaragar jelly, Biochemical journal 46(5) (1950) R41-R42.
[98] P. Hong, S. Koza, E.S.P. Bouvier, A review size-exclusion chromatography for the analysis of protein biotherapeutics and their aggregates, Journal of Liquid Chromatography & Related Technologies 35(20) (2012) 2923-2950.
[99] H.G. Barth, B.E. Boyes, C. Jackson, Size Exclusion Chromatography and Related Separation Techniques, Analytical Chemistry 70(12) (1998) 251-278.
[100] B.Q. Tran, C. Hernandez, P. Waridel, A. Potts, J. Barblan, F. Lisacek, M. Quadroni, Addressing Trypsin Bias in Large Scale (Phospho)proteome Analysis by Size Exclusion Chromatography and Secondary Digestion of Large Post-Trypsin Peptides, Journal of Proteome Research 10(2) (2011) 800-811.
[101] P.W. Johns, W.A. Jacobs, R.R. Phillips, R.J. McKenna, K.A. O'Kane, J.W. McEwen, Characterisation of peptide molecular mass distribution in commercial hydrolysates and hydrolysate-based nutritional products, Food Chemistry 125(3) (2011) 1041-1050.
[102] J. Qian, Q. Tang, B. Cronin, R. Markovich, A. Rustum, Development of a high performance size exclusion chromatography method to determine the stability of Human Serum Albumin in a lyophilized formulation of Interferon alfa-2b, Journal of Chromatography A 1194(1) (2008) 48-56.
[103] S. Fekete, A. Beck, J.-L. Veuthey, D. Guillarme, Theory and practice of size exclusion chromatography for the analysis of protein aggregates, Journal of Pharmaceutical and Biomedical Analysis 101 (2014) 161-173.
[104] G. Brusotti, E. Calleri, R. Colombo, G. Massolini, F. Rinaldi, C. Temporini, Advances on Size Exclusion Chromatography and Applications on the Analysis of Protein Biopharmaceuticals and Protein Aggregates: A Mini Review, Chromatographia 81(1) (2018) 3-23.
[105] K. Petritis, C. Elfakir, M. Dreux, A comparative study of commercial liquid chromatographic detectors for the analysis of underivatized amino acids, Journal of Chromatography A 961(1) (2002) 9-21.
[106] E.M. Fujinari, J. Damon Manes, Determination of molecular-mass distribution of food-grade protein hydrolyzates by size-exclusion chromatography and chemiluminescent nitrogen detection, Journal of Chromatography A 763(1) (1997) 323-329.
[107] D.S.T. Hsieh, C. Lin, E.R. Lang, N. Catsimpoolas, C.K. Rha, Molecular-weight distribution of soybean globulin peptides produced by peptic hydrolysis, Cereal Chemistry 56(4) (1979) 227-231.
[108] M.P.C. Silvestre, M. Hamon, M. Yvon, Analysis of Protein Hydrolyzates. 1. Use of Poly(2-hydroxyethylaspartamide)-Silica Column in Size Exclusion Chromatography for the Fractionation of Casein Hydrolyzates, Journal of Agricultural and Food Chemistry 42(12) (1994) 2778-2782.

[109] M.P.C. Silvestre, M.R. Silva, V.D.M. Silva, M.W.S. de Souza, C.D. Lopes, W.D. Afonso, Analysis of whey protein hydrolysates: peptide profile and ACE inhibitory activity, Brazilian Journal of Pharmaceutical Sciences 48(4) (2012) 747-757.
[110] J.S. Wang, M.M. Zhao, Q.Z. Zhao, Y. Bao, Y.M. Jiang, Characterization of hydrolysates derived from enzymatic hydrolysis of wheat gluten, J. Food Sci. 72(2) (2007) C103-C107.
[111] H.A. Morais, M.P.C. Silve, V.D.M. Silva, M.R. Silva, S. AnaCristina, J.N. Silveir, Correlation between the Degree of Hydrolysis and the Peptide Profile of Whey Protein Concentrate Hydrolysates: Effect of the Enzyme Type and Reaction Time, American Journal of Food Technology 8 (2013) 1-16.
[112] R.J. Fritsch, I. Krause, Electrophoresis, in: B. Caballero (Ed.), Encyclopedia of Food Sciences and Nutrition (Second Edition), Academic Press2003, pp. 2055-2062.
[113] J.R. Kasper, E.C. Andrews, C. Park, Product inhibition in native-state proteolysis, Plos One 9(10) (2014) 1-6.
[114] K.K. Møller, F.P. Rattray, J.C. Sørensen, Y. Ardö, Comparison of the hydrolysis of bovine κ-casein by camel and bovine chymosin: A kinetic and specificity study, Journal of Agricultural and Food Chemistry 60(21) (2012) 5454-5460.
[115] Q. Zhong, C.R. Daubert, Food Rheology, in: M. Kutz (Ed.), Handbook of Farm, Dairy and Food Machinery Engineering Academic Press 2013, pp. 403-426.
[116] B.P. Lamsal, S. Jung, L.A. Johnson, Rheological properties of soy protein hydrolysates obtained from limited enzymatic hydrolysis, LWT - Food Science and Technology 40(7) (2007) 1215-1223.
[117] M.C. Gómez-Guillén, B. Giménez, M.E. López-Caballero, M.P. Montero, Functional and bioactive properties of collagen and gelatin from alternative sources: A review, Food Hydrocolloids 25(8) (2011) 1813-1827.
[118] D.A. Skoog, J.J. Leary, Principles of Instrumental Analysis, 5 ed., Saunders College Pub. 1992.
[119] D.L. Pavia, G.M. Lampman, G.S. Kriz, J.R. Yyvyan, Introduction to Spectrosopy 5ed., Brooks/Cole, Cengage Lerning 2009.
[120] D.A. Lightner, H.F. Shurvell, J.B. Lambert, R.G. Cooks, Organic Structural Spectroscopy, 1 ed., Prentice-Hall 1998.
[121] A. Barth, Infrared spectroscopy of proteins, Biochimica Et Biophysica Acta-Bioenergetics 1767(9) (2007) 1073-1101.
[122] L. Nobili, Description d'un thermo-multiplicateur ou thermoscope électrique, Bibliothèque Universelle des Sciences, Belles-Lettres et Arts 2 (1830) 225-234.
[123] M. Vollmer, K.-P. Möllmann, Infrared Thermal Imaging: Fundamentals, Reserch and Applications 2ed., Wiley & Sons 2017.
[124] Y.M. Rabkin, Technological innovation in science - The adoption of infrared-spectroscopy by chemists, Isis 78(291) (1987) 31-54.
[125] H. Gershinowitz, The first infrared spectrometer, The Journal of Physical Chemistry 83(11) (1979) 1363-1365.
[126] T.J. Bruno, Basic Instrumental Techniques of Analytical Chemistry, in: W.M. Haynes (Ed.), CRC Handbook of Chemistry and Physics CRC Press2016.
[127] O. Abbas, N.K. Afseth, S. Armenta, M.J. Ayora-Canada, H. Azizian, V. Baeten, M. Baranska, A.S. Baarros, O. Bozkurt, L.R. Brewer, R. Burling-Clarige, M. Careche, P. Cormona, J.M. Chalmers, S.-M. Choi, M.A. Coimbra, D. Cozzolion, R.G. Dambergs, P. Dardenne, E. Da Silva, F. Dehareng, I. Delgadillo, A. Domfnguez-Vidal, A. Enfield, C.C. Fagan, S. Farquharson, S. Garrigues, A. Ghetler, M. Gishen, P.R. Griffiths, M. de la Guardia, T.M. Hancewicz, A.M. Herrero, D.S. Himmelsbach, F. Inscore, A.A. Ismail, F. Jimenez-Colmenero, K. Jørgensen, A. Kohler, J.K.G. Kramer, J.M. Lagaron, J. Lammertyn, B. Lendl,

E.C.Y. Li-Chan, M. Lin, A. Lopez-Rubio, X. Lu, C.-Y. Ma, H. Martens, J. Moss, M.M. Mossoba, M.P. Nelson, B. Nicolai, A. Nunes, C.P. O'Donnell, E. Okazaki, D.L. Phillips, J.A. Fernandez Pierna, E. Polshin, J. Popp, R.J. Priore, P.D.A. Pudney, Å. Randby, B. Rasco, L.E. Rodriguez-Saona, D. Rousseau, D.E. Rubio-Diaz, C. Ruiz-Capillas, I. Sanchez-Alonso, H. Schulz, J. Sedman, A. Sengupta, F. Severcan, C. Shende, P. Vermeulen, M. Strehle, D.L. Wetzel, P.J. Treado, M. Uddin, P. Williams, S.-N. Yuen, G. Sinnaeve, Applications of vibratonal spectroscopy in food science, 1 ed., John Wiley & Sons 2010.
[128] Z. Movasaghi, S. Rehman, I.U. Rehman, Fourier transform infrared (FTIR) spectroscopy of biological tissues, Applied Spectroscopy Reviews 43(2) (2008) 134-179.
[129] D.M. Byler, H. Susi, Examination of the secondary structure of proteins by deconvolved FTIR spectra, Biopolymers 25(3) (1986) 469-487.
[130] U. Böcker, R. Ofstad, H.C. Bertram, B. Egelandsdal, A. Kohler, Salt-induced changes in pork myofibrillar tissue investigated by FT-IR microspectroscopy and light microscopy, Journal of Agricultural and Food Chemistry 54(18) (2006) 6733-6740.
[131] P. Gelfand, R.J. Smith, E. Stavitski, D.R. Borchelt, L.M. Miller, Characterization of Protein Structural Changes in Living Cells Using Time-Lapsed FTIR Imaging, Analytical Chemistry 87(12) (2015) 6025-6031.
[132] S. Krimm, J. Bandekar, Vibrational Spectroscopy and Conformation of Peptides, Polypeptides, and Proteins, in: C.B. Anfinsen, J.T. Edsall, F.M. Richards (Eds.), Advances in Protein Chemistry, Academic Press 1986, pp. 181-364.
[133] K.A. Oberg, J.-M. Ruysschaert, E. Goormaghtigh, The optimization of protein secondary structure determination with infrared and circular dichroism spectra, European Journal of Biochemistry 271(14) (2004) 2937-2948.
[134] A. Barth, The infrared absorption of amino acid side chains, Progress in Biophysics and Molecular Biology 74(3-5) (2000) 141-173.
[135] N. Perisic, N.K. Afseth, R. Ofstad, A. Kohler, Monitoring Protein Structural Changes and Hydration in Bovine Meat Tissue Due to Salt Substitutes by Fourier Transform Infrared (FTIR) Microspectroscopy, Journal of Agricultural and Food Chemistry 59(18) (2011) 10052-10061.
[136] P.V. Andersen, E. Veiseth-Kent, J.P. Wold, Analyzing pH-induced changes in a myofibril model system with vibrational and fluorescence spectroscopy, Meat Science 125 (2017) 1-9.
[137] S.H. Arabi, B. Aghelnejad, C. Schwieger, A. Meister, A. Kerth, D. Hinderberger, Serum albumin hydrogels in broad pH and temperature ranges: characterization of their self-assembled structures and nanoscopic and macroscopic properties, Biomaterials science 6(3) (2018) 478-492.
[138] G. Martra, C. Deiana, Y. Sakhno, I. Barberis, M. Fabbiani, M. Pazzi, M. Vincenti, The Formation and Self-Assembly of Long Prebiotic Oligomers Produced by the Condensation of Unactivated Amino Acids on Oxide Surfaces, Angewandte Chemie International Edition 53(18) (2014) 4671-4674.
[139] H.-F. Okabayashi, H.-H. Kanbe, C.J. O'Connor, The role of an L-leucine residue on the conformations of glycyl-L-leucine oligomers and its N- or C-terminal dependence: infrared absorption and Raman scattering studies, European Biophysics Journal 45(1) (2016) 23-34.
[140] C. Ruckebusch, L. Duponchel, J.P. Huvenne, P. Legrand, N. Nedjar-Arroume, B. Lignot, P. Dhulster, D. Guillochon, Hydrolysis of hemoglobin surveyed by infrared spectroscopy II. Progress predicted by chemometrics, Analytica Chimica Acta 396(2-3) (1999) 241-251.
[141] C. Ruckebusch, L. Duponchel, J.P. Huvenne, Degree of hydrolysis from mid-infrared spectra, Analytica Chimica Acta 446(1-2) (2001) 257-268.

[142] C. Ruckebusch, B. Sombret, R. Froidevaux, J.P. Huvenne, On-line mid-infrared spectroscopic data and chemometrics for the monitoring of an enzymatic hydrolysis, Applied Spectroscopy 55(12) (2001) 1610-1617.
[143] G. Guler, E. Dzafic, M.M. Vorob'ev, V. Vogel, W. Mantele, Real time observation of proteolysis with Fourier transform infrared (FT-IR) and UV-circular dichroism spectroscopy: Watching a protease eat a protein, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 79(1) (2011) 104-111.
[144] G. Guler, M.M. Vorob'ev, V. Vogel, W. Mantele, Proteolytically-induced changes of secondary structural protein conformation of bovine serum albumin monitored by Fourier transform infrared (FT-IR) and UV-circular dichroism spectroscopy, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 161 (2016) 8-18.
[145] N.A. Poulsen, C.E. Eskildsen, M. Akkerman, L.B. Johansen, M.S. Hansen, P.W. Hansen, T. Skov, L.B. Larsen, Predicting hydrolysis of whey protein by mid-infrared spectroscopy, International Dairy Journal 61 (2016) 44-50.
[146] H. Ayvaz, R. Temizkan, Quick vacuum drying of liquid samples prior to ATR-FTIR spectral collection improves the quantitative prediction: a case study of milk adulteration, International Journal of Food Science & Technology 53(11) (2018) 2482-2489.
[147] R. Bro, A.K. Smilde, Principal component analysis, Analytical Methods 6(9) (2014) 2812-2831.
[148] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometrics and Intelligent Laboratory Systems 58(2) (2001) 109-130.
[149] N.K. Afseth, A. Kohler, Extended multiplicative signal correction in vibrational spectroscopy, a tutorial, Chemometrics and Intelligent Laboratory Systems 117 (2012) 92-99.
[150] J.L. Ilari, H. Martens, T. Isaksson, Determination of particle-size in powders by scatter correction in diffuse near-infrared reflectance, Applied Spectroscopy 42(5) (1988) 722-728.
[151] P. Geladi, D. Macdougall, H. Martens, Linearization and scatter-correction for near-infrared reflectance spectra of meat, Applied Spectroscopy 39(3) (1985) 491-500.
[152] K. Pearson, LIII. On lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2(11) (1901) 559-572.
[153] K.G. Dunn, Process Improvement Using Data, 2019. https://learnche.org/pid/contents. (Accessed August 18 2019).
[154] K. Tøndel, U.G. Indahl, A.B. Gjuvsland, J.O. Vik, P. Hunter, S.W. Omholt, H. Martens, Hierarchical Cluster-based Partial Least Squares Regression (HC-PLSR) is an efficient tool for metamodelling of nonlinear dynamic models, BMC Systems Biology 5:90(1) (2011).
[155] K.H. Liland, A. Kohler, V. Shapaval, Hot PLS—a framework for hierarchically ordered taxonomic classification by partial least squares, Chemometrics and Intelligent Laboratory Systems 138 (2014) 41-47.
[156] D. Perez-Guaita, J. Kuligowski, G. Quintás, S. Garrigues, M.d.l. Guardia, Modified locally weighted—Partial least squares regression improving clinical predictions from infrared spectra of human serum samples, Talanta 107 (2013) 368-375.
[157] O. Preisner, R. Guiomar, J. Machado, J.C. Menezes, J.A. Lopes, Application of Fourier Transform Infrared Spectroscopy and Chemometrics for Differentiation of Salmonella enterica Serovar Enteritidis Phage Types, Applied and Environmental Microbiology 76(11) (2010) 3538-3544.
[158] T. Næs, T. Isaksson, T. Fearn, T. Davies, A User-Friendly Guide to Multivariate Calibration and Classification, 1 ed., NIR Publications 2002.
[159] S. Raghavan, H.G. Kristinsson, ACE-inhibitory activity of tilapia protein hydrolysates, Food Chemistry 117(4) (2009) 582-588.

[160] M.A. Pfeffer, Angiotensin-converting enzyme inhibition in congestive heart failure: Benefit and perspective, American Heart Journal 126(3, Part 2) (1993) 789-793.
[161] K.A. Kristoffersen, K.H. Liland, U. Böcker, S.G. Wubshet, D. Lindberg, S.J. Horn, N.K. Afseth, FTIR-based hierarchical modeling for prediction of average molecular weights of protein hydrolysates, Talanta 205 (2019) 120084.
[162] R. Rajkhowa, B. Levin, S.L. Redmond, L.H. Li, L.J. Wang, J.R. Kanwar, M.D. Atlas, X.G. Wang, Structure and properties of biomedical films prepared from aqueous and acidic silk fibroin solutions, Journal of Biomedical Materials Research Part A 97A(1) (2011) 37-45.
[163] L.E. Valenti, M.B. Paci, C.P. De Pauli, C.E. Giacomelli, Infrared study of trifluoroacetic acid unpurified synthetic peptides in aqueous solution: Trifluoroacetic acid removal and band assignment, Analytical Biochemistry 410(1) (2011) 118-123.
[164] A. Ahmad, K.P. Madhusudanan, V. Bhakuni, Trichloroacetic acid and trifluoroacetic acid-induced unfolding of cytochrome c: stabilization of a native-like folded intermediate11CDRI communication number 5877, Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 1480(1-2) (2000) 201-210.
[165] A.S.N. Murthy, K.G. Rao, C.N.R. Rao, Molecular orbital study of the configuration protonation, and hydrogen bonding of secondary amides, Journal of the American Chemical Society 92(12) (1970) 3544-3548.
[166] J.L. Sudmeier, K.E. Schwartz, Protonation of carbonyl groups in peptides and amino-acids, Chemical Communications (London) (24) (1968) 1646-1648.
[167] I.M. Klotz, S.F. Russo, S. Hanlon, M.A. Stake, Protonation of Amides in a Helix-Breaking Solvent, Journal of the American Chemical Society 86(22) (1964) 4774-4778.
[168] J.W.O. Tam, I.M. Klotz, Protonation of amides by trifluoroacetic acid: infrared and nuclear magnetic resonance studies, Spectrochimica Acta Part A: Molecular Spectroscopy 29(4) (1973) 633-644.
[169] D. Spellman, E. McEvoy, G. O'Cuinn, R.J. FitzGerald, 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) (2003) 447-453.
[170] W. Saeys, N. Nguyen Do Trong, R. Van Beers, B.M. Nicolaï, Multivariate calibration of spectroscopic sensors for postharvest quality evaluation: A review, Postharvest Biology and Technology 158 (2019) 1-19.
[171] T. Naes, T. Isaksson, B. Kowalski, Locally weighted regression and scatter correction for near-infrared reflectance data, Analytical Chemistry 62(7) (1990) 664-673.
[172] R. Wu, T.B. McMahon, Protonation Sites and Conformations of Peptides of Glycine ($Gly_{1-5}H^+$) by IRMPD Spectroscopy, The Journal of Physical Chemistry B 113(25) (2009) 8767-8775.
[173] M. Boström, D.R.M. Williams, B.W. Ninham, Surface tension of electrolytes: Specific ion effects explained by dispersion forces, Langmuir 17(15) (2001) 4475-4478.
[174] Y.J. Zhang, P.S. Cremer, Interactions between macromolecules and ions: the Hofmeister series, Current Opinion in Chemical Biology 10(6) (2006) 658-663.
[175] L.H. Mujawar, A. van Amerongen, W. Norde, Influence of buffer composition on the distribution of inkjet printed protein molecules and the resulting spot morphology, Talanta 98 (2012) 1-6.
[176] L.H. Mujawar, A. van Amerongen, W. Norde, Influence of Pluronic F127 on the distribution and functionality of inkjet-printed biomolecules in porous nitrocellulose substrates, Talanta 131 (2015) 541-547.
[177] L.H. Mujawar, W. Norde, A. van Amerongen, Spot morphology of non-contact printed protein molecules on non-porous substrates with a range of hydrophobicities, Analyst 138(2) (2013) 518-524.

[178] L.H. Mujawar, J.G.M. Kuerten, D.P. Siregar, A. van Amerongen, W. Norde, Influence of the relative humidity on the morphology of inkjet printed spots of IgG on a non-porous substrate, RSC Advances 4(37) (2014) 19380-19388.
[179] R.G. Dadzie, H. Ma, E.E. Abano, W. Qu, S. Mao, Optimization of process conditions for production of angiotensin I-converting enzyme (ACE) inhibitory peptides from vital wheat gluten using response surface methodology, Food Science and Biotechnology 22(6) (2013) 1531-1537.
[180] S. Wu, J. Sun, Z. Tong, X. Lan, Z. Zhao, D. Liao, Optimization of hydrolysis conditions for the production of angiotensin-I converting enzyme-inhibitory peptides and isolation of a novel peptide from lizard fish (Saurida elongata) muscle protein hydrolysate, Mar Drugs 10(5) (2012) 1066-1080.
[181] G. ElMasry, J.P. Wold, High-speed assessment of fat and water content distribution in fish fillets using online imaging spectroscopy, Journal of Agricultural and Food Chemistry 56(17) (2008) 7672-7677.
[182] J.P. Wold, M. Kermit, A. Woll, Rapid nondestructive determination of edible meat content in crabs (*Cancer Pagurus*) by near-infrared imaging spectroscopy, Applied Spectroscopy 64(7) (2010) 691-699.
[183] J.P. Wold, M. O'Farrell, M. Høy, J. Tschudi, On-line determination and control of fat content in batches of beef trimmings by NIR imaging spectroscopy, Meat Science 89(3) (2011) 317-324.
[184] J.P. Wold, E. Veiseth-Kent, V. Høst, A. Løvland, Rapid on-line detection and grading of wooden breast myopathy in chicken fillets by near-infrared spectroscopy, Plos One 12(3) (2017) 1-16.
[185] I. Måge, T. Næs, Optimising production cost and end-product quality when raw material quality is varying, Journal of Chemometrics 21(10-11) (2007) 440-450.
[186] I. Måge, B.-H. Mevik, T. Næs, Regression models with process variables and parallel blocks of raw material measurements, Journal of Chemometrics 22(8) (2008) 443-456.
[187] J. Ollesch, S.L. Drees, H.M. Heise, T. Behrens, T. Bruning, K. Gerwert, FTIR spectroscopy of biofluids revisited: an automated approach to spectral biomarker identification, Analyst 138(14) (2013) 4092-4102.
[188] Y. Xiong, V. Shapaval, A. Kohler, P.J. From, A Laboratory-Built Fully Automated Ultrasonication Robot for Filamentous Fungi Homogenization, Slas Technology: Translating Life Sciences Innovation (2019) 1-13.
[189] M. Sieger, G. Kos, M. Sulyok, M. Godejohann, R. Krska, B. Mizaikoff, Portable Infrared Laser Spectroscopy for On-site Mycotoxin Analysis, Scientific Reports 7 (2017) 1-6.

# Paper I

Talanta 205 (2019) 120084

Contents lists available at ScienceDirect

Talanta

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

# FTIR-based hierarchical modeling for prediction of average molecular weights of protein hydrolysates

Kenneth Aase Kristoffersen[a,b,*], Kristian Hovde Liland[c], Ulrike Böcker[a], Sileshi Gizachew Wubshet[a], Diana Lindberg[a], Svein Jarle Horn[b], Nils Kristian Afseth[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] Faculty of Science and Technology, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, N-1432, Ås, Norway

ARTICLE INFO

*Keywords:*
Enzymatic protein hydrolysis
FTIR
Data analysis
Hierarchical modeling

ABSTRACT

In the presented study, Fourier-transform infrared (FTIR) spectroscopy is used to predict the average molecular weight of protein hydrolysates produced from protein-rich by-products from food industry using commercial enzymes. Enzymatic protein hydrolysis is a well-established method for production of protein-rich formulations, recognized for its potential to valorize food-processing by-products. The monitoring of such processes is still a significant challenge as the existing classical analytical methods are not easily applicable to industrial setups. In this study, we are reporting a generic FTIR-based approach for monitoring the average molecular weights of proteins during enzymatic hydrolysis of by-products from the food industry. A total of 885 hydrolysate samples from enzymatic protein hydrolysis reactions of poultry and fish by-products using different enzymes were studied. FTIR spectra acquired from dry-films of the hydrolysates were used to build partial least squares regression (PLSR) models. The most accurate predictions were obtained using a hierarchical PLSR approach involving supervised classification of the FTIR spectra according to raw material quality and enzyme used in the hydrolysis process, and subsequent local regression models tuned to specific enzyme-raw material combinations. The results clearly underline the potential of using FTIR for monitoring protein sizes during enzymatic protein hydrolysis in industrial settings, while also paving the way for measurements of protein sizes in other applications.

## 1. Introduction

FTIR spectroscopy has become an established method for protein and peptide structural characterization over the last few decades. This is due to the detailed structural information found in FTIR spectra, where the repeated amino acid building blocks of proteins and peptides give rise to nine distinctive infrared (IR) absorption bands (i.e., the amide bands) [1,2]. The inherent ability of FTIR spectroscopy to monitor the protein backbone can also provide a range of possibilities to study parameters related to protein secondary structures. These parameters include hydration and solvent effects, pH and peptide size [3–8]. Protein size estimations can potentially have practical applications in a range of different fields, one of them being enzymatic protein hydrolysis. This process represents an efficient and suitable method to extract protein from food processing residuals, involving the breakdown of proteins into peptides and free amino acids. The success of a hydrolysis process can be measured by its ability to produce maximum possible yield of a high-quality product within the required specifications and with limited or no batch-to-batch variation. Currently, there is a lack of fast and reliable analytical monitoring tools that can be used to achieve such process control.

In enzymatic protein hydrolysis reactions, a proteolytic enzyme catalyzes the hydrolysis of peptide bonds. The reaction results in the formation of C-terminals ($COO^-$) and N-terminals ($NH_3^+$), consequently changing both the primary and secondary structure of the protein or peptide. Several studies on pure model proteins like hemoglobin, β-lactoglobulin, β-casein, and bovine serum albumin have demonstrated that FTIR spectroscopy can be used to monitor proteolytic reactions. [9–13] This was extended to more complex protein-rich matrices when FTIR was employed as a tool to predict degree of hydrolysis (DH%) values for trypsin-catalyzed hydrolysis of whey proteins [14]. Recently, the applicability of monitoring proteolysis using FTIR was further expanded to salmon and poultry-based substrates [15]. Wubshet et al. reported an FTIR-based multivariate approach for monitoring the change in weight average molecular weight ($M_w$) during enzymatic hydrolysis of chicken by-products [16]. In that study, $M_w$ was

* Corresponding author. Nofima - Norwegian Institute of Food, Fisheries and Aquaculture Research, P.O. Box 210, N-1431, Ås, Norway.
*E-mail address:* Kenneth.kristoffersen@nofima.no (K.A. Kristoffersen).

https://doi.org/10.1016/j.talanta.2019.06.084
Received 26 January 2019; Received in revised form 19 June 2019; Accepted 21 June 2019
Available online 22 June 2019

calculated from size exclusion chromatography (SEC), establishing it as a reference method for FTIR spectra calibration models and as a measure of the extent of protein hydrolysis.

Several studies have revealed that amide absorptions (i.e., amide I at ~1650 $cm^{-1}$), $NH_3^+$ deformation (1516 $cm^{-1}$), and $COO^-$ stretching (1400 $cm^{-1}$) are important for prediction of $M_w$ or DH% [14–16]. From these studies, it is also apparent that protein hydrolysates originating from different raw materials and even different enzymes will display different FTIR fingerprints. However, when the aim is to establish a generic prediction model for $M_w$, such spectral differences may be a challenge. This was illustrated in a recent study by Wubshet et al. [16]. Here, the coefficients of determination between FTIR spectra and $M_w$ were reduced when samples of different raw material origins were combined in a single model, compared to separate modeling according to raw material origin. Different spectral signatures of the raw materials were designated as the main reason for this. Removing these types of spectral variations through mathematical pre-processing is hard and, in most cases, not possible. An alternative approach will therefore be to exploit these spectral differences. This can be achieved through a two-level model, where the spectra are assigned to predefined groups consisting of known enzyme-raw material combinations in the first level using a supervised classification model. On the second level the spectra can be subjected to a local regression model tuned to specific enzyme-raw material combinations. Hierarchical modeling through two-level strategies has previously been exploited in methods such as hierarchical cluster-based partial least squares regression (HC-PLSR) and hierarchically ordered taxonomic classification by partial least squares (Hot PLS) in applications including nonlinear dynamic models and taxonomic classification, respectively [17,18]. Successful applications of FTIR-based two-level partial least squares (PLS) modeling have been demonstrated in the determination of clinical parameters such as urea and glucose as well as complex protein structures [19,20].

In the present study, the objective was to establish and study the relationship between $M_w$ and FTIR spectra in an extensive set of protein hydrolysates. Hence, a total of 885 hydrolysates from enzymatic protein hydrolysis of poultry and fish by-products using five different enzymes were studied. The two-level regression model, tuned to various combinations of raw material and enzymes, was compared to the standard PLSR model in order to demonstrate the power of a two-level regression approach. To the best of our knowledge, this is the first time a link between $M_w$ of proteins and FTIR spectra has been studied and established for an extensive set of samples.

## 2. Materials and methods

### 2.1. Materials

Protease from *Bacillus licheniformis* (Alcalase, 2.4 U/g) was purchased from Sigma-Aldrich (St. Louis, MO, USA). Protamex and Flavourzyme was obtained from Novozymes (Bagsværd, Denmark), Papain LSG 100 from Enzybel (Waterloo, Belgium) and Corolase 2TS from AB Enzymes (Darmstadt, Germany). Analytical grade acetonitrile, trifluoroacetic acid, monosodium phosphate and molecular weight standards, i.e., bovine serum albumin, albumin from chicken egg white, carbonic anhydrase from bovine erythrocytes, lysozyme, cytochrome c from bovine heart, aprotinin from bovine lung, insulin chain B oxidized from bovine pancreas, angiotensin II human, bradykinin fragment 1-7, Val-Tyr-Val, and tryptophan were purchased from Sigma-Aldrich (St. Louis, MO, USA). Water used for HPLC was purified by deionization and 0.22 µm membrane filtration (MilliporeSigma, Burlington, MA, USA).

### 2.2. Raw materials

Protein-rich raw materials derived from chicken, turkey, salmon and mackerel were hydrolyzed by a selection of commercially available enzymes (see Table 1). The poultry raw materials (i.e., chicken mechanical deboning residue (CMDR), heat treated chicken mechanical deboning residue (hCMDR), chicken skin (CS), chicken bone (CB), turkey carcasses (TC) and turkey mechanical deboning residue (TMDR)) were supplied by a Norwegian slaughterhouse (Nortura, Hærland, Norway). Chicken fillets/muscle (CM) were purchased from a local grocery store in Ås, Norway. Salmon raw materials (i.e., heads (SH), bone (SB) and skin (SS)) were supplied by Nutrimar (Kverva, Norway). Mackerel raw materials (Ma) were supplied by Pelagia Tromsø (Tromsdalen, Norway). All samples were minced, packed in plastic bags and stored at −20 °C until further use.

**Table 1**
An overview of samples and hydrolysis reaction conditions.

| Sample name[a] | Enzyme[b] | Enzyme loading (w/w)%[c] | Water (mL)[d] | Raw material (g)[e] | No. of samples[f] |
|---|---|---|---|---|---|
| CMDRA | Alcalase | 1.5 | 1000 | 500 | 89 |
| CMDRPa | Papain | 1.5 | 1000 | 500 | 24 |
| CMDRPr | Protamex | 1.5 | 1000 | 500 | 36 |
| hCMDRA | Alcalase | 1.5 | 1000 | 500 | 23 |
| hCMDRPa | Papain | 1.5 | 1000 | 500 | 24 |
| hCMDRPr | Protamex | 1.5 | 1000 | 500 | 12 |
| CMA | Alcalase | 1.5 | 1000 | 500 | 87 |
| CMPa | Papain | 1.5 | 1000 | 500 | 23 |
| CMPr | Protamex | 1.5 | 1000 | 500 | 12 |
| CSA | Alcalase | 1.5 | 1000 | 500 | 22 |
| CSPa | Papain | 1.5 | 1000 | 500 | 24 |
| CSPr | Protamex | 1.5 | 1000 | 500 | 12 |
| CBA | Alcalase | 1.5 | 1000 | 500 | 12 |
| CBPa | Papain | 1.5 | 1000 | 500 | 12 |
| CBPr | Protamex | 1.5 | 1000 | 500 | 12 |
| TCA | Alcalase | 1 | 1000 | 500 | 22 |
| TCC | Corolase 2TS | 1 | 1000 | 500 | 24 |
| TCF | Flavourzyme | 1 | 1000 | 500 | 23 |
| TMDRA | Alcalase | 1 | 1000 | 500 | 24 |
| TMDRC | Corolase 2TS | 1 | 1000 | 500 | 24 |
| TMDRF | Flavourzyme | 1 | 1000 | 500 | 24 |
| SHA | Alcalase | 1 | 400 | 400 | 130 |
| SSA | Alcalase | 1 | 400 | 400 | 132 |
| SBA | Alcalase | 1 | 400 | 400 | 11 |
| Ma | None | 0 | 400 | 400 | 12 |
| MaA | Alcalase | 1 | 400 | 400 | 12 |
| MaPa | Papain | 1 | 400 | 400 | 11 |
| MaF | Flavourzyme | 1 | 400 | 400 | 12 |
| All | | | | | 885 |

[a] Raw materials are defined in chapter 2.2. The abbreviation for the enzymes used are added to the sample name.
[b] Alcalase, 2.4 U/g (A), Protamex (Pr), Flavourzyme (F), Papain LSG 100 (Pa) and Corolase 2TS (C).
[c] Enzyme loading relative to wet weight raw material.
[d] Water added to reaction mixture.
[e] Raw material loading.
[f] Number of samples in each enzyme-raw material group.

### 2.3. Enzymatic hydrolysis and sampling

All hydrolysis reactions were performed according to a previously published protocol using a Reactor-Ready™ jacketed reaction vessel (Radleys, Essex, United Kingdom) [16]. Water circulating through the vessel jacket was kept at 50 °C and was supplied using a JULABO circulator pump (JULABO GmbH, Seelbach, Germany). Raw materials and water were mixed in the ratios presented in Table 1. All reaction mixtures were thoroughly mixed and heated until the suspensions reached 50 ± 1 °C. This was followed by addition of 1-1.5% enzyme w/w to wet substrate weight. The reaction time, from addition of the enzyme, were 60 or 80 min. During the hydrolysis, aliquots of approximately 7 mL were collected at 11 or 12 time points (0.5, 2.5, 5, 7.5, 10, 15, 20, 30, 40, 50, 60 and 80 min, respectively). Many of the reactions were repeated multiple times, as seen in Table 1. After collecting the samples

from the reaction vessel, the enzyme was thermally inactivated before being allowed to cool to room temperature. The samples were then centrifuged to separate the mixtures into three phases: Solid, water and fat. The water phase was collected, and analytical samples of protein hydrolysate were prepared by filtration through Millex-HV PVDF 0.45 µm 33 mm filter (MilliporeSigma, Burlington, MA, USA).

### 2.4. Size exclusion chromatography

SEC analysis was preformed according to Wubshet et al. [16]. The 2 mg/mL solutions of standards and the filtrates of the water phases collected from the hydrolysis were directly used as injection solutions without further modifications. The injection volumes were of 7 µL for the fish samples and 10 µL for the standards and poultry samples. Chromatographic separations of standards and samples were performed with an Agilent 1200 series instrument (Agilent Technologies, Santa Clara, CA, USA). Separation was performed at 25 °C using BioSep-SEC-s2000 (300 × 7.8 mm) columns from Phenomenex (Torrence, CA, USA). The mobile phase consisted of a mixture of acetonitrile and ultrapure water in a proportion 30:70 (v/v), containing 0.05% trifluoroacetic acid. Isocratic elution was carried out using a flow rate of 0.9 mL/min for 17.0 min. Between 17.0 and 17.1 min the mobile phase was changed to $NaH_2PO_4$ (0.10 M) and maintained until 20.0 min for column cleaning. Elution conditions were restored between minute 20.0 and 20.1 and the column was equilibrated for an additional 25 min. Chromatographic runs were controlled from OpenLAB CDS Rev. C. 01.07 (Agilent Technologies, Santa Clara, CA, USA). From chromatographic runs of both the standards and hydrolysate samples presented in Table 1, a UV trace of 214 nm was used. For the analytical standards, retention times were obtained from the automatic peak-picking algorithm of OpenLAB CDS. The average retention times from triplicate measurements of the standards were used to construct the calibration curves. Calibration data for one of the columns used including average retention times and standard deviation for all the analytical standers, are presented in the supporting information (SI) in Table S–1. Finally, $M_w$ was calculated from the UV trace of a single chromatographic run for each of the hydrolysate samples. Calculations of the $M_w$ were performed using PSS winGPC UniChrom V 8.00 (Polymer Standards Service, Mainz, Germany). The calculation from the software was based on a slicing method, similar to those previously used for analysis of protein hydrolysates [21].

### 2.5. FTIR spectroscopy

From each of the filtered protein hydrolysates, aliquots (5-7.5 µL) were deposited on 96-well IR-transparent Si-plates (Bruker, Billerica, MA, USA) and dried at room temperature for at least 30 min to form dry-films as described by Böcker et al. [15]. 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).

### 2.6. Data analysis

Pre-processing of FTIR spectra was performed using Savitzky-Golay 2nd derivative smoothing (window width 11 pt, 3rd order polynomial smoothing) followed by extended multiplicative signal correction (EMSC) with 2nd order polynomial correction with the mean spectrum as reference. This pre-processing approach was used to reduce the scattering effects in the spectra and suppress the effect of the varying thickness of the dry-films [22]. For all subsequent data analysis, the region from 1800 $cm^{-1}$ to 700 $cm^{-1}$ was used. The pre-processed FTIR spectra were further used for prediction of $M_w$. Two PLSR approaches were applied and compared: (1) Standard PLSR (henceforth denoted as the one-level PLSR model) and (2) a two-level modeling approach. In the first approach, PLSR was applied to the pre-processed FTIR spectra from all subgroups of enzymes and raw materials and the corresponding $M_w$ values obtained from size exclusion chromatography. All the 885 FTIR spectra and subgroup datasets according to raw material origin and enzymes used (see Table 1) were modeled together. In the two-level modeling approach, spectra were first classified into 28 subgroups (as defined in Table 1) and then subjected to regression models tuned to each raw material and enzyme combination. Both the classification and regression models used PLSR for dimension reduction. In the classification a variant called canonical partial least squares (CPLS) was chosen for its ability to give simpler models and the response was a dummy coded representation of the subgroup [23]. PLS finds small subspaces in the high-dimensional FTIR data that span most of the covariation between spectra and response. For classification, the scores from the CPLS model were subjected to linear discriminant analysis (LDA) to obtain class memberships, while ordinary least squares regression was used on the scores when predicting $M_w$. All the modeling, classification and prediction were wrapped in a leave-one-out cross-validation to make sure that none of the spectra interfered with any of the models that it was to be predicted from, thus eliminating sources of information bleeding. All statistical analyses were performed using the MATLAB software (R2018a, The MathWorks, Inc., Natick, MA, USA).

## 3. Results and discussion

### 3.1. Average molecular weight and FTIR profiling

A total of 885 hydrolysate samples were prepared in the current study, collected at different time points during enzymatic protein hydrolysis of a variety of raw materials and enzymes (see Table 1). All samples were analyzed by SEC and corresponding $M_w$ were calculated. For a given protein hydrolysate, $M_w$ is a highly descriptive molecular weight distribution parameter that can serve as a measure of the extent of hydrolysis. The relationship between hydrolysis time and $M_w$ for two

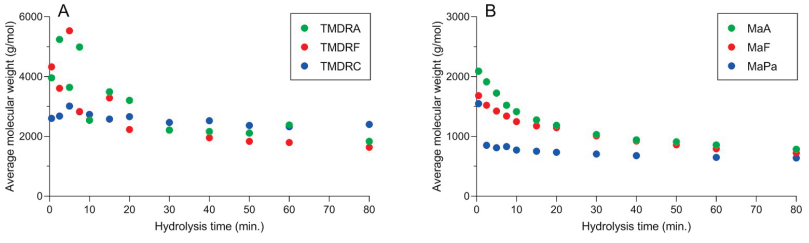



**Fig. 1.** $M_w$ plotted against the hydrolysis time for protein hydrolysates obtained using four enzymes and two different raw materials from A) Turkey and B) Mackerel. Raw materials and enzyme abbreviations are defined in chapter 2.2. and Table 1.

raw materials (turkey and mackerel) is seen in Fig. 1. The maximum, minimum and average $M_w$ for all raw material and enzyme combinations are displayed in SI, Table S–2.

A large variation in $M_w$ between samples prepared from different raw materials and enzymes was observed. The general trend was a fast reduction in $M_w$ at the start of the process followed by a slower change with increasing hydrolysis time, as shown in Fig. 1. Similar observations have been reported previously for hydrolysis reactions using different complex substrates and enzymes [16,21]. For many of the reactions, significant fluctuations were also observed in the first 20-30 min as shown in Fig. 1. It is reasonable to explain this observation by the substantial heterogeneity of the reaction mixtures, as raw materials were only ground to a semi-homogenous mass to simulate industry-relevant pretreatment before hydrolysis [24].

The turkey hydrolysates generally contained larger amounts of longer peptides at the beginning of the process, as compared to almost all of the other raw materials. The difference could be explained by differences in amount of structural proteins between the two poultry species. Turkey contains more collagenous compounds than chicken, and all proteases used have relatively higher specificity for peptide bonds prevalent within myofibrillar proteins. Another possible explanation is the formation of relatively higher amounts of "virtual intermediate peptides" after liberation of peptides within the collagen-rich turkey hydrolysates. Virtual intermediate peptides are aggregates of cleaved peptides, limiting the accessibility of the specific peptide bonds for peptide hydrolysis [25].

Depending on both the raw material quality and the enzyme used, enzymatic liberation of peptides and free amino acids during an enzymatic hydrolysis process of complex protein-rich substrates will result in large variations in the water phase composition. This is clearly observed in the dry-film FTIR spectra of all protein hydrolysates, especially in the spectral region 1800-1300 $cm^{-1}$. Fig. 2 displays the second derivative FTIR spectra representing hydrolysis time-series using two enzymes on two raw materials. The most important bands for describing these variations are marked, and include the $NH_3^+$ deformation (1516 $cm^{-1}$), the $COO^-$ stretching (1400 $cm^{-1}$), the amide I (~1650 $cm^{-1}$), and the amide II (~1550 $cm^{-1}$) bands [15,16]. Additional examples of second derivative FTIR time-series of chicken and salmon hydrolysis reactions are provided in SI Fig. S-1.

A few general trends can be deduced from the second derivative FTIR spectra in Fig. 2. The dominating variation seen in the amide I and amide II bands are related to changes in the corresponding secondary structures of the proteins. As the larger proteins are broken down into smaller peptide fragments, the amide I and II changes and simplifies accordingly (i.e., there are fewer peaks in the second derivative spectra after 80 min than in the beginning of the hydrolysis). These changes are most pronounced in hydrolysis of the turkey samples due to the complex protein composition of this raw material. Another observation is the differences seen between the enzymes used. For Alcalase, large changes are observed in all the four bands marked in Fig. 2, while for Flavourzyme, more dominant changes are seen in the signals from the C- and N-terminals. These developments can be explained by the fact that Alcalase contains mostly endo-peptidases, while Flavourzyme, which mainly contains exo-peptidases, releases more free amino acids into the water phase [26].

### 3.2. Multivariate calibrations (PLSR)

The hydrolysates analyzed in the current study were produced using a variety of raw materials and enzymes. This resulted in hydrolysates with large variations in composition and size distribution, as illustrated in the previous section. Due to the extensive spectral variations seen in the FTIR spectra of the hydrolysates, two multivariate regression approaches were used and compared to establish the relationship between the FTIR spectra and the corresponding $M_w$ values: (1) A one-level PLSR model and (2) a two-level PLSR model. In the one-level PLSR approach, the different main subgroups of Table 1 were combined in one single PLSR model. The results are provided in Table 2, and as shown in the table, PLSR models with moderate to high cross-validated coefficients of determination ($R^2$) were obtained. The exception was the turkey protein hydrolysates where the $R^2$ value was 0.455. The PLSR results of Table 2 can also be presented in measured vs. predicted plots, as shown in Fig. 3. Here the raw material groups are color-coded. Fig. 3 reveals that for the one-level PLSR approach, the prediction of lower and higher $M_w$ values were less accurate than the region between approx. 1500-4000 g $mol^{-1}$. The most challenging samples, as also indicated in the previous section, are the hydrolysates originating from turkey.

The results from merging all the data and constructing a one-level PLSR model are shown in Fig. 3 and Table 2. As judged by $R^2$ and the root mean square error of cross-validation (RMSECV), the model is only moderately good. Thus, a two-level PLSR model was utilized to see if any improvements could be achieved. In the two-level modeling approach, the spectra were first classified into 28 subgroups of raw materials + enzymes as defined in Table 1 and then subjected to

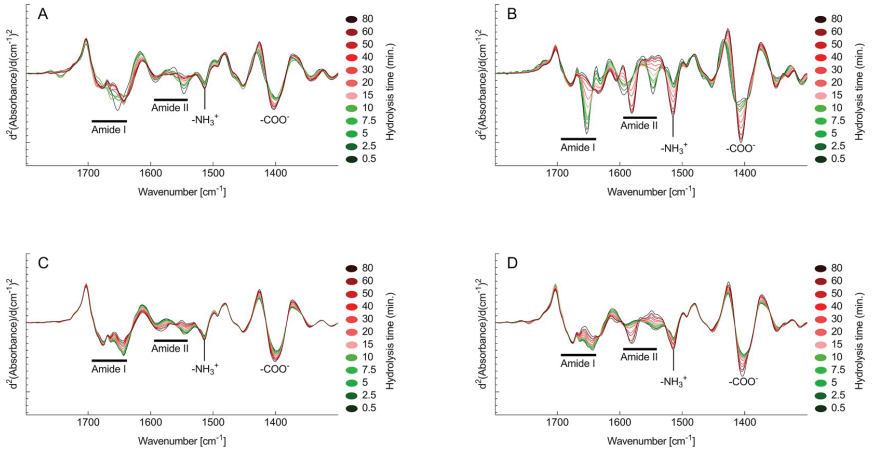



**Fig. 2.** Second derivative FTIR spectra (1800-1300 $cm^{-1}$) of hydrolysate time-series from hydrolysis reactions using two different enzymes and two different raw materials. A and B) Turkey mechanical debone residue. C and D) Mackerel. A and C) Alcalase. B and D) Flavourzyme.

**Table 2**
One-level PLSR model results for different groups of protein hydrolysis samples.

| One-level PLSR model | No. of samples | $R^2$ | RMSECV (g mol$^{-1}$) |
|---|---|---|---|
| **All** | 885 | 0.834 | 446 |
| **Poultry** | 565 | 0.817 | 449 |
| Chicken | 424 | 0.929 | 285 |
| Turkey | 141 | 0.455 | 667 |
| **Fish** | 320 | 0.934 | 298 |
| Salmon | 273 | 0.926 | 299 |
| Mackerel | 47 | 0.992 | 57 |

regression models tuned to each raw material + enzyme subgroup. Both the classification and regression models used PLS for dimension reduction. A score plot showing the two first PLS components of the classification step is provided in Fig. 4. It can be seen from the figure, using the two first components of the classification model, that 81.25% of the samples are correctly classified into subgroups. The score plot provides a very good illustration of the raw material effect on the FTIR spectra. In the plot, there is a main separation between fish raw materials (lower (i.e., mackerel) and left (i.e., salmon) part of the plot) and poultry materials (right and upper part of the plot). The chicken raw materials are found along the whole length of the second component, whereas most of the turkey raw materials are found in the middle part of the plot. There are also significant overlaps between some of the raw material subgroups, and a total of 24 components were needed for correctly classifying 884 of the 885 samples in the current sample set (data not shown). Since this is a supervised classification, it is also interesting to note that the effect of the hydrolysis time, which is one of the major parameters contributing to protein size differences, is virtually absent in the score plot.

The results of the two-level PLSR model are provided in Table 3 and Fig. 5. The table reveals that for all groups, there is a considerable improvement in regression results compared to Table 2. The exception is the mackerel data, which is very well modeled also using the one-level approach. A similar trend is shown in Fig. 5 where large improvements are observed for all raw material groups, especially in the $M_w$ regions that were more challenging using the one-level approach. This shows that using the two-level approach is a feasible tool for quantifying generic features from highly detailed spectroscopic measurements of samples of different origin. However, even when using the two-level approach, the estimation errors for prediction of $M_w$ higher than 4000 g mol$^{-1}$ are still high, especially for the turkey hydrolysates.

As previously suggested, the error in prediction of higher $M_w$ may be largely related to inaccuracies of the reference analysis itself. This is not only due to the SEC column exclusion range, but also partly due to the detection method used. In this study a UV detector was used measuring peptide bonds at 214 nm, which is the most common detection method and wavelength used for this purpose. However, there are some limitations to this method as free amino acids are barely detected at this wavelength, while proteins and peptides are detected by absorption contributions from both peptide bonds and side-groups [27,28]. This, together with poor retention of larger peptides and protein fragments, will result in scaling errors, which in turn will affect how well the

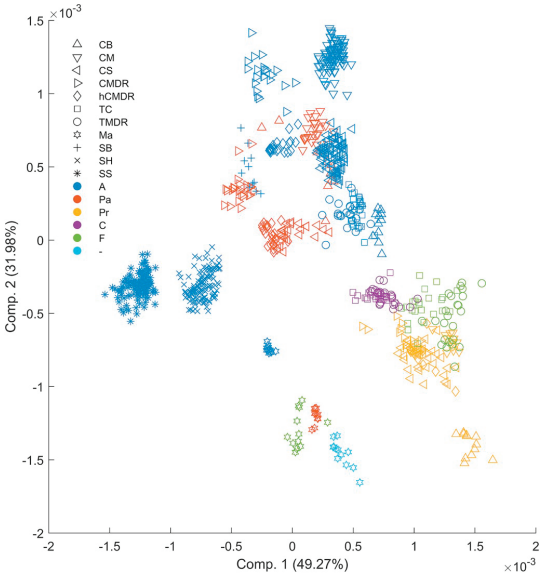



**Fig. 4.** Score plot with the two first components of the PLSR classification. The shape of the points denote raw material origin, whereas enzymes are denoted by color. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

**Table 3**
Two-level PLSR model results for different groups of protein hydrolysis samples.

| Two-level PLSR model | No. of samples | $R^2$ | RMSECV (g mol$^{-1}$) |
|---|---|---|---|
| **All** | 885 | 0.944 | 260 |
| **Poultry** | 565 | 0.926 | 286 |
| Chicken | 424 | 0.989 | 114 |
| Turkey | 141 | 0.651 | 534 |
| **Fish** | 320 | 0.970 | 201 |
| Salmon | 273 | 0.962 | 217 |
| Mackerel | 47 | 0.995 | 44 |

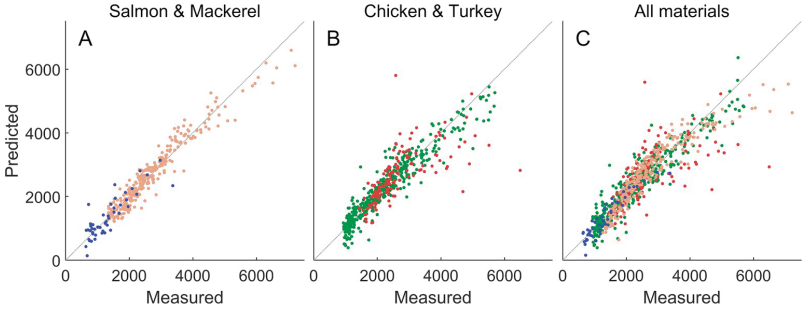



**Fig. 3.** One-level PLSR results, (blue) mackerel, (orange) salmon, (green) chicken and (red) turkey. A) 320 Salmon and mackerel samples. B) 565 Chicken and turkey samples. C) All 885 samples. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

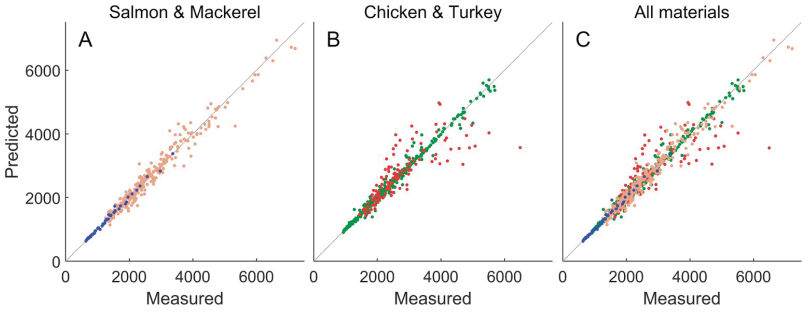



**Fig. 5.** Two-level PLSR results, (blue) mackerel, (orange) salmon, (green) chicken and (red) turkey. A) 320 Salmon and mackerel samples. B) 565 Chicken and turkey samples. C) All 885 samples. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

calculated $M_w$ reflects the actual molecular weight distribution in the sample.

Another potential source of error affecting the relationship between the FTIR spectra and $M_w$ values, is the inherent levels of chemical detail reflected in the FTIR spectra. When the hydrolysates have high molecular weights, larger protein and peptide fragments having a secondary structure will dominate the FTIR spectra, particularly in the amide I and the amide II regions. At lower molecular weights, where the enzymes have broken down some of these larger fragments, the spectral features related to secondary structures will be less pronounced. In addition, the raw materials used in the current study have different complexity levels. Fish raw materials generally have less complex protein composition than poultry raw materials. This will lead to less complexity in the amide I and amide II regions, as illustrated in Fig. 2, which in turn influences the possibility of making an adequate protein size calibration covering many different raw materials. The complexity differences in the FTIR amide bands related to secondary structures between subgroups also serve as a very good illustration of why the two-level PLSR outperforms the one-level PLSR approach in the current study. The regression coefficients of the PLSR models provide a good support for the effects of raw material on prediction accuracy. The regression coefficients of two PLSR models (i.e., chicken muscle hydrolyzed with Alcalase and turkey mechanical debone residue hydrolyzed with Flavourzyme) are shown in Fig. 6. Chicken muscle hydrolysate samples are expected to contain the least complex protein composition of the two raw materials and the $COO^-$ stretching band (around 1400 $cm^{-1}$) is therefore a major feature of the regression coefficients. For the turkey mechanical debone residue, on the other hand, features in the amide region are more dominant. The regression coefficients of the other combinations of raw material and enzyme are presented in SI, Fig. S-2.

The use of two-level or hierarchical modeling is usually related to three conditions: (1) The reference value of interest relates differently to the recorded spectra in different subgroups of the sample material, thus leading to a higher prediction accuracy if a spectrum is predicted with the correct local model. (2) The spectral signals that are used for classifying into subgroups are stable and possibly distinct from the signals used for the reference predictions. (3) The precision of the classification into subgroups is high and/or the consequence of a wrong classification is low because similar subgroups have similar prediction models. If the data being analyzed does not adhere to these conditions, a hierarchical approach may have little value or be negative for the overall prediction accuracy. The hydrolysates analyzed in this study fit perfectly to all three conditions. The dominant variation in the spectra is due to the differences between subgroups (condition 2), while this source of variation is small inside each subgroup and across closely related subgroups (conditions 1 and 3).

One usually imagines a hierarchical model as a structure where samples enter at the top, getting classified through one or more levels

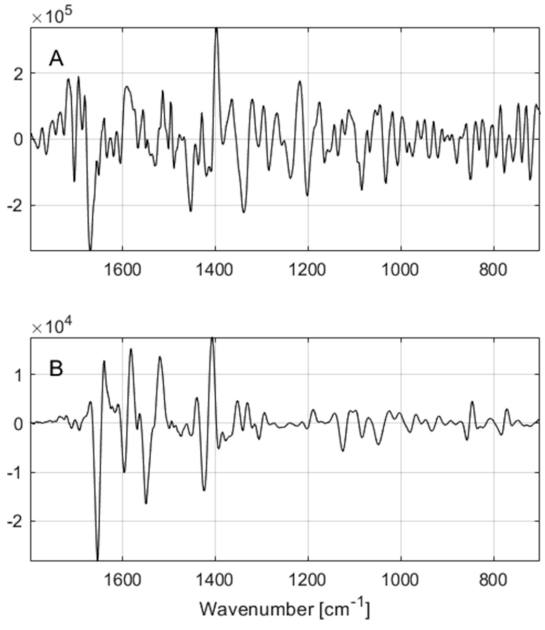



**Fig. 6.** Regression coefficients of the PLSR models of all hydrolysate time-series produced using: A) Chicken muscle and Alcalase. B) Turkey mechanical debone residue and Flavourzyme.

on their way down, while predictions (or classifications) fall out of the bottom of the model. In the two-level PLS model, the upper level is thus the global subgroup classifier (CPLS + LDA), while the lower level consists of a set of PLSR models that return predicted $M_w$ values. The upper level in the approach uses known subgroups as classes. In cases where such subgroups follow the above-mentioned conditions, they are well suited for this type of modeling. If the subgroups were unknown, they might still be possible to infer from the structure of the spectra, e.g. using the clustering approach of HC-PLS. More complex relations between subgroups with different structures in the signals might benefit from more levels in the modeling, e.g. like the Hot PLS approach. In all the hierarchical approaches, a balance between the local adaptiveness of the model, i.e., how finely split the subgroups are, and the size of the subgroups, i.e., how robust and precise the prediction models are, must be found. In this case, optimization for this was not performed, but a good balance was achieved by the sheer number of samples and the

design of the enzyme-raw material combinations that were chosen. This two-level approach may also be further developed to handle unspecified hydrolysate samples (with regard to enzyme and raw material used to hydrolysate). For this, an unsupervised classification system is needed to group samples with similarities. An example of this has been presented by Perez-Guaita D. et al. [19].

It is important to note that full cross-validation was used for all regression models in the present study. As the sample size of the different local regression models in the hierarchical approach varied from 11 samples to 132 samples, the cross-validation approach was the only validation allowing to appropriately compare modeling results. Segmented cross-validation leaving out sets of replicates corresponding to each material-enzyme-time combination was tested, but these results were closely comparable to the ones presented here (< 1% difference in $R^2$, data not shown). In future work, when larger subgroups are present, it will be essential to also employ proper test set validation of the local regression models.

The results of this study clearly illustrate the potential of using FTIR for quantifying protein sizes in a range of different protein hydrolysates. The study also provides a feasible solution for building a generic calibration for protein sizes in hydrolysates. The approach of hierarchical modeling is also expected to be a potential solution in other FTIR approaches where the aim is to quantify a generic component in different raw materials (e.g. fatty acid quantification across microbial strains). As the use of dry-film FTIR for automated high-throughput analysis including automated sample handling and robotics is gaining increasing attention, a commercial system for protein size estimations in enzymatic protein hydrolysis industry could thus be expected when proper technical developments are made. A tool for protein size estimation would potentially also find applications in a range of different fields, including reaction kinetics, *in vitro* protein digestion, protein production by fermentation, and characterization of protein and peptide compositions of dairy products.

## 4. Conclusion

In the present study, we have shown that $M_w$ of protein hydrolysates can be predicted with high accuracy using FTIR spectroscopy. The best result was obtained using a hierarchical PLSR approach where FTIR spectra of the protein hydrolysates were classified according to raw material type and enzyme prior to local modeling. This shows that prediction of protein sizes in protein hydrolysates can be achieved for a range of different raw materials using a single mathematical model. The results therefore demonstrate the potential of using FTIR for monitoring protein sizes during enzymatic protein hydrolysis in industrial settings, while also paving the way for measurements of protein sizes in other applications.

## CRediT authorship contribution statement

**Kenneth Aase Kristoffersen:** Writing - original draft, Data curation, Conceptualization, Visualization, Investigation, Formal analysis. **Kristian Hovde Liland:** Writing - original draft, Software, Methodology, Conceptualization, Visualization, Investigation. **Ulrike Böcker:** Data curation, Conceptualization, Writing - review & editing. **Sileshi Gizachew Wubshet:** Conceptualization, Writing - review & editing. **Diana Lindberg:** Writing - review & editing. **Svein Jarle Horn:** Writing - review & editing. **Nils Kristian Afseth:** Writing - original draft, Writing - review & editing, Conceptualization, Investigation, Methodology.

## Acknowledgement

Ria Das, Ornella Fleury, Rasmus Karstad and Marte Ryen Dalsnes are acknowledged for excellent technical assistance. Financial support from the Norwegian Fund for Research Fees for Agricultural Products, Norway through the projects “FoodSMaCK” (no. 262308) and “SunnMat” (no. 262300), and from the Norwegian Research Council, Norway through the project “Notably” (no. 280709) is greatly acknowledged. Internal funding from Nofima through the project “PepTek” is also greatly acknowledged.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.talanta.2019.06.084.

## References

[1] U. Böcker, R. Ofstad, H.C. Bertram, B. Egelandsdal, A. Kohler, Salt-induced changes in pork myofibrillar tissue investigated by FT-IR microspectroscopy and light microscopy, J. Agric. Food Chem. 54 (18) (2006) 6733–6740.
[2] P. Gelfand, R.J. Smith, E. Stavitski, D.R. Borchelt, L.M. Miller, Characterization of protein structural changes in living cells using time-lapsed FTIR imaging, Anal. Chem. 87 (12) (2015) 6025–6031.
[3] A. Barth, The infrared absorption of amino acid side chains, Prog. Biophys. Mol. Biol. 74 (3) (2000) 141–173.
[4] N. Perisic, N.K. Afseth, R. Ofstad, A. Kohler, Monitoring protein structural changes and hydration in bovine meat tissue due to salt substitutes by fourier transform infrared (FTIR) microspectroscopy, J. Agric. Food Chem. 59 (18) (2011) 10052–10061.
[5] P.V. Andersen, E. Veiseth-Kent, J.P. Wold, Analyzing pH-induced changes in a myofibril model system with vibrational and fluorescence spectroscopy, Meat Sci. 125 (2017) 1–9.
[6] S.H. Arabi, B. Aghelnejad, C. Schwieger, A. Meister, A. Kerth, D. Hinderberger, Serum albumin hydrogels in broad pH and temperature ranges: Characterization of their self-assembled structures and nanoscopic and macroscopic properties, Biomater. Sci. 6 (3) (2018) 478–492.
[7] G. Martra, C. Deiana, Y. Sakhno, I. Barberis, M. Fabbiani, M. Pazzi, M. Vincenti, The Formation and self-assembly of long prebiotic oligomers produced by the condensation of unactivated amino acids on oxide surfaces, Angew. Chem. Int. Ed. 53 (18) (2014) 4671–4674.
[8] H.-F. Okabayashi, H.-H. Kanbe, C.J. O'Connor, The role of an L-leucine residue on the conformations of glycyl-L-leucine oligomers and its N- or C-terminal dependence: Infrared absorption and Raman scattering studies, Eur. Biophys. J. 45 (1) (2016) 23–34.
[9] C. Ruckebusch, L. Duponchel, J.P. Huvenne, P. Legrand, N. Nedjar-Arroume, B. Lignot, P. Dhulster, D. Guillochon, Hydrolysis of hemoglobin surveyed by infrared spectroscopy II. Progress predicted by chemometrics, Anal. Chim. Acta 396 (2–3) (1999) 241–251.
[10] C. Ruckebusch, L. Duponchel, J.P. Huvenne, Degree of hydrolysis from mid-infrared spectra, Anal. Chim. Acta 446 (1–2) (2001) 257–268.
[11] C. Ruckebusch, B. Sombret, R. Froidevaux, J.P. Huvenne, On-line mid-infrared spectroscopic data and chemometrics for the monitoring of an enzymatic hydrolysis, Appl. Spectrosc. 55 (12) (2001) 1610–1617.
[12] G. Guler, E. Dzafic, M.M. Vorob'ev, V. Vogel, W. Mantele, Real time observation of proteolysis with Fourier transform infrared (FT-IR) and UV-circular dichroism spectroscopy: Watching a protease eat a protein, Spectrochim. Acta Mol. Biomol. Spectrosc. 79 (1) (2011) 104–111.
[13] G. Guler, M.M. Vorob'ev, V. Vogel, W. Mantele, Proteolytically-induced changes of secondary structural protein conformation of bovine serum albumin monitored by Fourier transform infrared (FT-IR) and UV-circular dichroism spectroscopy, Spectrochim. Acta Mol. Biomol. Spectrosc. 161 (2016) 8–18.
[14] N.A. Poulsen, C.E. Eskildsen, M. Akkerman, L.B. Johansen, M.S. Hansen, P.W. Hansen, T. Skov, L.B. Larsen, Predicting hydrolysis of whey protein by mid-infrared spectroscopy, Int. Dairy J. 61 (2016) 44–50.
[15] U. Böcker, S.G. Wubshet, D. Lindberg, N.K. Afseth, Fourier-transform infrared spectroscopy for characterization of protein chain reductions in enzymatic reactions, Analyst 142 (15) (2017) 2812–2818.
[16] S.G. Wubshet, I. Mage, U. Böcker, D. Lindberg, S.H. Knutsen, A. Rieder, D.A. Rodriguez, N.K. Afseth, FTIR as a rapid tool for monitoring molecular weight distribution during enzymatic protein hydrolysis of food processing by-products, Anal. Methods 9 (29) (2017) 4247–4254.
[17] K. Tøndel, U.G. Indahl, A.B. Gjuvsland, J.O. Vik, P. Hunter, S.W. Omholt, H. Martens, Hierarchical Cluster-based Partial Least Squares Regression (HC-PLSR) is an efficient tool for metamodelling of nonlinear dynamic models, BMC Syst. Biol. 5 (1) (2011) 90.
[18] K.H. Liland, A. Kohler, V. Shapaval, Hot PLS—a framework for hierarchically ordered taxonomic classification by partial least squares, Chemometr. Intell. Lab. Syst. 138 (2014) 41–47.
[19] D. Perez-Guaita, J. Kuligowski, G. Quintás, S. Garrigues, M.d.l. Guardia, Modified locally weighted—partial least squares regression improving clinical predictions from infrared spectra of human serum samples, Talanta 107 (2013) 368–375.
[20] O. Preisner, R. Guiomar, J. Machado, J.C. Menezes, J.A. Lopes, Application of fourier transform infrared spectroscopy and chemometrics for differentiation of Salmonella enterica serovar enteritidis phage types, Appl. Environ. Microbiol. 76 (11) (2010) 3538–3544.
[21] D.S.T. Hsieh, C. Lin, E.R. Lang, N. Catsimpoolas, C.K. Rha, Molecular-weight

distribution of soybean globulin peptides produced by peptic hydrolysis, Cereal Chem. 56 (4) (1979) 227–231.
[22] N.K. Afseth, A. Kohler, Extended multiplicative signal correction in vibrational spectroscopy, a tutorial, Chemometr. Intell. Lab. Syst. 117 (2012) 92–99.
[23] U.G. Indahl, K.H. Liland, T. Naes, Canonical partial least squares—a unified PLS approach to classification and regression problems, J. Chemom. 23 (9) (2009) 495–504.
[24] M.C. Archer, J.O. Ragnarsson, S.R. Tannenbaum, D.I.C. Wang, Enzymatic solubilization of an insoluble substrate, fish protein concentrate: Process and kinetic considerations, Biotechnol. Bioeng. 15 (1) (1973) 181–196.
[25] R. Muñoz-Tamayo, J. de Groot, P.A. Wierenga, H. Gruppen, M.H. Zwietering, L. Sijtsma, Modeling peptide formation during the hydrolysis of β-casein by Lactococcus lactis, Process Biochem. 47 (1) (2012) 83–93.
[26] M. Merz, W. Claaßen, D. Appel, P. Berends, S. Rabe, I. Blank, T. Stressler, L. Fischer, Characterization of commercially available peptidases in respect of the production of protein hydrolysates with defined compositions using a three-step methodology, J. Mol. Catal. B Enzym. 127 (2016) 1–10.
[27] P. Hong, S. Koza, E.S.P. Bouvier, A review size-exclusion chromatography for the analysis of protein biotherapeutics and there aggregates, J. Liq. Chromatogr. Relat. Technol. 35 (20) (2012) 2923–2950.
[28] G. Brusotti, E. Calleri, R. Colombo, G. Massolini, F. Rinaldi, C. Temporini, Advances on size exclusion chromatography and applications on the analysis of protein biopharmaceuticals and protein aggregates: A mini review, Chromatographia 81 (1) (2018) 3–23.

# Paper II

# 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*], 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: Kenneth Aase Kristoffersen, Kenneth.kristoffersen@nofima.no*

## Abstract

Fourier-transform infrared (FTIR) spectroscopy was applied to predict 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 could therefore be essential for process control and optimization in industry. In the present study the intercorrelation and the complimentary nature of these parameters were investigated and linked by a partial least square regression (PLSR) model predicting DH% from molecular weight distributions. Finally, we developed PLSR models based on dry-film FTIR spectroscopy for prediction of both DH% and $M_w$. Here, spectral changes in the amide region underline the need for dry-film spectra, as the amide I band was found to be important for the two calibration models. 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 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 prediction of DH%. However, such 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 reactions 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 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). 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 ºC. 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.

$$\mathrm{DH\%} = \frac{h}{h_{tot}} \times 100\%$$

Equation 1

## 2.5 Size exclusion chromatography

SEC was preformed according to Wubshet et al. using 2 mg/mL solutions of standards and rehydrated hydrolysate samples (1% w/v, filtrated using Millex-HV PVDF 0.45 µm 33 mm filters (MilliporeSigma, Burlington, MA, USA)) as injection solutions (Wubshet, et al., 2017). Chromatographic separation of standards and samples was performed with a Thermo Scientific Dionex UltiMate 3000 Standard System (Thermo Fisher Scientific). The injection volume was 10 µL for the standards and 15 µL for samples. Separation was performed at 25 °C using a BioSep-SEC-s2000 column (300 × 7.8 mm, Phenomenex, Torrence, CA, USA). The mobile phase consisted of a mixture of acetonitrile and ultrapure water in a proportion 30:70 (v/v), containing 0.05 % trifluoroacetic acid. Isocratic elution was carried out using a flow rate of 0.9 mL/min for 20.0 minutes. Between 20.0 and 20.1 minutes the mobile phase was changed to $NaH_2PO_4$ (0.10 M) and maintained until 23.0 minutes for column cleaning. Elution conditions were restored between minute 23.0 and 23.1 and the column was equilibrated for an additional 27 minutes. Chromatographic runs were controlled from Chromeleon™ Chromatography Data System (CDS) software (Thermo Fisher Scientific). From chromatographic runs of both the standards and hydrolysates, a UV trace of 214 nm was used. The retention times of analytical standards were obtained by manual peak-picking. The retention times of the standards were used to construct a third polynomial ($r^2$=0.97) fitted

169 calibration curve (Vander Heyden, Popovici, & Schoenmakers, 2002). The retention times for
170 the standards are presented in SI Table S-1. Finally, $M_w$ were calculated using PSS winGPC
171 UniChrom V 8.00 (Polymer Standards Service, Mainz, Germany) for each chromatogram.
172 The calculation from the software was based on a slicing method, similar to those previously
173 used for analysis of protein hydrolysates (Hsieh, Lin, Lang, Catsimpoolas, & Rha, 1979). The
174 calculated $M_w$ values are presented in SI Table S-3.

175

176 **2.6 Reproducibility study**

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

182

183 **2.7 FTIR spectroscopy**

184 The samples for FTIR measurements were prepared by rehydration and filtration of the
185 supernatant (5% w/v, filtrated using Millex-HV PVDF 0.45 µm 33 mm filters
186 (MilliporeSigma, Burlington, MA, USA)). For all of the hydrolysates aliquots of 5 µL were
187 deposited on 96-well IR-transparent Si-plates (Bruker, Billerica, MA, USA) and dried at room
188 temperature for at least 30 minutes to form dry-films as described by Böcker et al. (Böcker,
189 Wubshet, Lindberg, & Afseth, 2017). From each hydrolysate sample, five aliquots were
190 deposited to allow for replicate measurements. FTIR measurements were performed using a
191 High Throughput Screening eXTension (HTS-XT) unit coupled to a Tensor 27 spectrometer
192 (Bruker, Billerica, MA, USA). The spectra were recorded in the region between 4000 and 400

193 $cm^{-1}$ with a spectral resolution of 4 $cm^{-1}$ and an aperture of 5.0 mm. For each spectrum, 40
194 interferograms were collected and averaged. Data acquisition was controlled using Opus v6.5
195 (Bruker, Billerica, MA, USA).

196

197 **2.8 Data analysis**

198 Pre-processing of FTIR spectra was performed using Savitzky-Golay algorithm with a
199 polynomial degree of two and a window size of 13 points. The second-derivative spectra were
200 then normalized by applying extended multiplicative signal correction (EMSC) (Afseth &
201 Kohler, 2012). The pre-treated data sets from the different samples were subjected to partial
202 least squares regression (PLSR) modeling to predict DH% and $M_w$. For the PLSR, the spectral
203 region from 1800-700 $cm^{-1}$ was used. The regions from 5-20 min of the SEC chromatogram
204 were also subjected to pre-processing using area normalization. The pre-treated datasets from
205 5-16 min of the SEC chromatograms were then subjected to PLSR modeling to predict DH%.
206 Data processing and analysis were carried out using The Unscrambler® X v10.3 (CAMO
207 Software AS, Oslo, Norway).

208

209 # 3 Results and discussion

210 Three dairy protein sources (WPC80, WMP and MPO) were hydrolyzed for up to three hours
211 using two commercially available enzyme products (Alcalase (A) and Flavourzyme (F)). The
212 reactions were stopped by thermal inactivation of the enzymes before samples were prepared
213 for FTIR, SEC, nitrogen content and DH% analysis. The raw materials used are known to
214 consist of a mixture of many different proteins. Whole milk protein consist of approximately
215 20% whey proteins with major components α-lactalbumin, β-lactoglobulin, and 80% caseins,

divided into major subclasses α- ($\alpha_{S1}$- and $\alpha_{S2}$-), 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 yielding single amino acids or di- or tri-peptides. As such, these are representative for 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 peptides from the peptide ends, and instead 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, in other words, the none-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 Cross Validated (RMSECV) of 0.86 % and a cross-validated coefficient of determination ($R^2$) 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 underestimation of the higher DH% values and 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 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 good representation of the actual MWD.

The FTIR-based PLSR prediction models displayed in Fig. 4 provided prediction models with a 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% were predicted from FTIR spectra, indicating that the limitation 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 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. Interestingly, there are features in the PLSR models appearing to describe the inherent differences. This shows the industrial relevance of dry-film FTIR-based models, as it is opens for rapid measurements of DH% and $M_w$ as a tool to monitor quality development during EPH processes.

## 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 prediction of both $M_w$ and DH% in milk protein hydrolysates in industrial settings.

## 5 Acknowledgements

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 weight of peptides released in the course of enzymatic proteolysis. *Journal of Chromatography B, 1105*, 1-9.

Böcker, 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.

Chi, C. F., Cao, Z. H., Wang, B., Hu, F. Y., Li, Z. R., & Zhang, B. (2014). Antioxidant and Functional Properties of Collagen Hydrolysates from Spanish Mackerel Skin as Influenced by Average Molecular Weight. *Molecules, 19*(8), 11211-11230.

Church, F. C., Swaisgood, H. E., Porter, D. H., & Catignani, G. L. (1983). Spectrophotometric Assay Using o-Phthaldialdehyde for Determination of Proteolysis in Milk and Isolated Milk Proteins. *Journal of Dairy Science, 66*(6), 1219-1227.

Gellrich, K., Meyer, H. H. D., & Wiedemann, S. (2014). Composition of major proteins in cow milk differing in mean protein concentration during the first 155 days of lactation and the influence of season as well as short-term restricted feeding in early and mid-lactation. *Czech Journal of Animal Science, 59*(3), 97-106.

Guler, G., Dzafic, E., Vorob'ev, M. M., Vogel, V., & Mantele, W. (2011). Real time observation of proteolysis with Fourier transform infrared (FT-IR) and UV-circular dichroism spectroscopy: Watching a protease eat a protein. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 79*(1), 104-111.

Guler, G., Vorob'ev, M. M., Vogel, V., & Mantele, W. (2016). Proteolytically-induced changes of secondary structural protein conformation of bovine serum albumin monitored by Fourier transform infrared (FT-IR) and UV-circular dichroism spectroscopy. *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy, 161*, 8-18.

Hong, P., Koza, S., & Bouvier, E. S. P. (2012). A review Size-exclusion chromatography for the analysis of protein biotherapeutics and their aggregates. *Journal of Liquid Chromatography & Related Technologies, 35*(20), 2923-2950.

Hsieh, D. S. T., Lin, C., Lang, E. R., Catsimpoolas, N., & Rha, C. K. (1979). Molecular-weight distribution of soybean globulin peptides produced by peptic hydrolysis. *Cereal Chemistry, 56*(4), 227-231.

Kristoffersen, K. A., Liland, K. H., Böcker, U., Wubshet, S. G., Lindberg, D., Horn, S. J., & Afseth, N. K. (2019). FTIR-based hierarchical modeling for prediction of average molecular weights of protein hydrolysates. *Talanta, 205*, 120084.

Kuipers, B. J. H., & Gruppen, H. (2007). Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative reverse phase high-performance liquid chromatography-mass spectrometry analysis. *Journal of Agricultural and Food Chemistry, 55*(14), 5445-5451.

Li, Z., Wang, B., Chi, C., Gong, Y., Luo, H., & Ding, G. (2013). Influence of average molecular weight on antioxidant and functional properties of cartilage collagen hydrolysates from Sphyrna lewini, Dasyatis akjei and Raja porosa. *Food Research International, 51*(1), 283-293.

Lordan, S., Ross, R. P., & Stanton, C. (2011). Marine Bioactives as Functional Food Ingredients: Potential to Reduce the Incidence of Chronic Diseases. *Marine Drugs, 9*(6), 1056-1100.

Martinez-Maqueda, D., Miralles, B., Recio, I., & Hernandez-Ledesma, B. (2012). Antihypertensive peptides from food proteins: a review. *Food & Function, 3*(4), 350-361.

Merz, M., Claaßen, W., Appel, D., Berends, P., Rabe, S., Blank, I., Stressler, T., & Fischer, L. (2016). Characterization of commercially available peptidases in respect of the production of protein hydrolysates with defined compositions using a three-step 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.

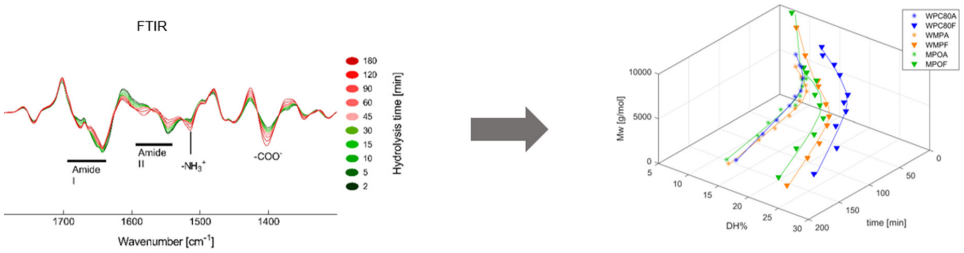
## Tables

**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).

# Figures




**Graphical abstract**

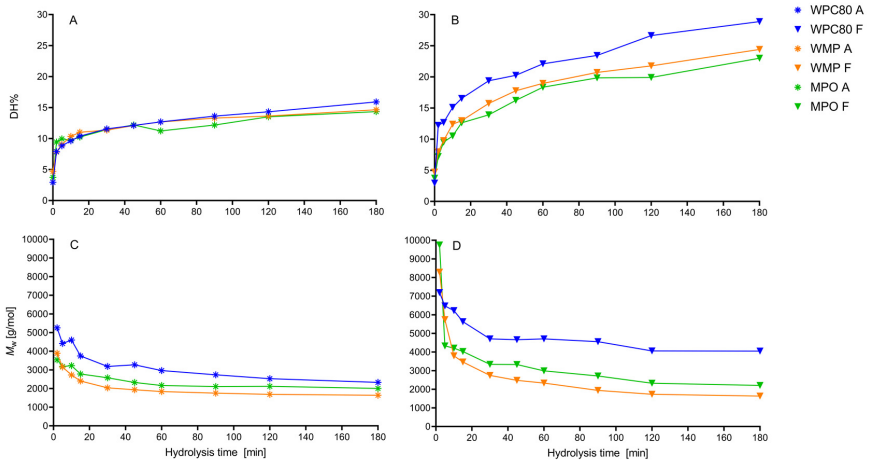
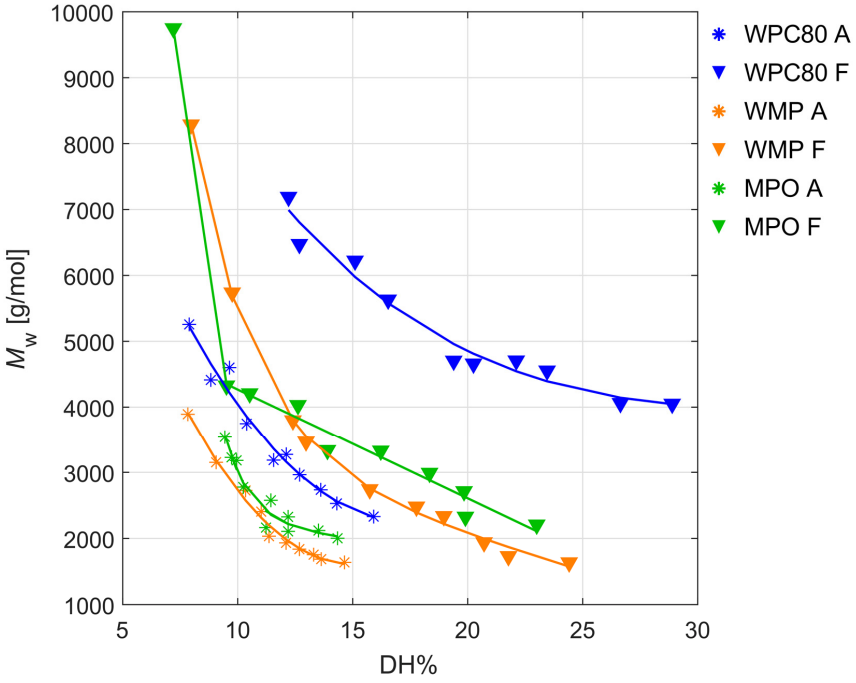


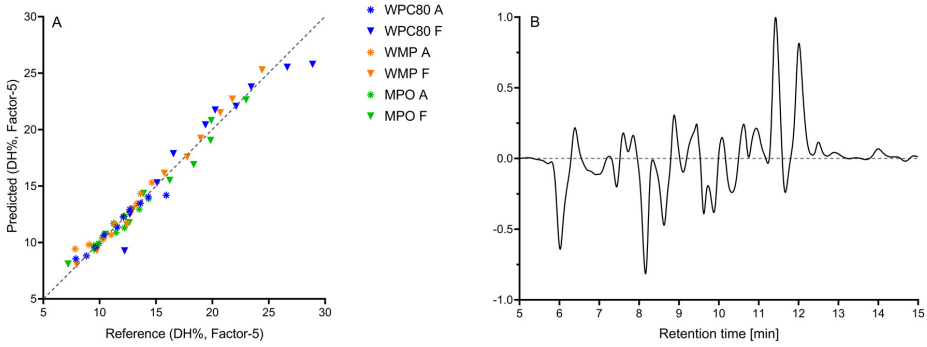

**Fig. 1:** DH% development (over time) for hydrolysis reactions of three different substrates using A) Alcalase and B) Flavourzyme. $M_w$ development (over time) for hydrolysis reactions of three different substrates using C) Alcalase and D) Flavourzyme.




508

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

510




511

512 **Fig. 3:** A) PLSR correlation plot of DH% for six hydrolysis reactions. Predicted from area

513 normalized and mid-centered chromatograms. B) The normalized regression coefficients of

514 the chromatograms for the PLSR model.

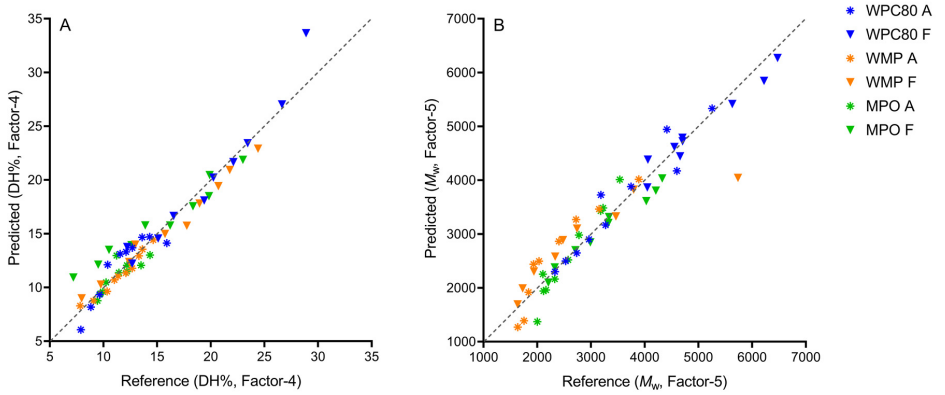
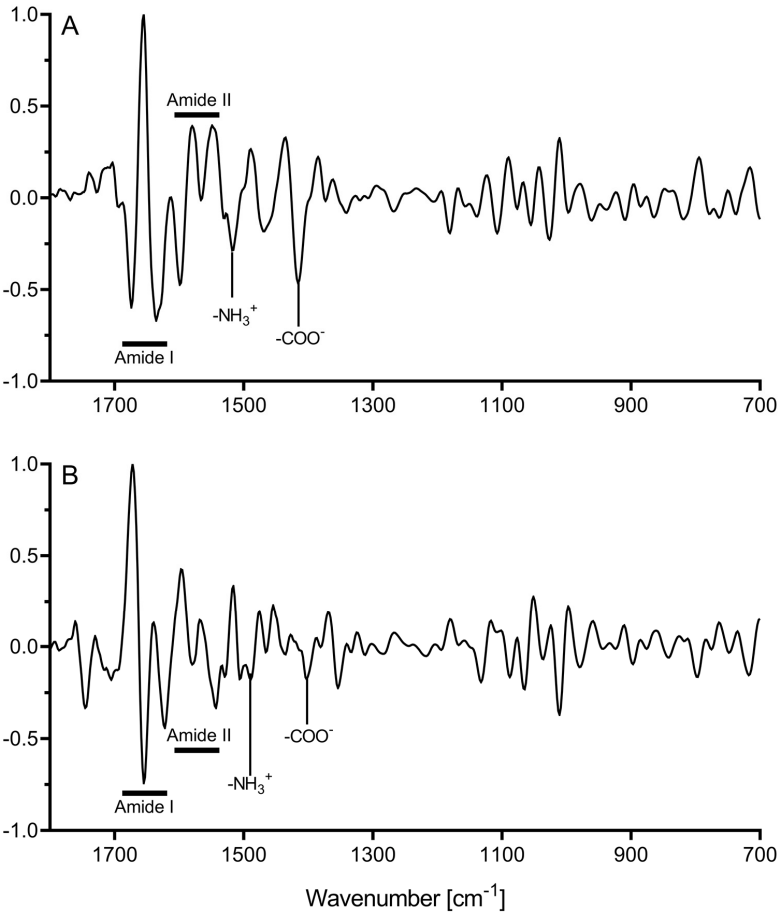



**Fig. 4:** A) PLSR correlation plot DH%, 60 samples B) PLSR correlation plot of $M_w$, 57 samples.




520

521 **Fig. 5:** Normalized regression coefficients for PLSR model of DH% (Fig. 5A) and $M_w$ (Fig.

522 5B).

523

524

# Supporting Information

## 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*], 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: Kenneth Aase Kristoffersen, Kenneth.kristoffersen@nofima.no*

This document contains:

Six tables; the first table contains retention times for all standards used to calculate $M_w$, the second contains DH% values and the third contains $M_w$ values. The fourth table contains measured protein content for all samples included in this study and the two last tables contains data from the validation study of the DH% and the $M_w$ method. Two figures are also included, the first containing all SEC chromatograms and the second figure contains second derivative FTIR spectra for the six different hydrolysis time series.

547 **Table S-1:** Retention times for calibration standards used.

| **Compound name** | M, Wt, (g/mol) | RT (min) |
|---|---|---|
| Albumin from chicken egg white | 44287 | 6.071 |
| Carbonic anhydrase | 29000 | 6.003 |
| Lysozyme | 14300 | 6.626 |
| Aprotinin from bovine lung | 6511 | 6.865 |
| Insulin chain B oxidized from bovine pancreas | 3496 | 8.763 |
| Renin substrate tetradecapeptide porcine | 1759 | 8.133 |
| Angiotensin II human | 1046 | 8.724 |
| Bradykinin Fragment 1-7 | 757 | 9.208 |
| [D-Ala$^2$]-leucine enkephalin | 570 | 11.377 |
| Val-Tyr-Val | 379 | 10.925 |
| Tryptophan | 204 | 11.950 |

548 The column was calibrated with three replicates of each standers for the first time use on the HPLC system. The
549 column was recalibrated with one replicate of each standard prior to the reproducibility study.

550

551

552 **Table S-2:** DH% values for all EPH samples.

| Time (min) | WPC80 A | WMP A | MPO A | WPC80 F | WMP F | MPO F |
|---|---|---|---|---|---|---|
| **2** | 7.89 | 7.83 | 9.44 | 12.21 | 7.97 | 7.20 |
| **5** | 8.83 | 9.06 | 9.95 | 12.68 | 9.76 | 9.51 |
| **10** | 9.64 | 10.34 | 9.73 | 15.10 | 12.40 | 10.52 |
| **15** | 10.39 | 11.02 | 10.24 | 16.54 | 12.97 | 12.62 |
| **30** | 11.56 | 11.36 | 11.44 | 19.39 | 15.74 | 13.90 |
| **45** | 12.11 | 12.10 | 12.19 | 20.25 | 17.77 | 16.22 |
| **60** | 12.69 | 12.68 | 11.23 | 22.11 | 18.96 | 18.34 |
| **90** | 13.61 | 13.30 | 12.19 | 23.45 | 20.71 | 19.84 |
| **120** | 14.31 | 13.63 | 13.51 | 26.64 | 21.77 | 19.90 |
| **180** | 15.91 | 14.64 | 14.34 | 28.89 | 24.41 | 23.00 |

553 All samples were measured in triplicates.

554

555 **Table S-3:** $M_w$ (g/mol) values for all EPH samples.

| Time (min) | WPC80 A | WMP A | MPO A | WPC80 F | WMP F | MPO F |
|---|---|---|---|---|---|---|
| **2** | 5256.50 | 3891.40 | 3538.20 | 7189.80 | 8288.10 | 9748.40 |
| **5** | 4413.40 | 3151.70 | 3180.70 | 6474.50 | 5737.90 | 4326.50 |
| **10** | 4600.70 | 2725.30 | 3226.30 | 6221.70 | 3796.10 | 4208.00 |
| **15** | 3746.90 | 2405.30 | 2776.60 | 5630.70 | 3465.90 | 4029.80 |
| **30** | 3185.20 | 2034.20 | 2576.60 | 4703.60 | 2741.10 | 3335.50 |
| **45** | 3272.00 | 1933.20 | 2326.00 | 4659.60 | 2478.60 | 3326.20 |
| **60** | 2964.40 | 1838.40 | 2164.00 | 4706.10 | 2334.90 | 2989.40 |
| **90** | 2735.10 | 1754.80 | 2107.10 | 4553.60 | 1939.30 | 2711.20 |
| **120** | 2528.70 | 1687.10 | 2119.90 | 4060.50 | 1730.00 | 2325.50 |
| **180** | 2330.40 | 1638.10 | 2002.70 | 4050.10 | 1636.10 | 2207.00 |

556 All samples were measured ones.

**Table S-4:** Protein content in percent (6.25*)*. Values for all EPH samples.

| Time (min) | WPC80 A | WMP A | MPO A | WPC80 F | WMP F | MPO F |
|---|---|---|---|---|---|---|
| **2** | 60.87 | 22.31 | 26.38 | 61.26 | 23.56 | 31.13 |
| **5** | 61.24 | 27.63 | 26.81 | 61.16 | 21.81 | 22.94 |
| **10** | 63.13 | 27.81 | 26.25 | 60.16 | 19.25 | 23.44 |
| **15** | 63.25 | 30.19 | 30.25 | 59.68 | 21.44 | 21.69 |
| **30** | 65.50 | 32.19 | 29.88 | 57.95 | 21.44 | 24.88 |
| **45** | 66.75 | 34.19 | 33.19 | 60.18 | 23.38 | 24.31 |
| **60** | 68.31 | 34.13 | 34.31 | 60.15 | 23.75 | 24.19 |
| **90** | 69.13 | 36.69 | 32.50 | 61.54 | 27.00 | 24.50 |
| **120** | 70.19 | 36.88 | 33.31 | 62.24 | 29.00 | 27.25 |
| **180** | 70.75 | 33.06 | 34.44 | 63.31 | 29.75 | 28.38 |

All samples were measured in triplicates.

560 **Table S-5:** An overview of the Intra- and interday validation study of the DH% method.

| | | Intraday (n=6) | | | Interday (n=6) | | |
|---|---|---|---|---|---|---|---|
| **TNBS DH%** | **Hydrolysis time (min)** | **Average** | **SD** | **%RSD** | **Average** | **SD** | **%RSD** |
| WPC80 5 A | 5 | 9.28 | 0.59 | 6.31 | 8.92 | 0.42 | 4.70 |
| WPC80 30 A | 30 | 12.13 | 0.46 | 3.80 | 11.52 | 0.70 | 6.08 |
| WPC80 120 A | 120 | 14.76 | 0.40 | 2.74 | 14.45 | 0.51 | 3.54 |
| WPC80 5 F | 5 | 13.30 | 0.80 | 6.01 | 12.69 | 0.50 | 3.94 |
| WPC80 30 F | 30 | 19.49 | 0.72 | 3.68 | 18.89 | 0.69 | 3.67 |
| WPC80 120 F | 120 | 27.06 | 0.24 | 0.89 | 26.70 | 0.75 | 2.79 |
| WMP 5 A | 5 | 8.17 | 1.00 | 12.18 | 8.41 | 0.96 | 11.36 |
| WMP 30 A | 30 | 11.53 | 0.55 | 4.77 | 11.95 | 0.94 | 7.90 |
| WMP 120 A | 120 | 13.75 | 0.41 | 3.00 | 14.07 | 0.88 | 6.25 |
| WMP 5 F | 5 | 9.99 | 0.68 | 6.79 | 9.19 | 0.55 | 5.98 |
| WMP 30 F | 30 | 15.97 | 0.61 | 3.83 | 16.35 | 0.84 | 5.15 |
| WMP 120 F | 120 | 21.84 | 0.46 | 2.13 | 22.09 | 0.93 | 4.20 |
| MPO 5 A | 5 | 9.40 | 0.87 | 9.21 | 10.28 | 0.89 | 8.68 |
| MPO 30 A | 30 | 11.52 | 0.68 | 5.86 | 11.38 | 0.70 | 6.14 |
| MPO 120 A | 120 | 13.65 | 0.84 | 6.13 | 13.32 | 1.06 | 7.94 |
| MPO 5 F | 5 | 9.68 | 0.63 | 6.53 | 9.76 | 0.85 | 8.69 |
| MPO 30 F | 30 | 13.91 | 0.31 | 2.19 | 13.77 | 1.25 | 9.08 |
| MPO 120 F | 120 | 19.99 | 0.47 | 2.34 | 20.06 | 0.46 | 2.27 |

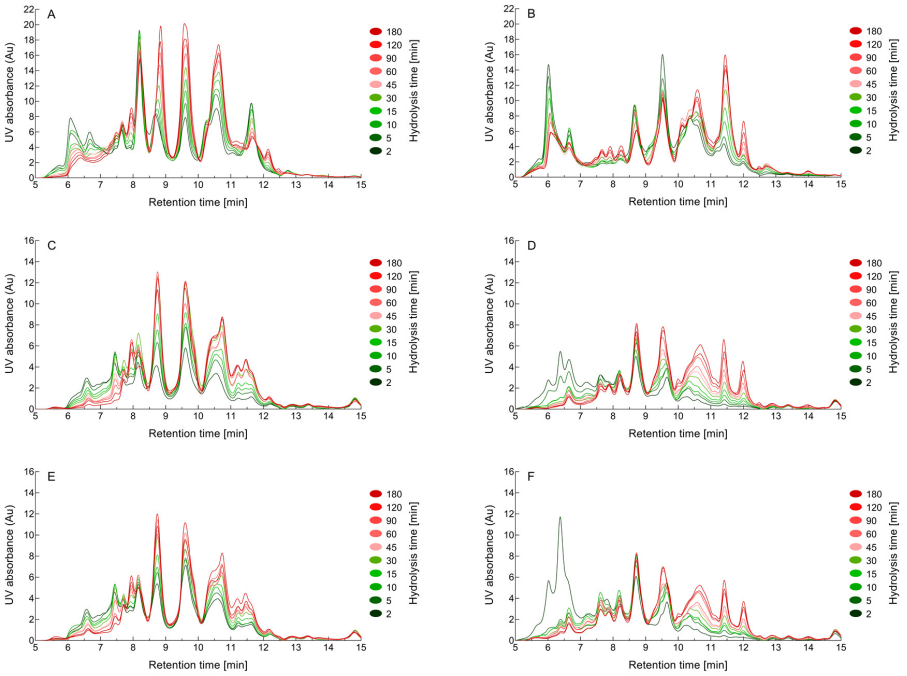
561

562

**Table S-6:** An overview of the Intra- and interday validation study of the $M_w$ method

| | | Intraday (n=6) | | | Interday (n=6) | | |
|---|---|---|---|---|---|---|---|
| $M_w$ | Hydrolysis time (min) | Average | SD | %RSD | Average | SD | %RSD |
| WPC80 5 A | 5 | 4231.88 | 31.68 | 0.75 | 4449.42 | 110.24 | 2.48 |
| WPC80 30 A | 30 | 3140.12 | 39.02 | 1.24 | 3232.37 | 49.30 | 1.53 |
| WPC80 120 A | 120 | 2523.26[a] | 19.06 | 0.76 | 2568.82 | 39.25 | 1.53 |
| WPC80 5 F | 5 | 6474.43 | 11.19 | 0.17 | 6463.65 | 29.68 | 0.46 |
| WPC80 30 F | 30 | 4582.93 | 17.80 | 0.39 | 4757.80 | 109.04 | 2.29 |
| WPC80 120 F | 120 | 3867.70 | 8.10 | 0.21 | 4104.78 | 128.72 | 3.14 |
| WMP 5 A | 5 | 3238.45 | 21.57 | 0.67 | 3206.03 | 38.87 | 1.21 |
| WMP 30 A | 30 | 2071.22 | 10.10 | 0.49 | 2055.93 | 24.73 | 1.20 |
| WMP 120 A | 120 | 1692.35 | 6.19 | 0.37 | 1701.18 | 15.46 | 0.91 |
| WMP 5 F | 5 | 5595.62 | 23.92 | 0.43 | 5716.60 | 206.91 | 3.62 |
| WMP 30 F | 30 | 2799.32 | 18.58 | 0.66 | 2800.03 | 40.23 | 1.44 |
| WMP 120 F | 120 | 1795.92 | 25.73 | 1.43 | 1765.40 | 23.69 | 1.34 |
| MPO 5 A | 5 | 3258.40 | 24.47 | 0.75 | 3251.35 | 56.66 | 1.74 |
| MPO 30 A | 30 | 2627.40 | 20.07 | 0.76 | 2614.33 | 33.66 | 1.29 |
| MPO 120 A | 120 | 2002.92 | 19.17 | 0.96 | 2085.52 | 53.23 | 2.55 |
| MPO 5 F | 5 | 4049.52 | 82.33 | 2.03 | 4294.68 | 141.03 | 3.28 |
| MPO 30 F | 30 | 3330.90 | 37.74 | 1.13 | 3379.55 | 59.35 | 1.76 |
| MPO 120 F | 120 | 2380.97 | 27.70 | 1.16 | 2388.93 | 46.80 | 1.96 |

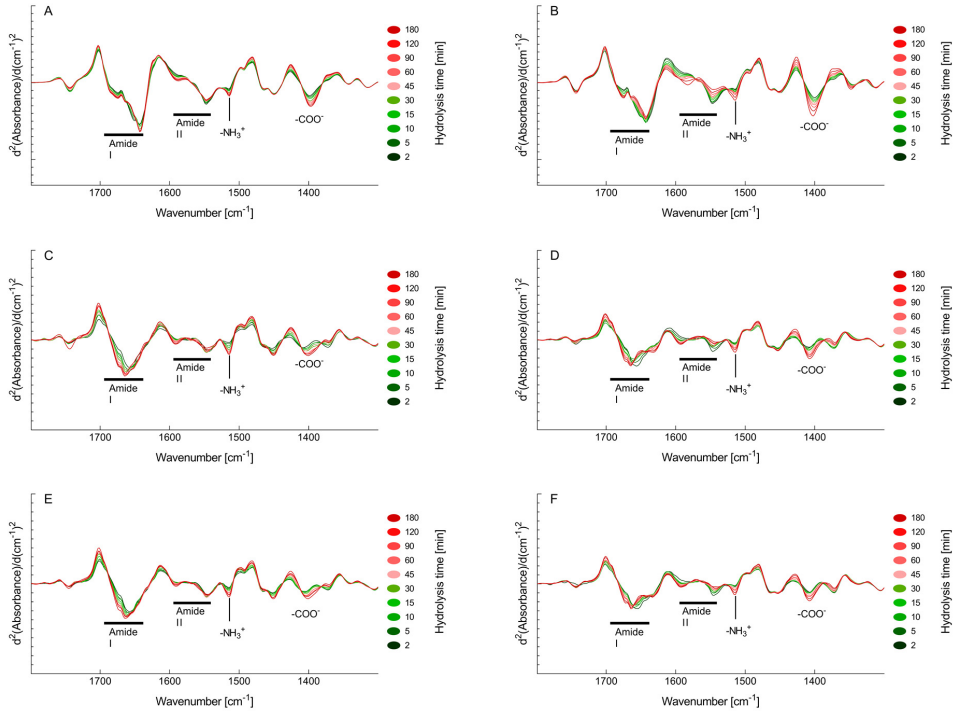
a) One of the measurements were removed due to injection error (n=5).




566

567 **Fig S-1:** All SEC chromatograms from 5-15 min of the six EPH time series studied: A)
568 WPC80 hydrolyzed with Alcalase. B) WPC80 hydrolyzed with Flavourzyme. C) WMP
569 hydrolyzed with Alcalase. D) WMP hydrolyzed with Flavourzyme. E) MPO hydrolyzed with
570 Alcalase and F) MPO hydrolyzed with Flavourzyme.

571




572

573 **Fig S-2:** All second derivative FTIR spectra from 1800-1300 $cm^{-1}$ of the six EPH time series
574 studied: A) WPC80 hydrolyzed with Alcalase. B) WPC80 hydrolyzed with Flavourzyme. C)
575 WMP hydrolyzed with Alcalase. D) WMP hydrolyzed with Flavourzyme. E) MPO
576 hydrolyzed with Alcalase and F) MPO hydrolyzed with Flavourzyme.

577

# Paper III

# Fourier-transform infrared spectroscopy for protein hydrolysate characterisation using dry-films treated with trifluoroacetic acid

Kenneth Aase Kristoffersen[a,b*], Aart van Amerongen[c], Ulrike Böcker[a], Diana Lindberg[a], Sileshi Gizachew Wubshet[a], Heleen de Vogel-van den Bosch[c], Svein Jarle Horn[b] and Nils Kristian Afseth[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: Kenneth Aase Kristoffersen, kenneth.kristoffersen@nofima.no*

# Abstract

In this study, we explore the potential of using Fourier-transform infrared (FTIR) spectra of trifluoroacetate-protein complexes for monitoring enzymatic protein degradation and protein characterization in general. The idea of treating dry-films of proteins with trifluoroacetic acid (TFA) prior to FTIR analysis is based on the unique properties of TFA. By adding a large excess of TFA to protein hydrolysate samples, the possible protonation sites of the proteins and peptides will be saturated. In addition, TFA has a low boiling point when protonated, and it also has complex forming abilities. When forming TFA-treated dry-films of hydrolysates, the excess TFA acid will evaporate and the deprotonated acid ($CF_3COO^-$) will interact as a counter ion with the positive charges of the sample materials. In the study, spectral changes in TFA-treated dry-films of a pure protein as well as complex protein mixtures were compared to the FTIR fingerprints of untreated dry-films. The results show that time-dependent information related to enzymatic protein hydrolysis (EPH) and, consequently, on the characteristics of a protein hydrolysate can be obtained. With additional developments, FTIR on dry-films treated with TFA may be regarded as a potential future tool for analysis of all types of proteolytic reactions in the laboratory as well as in industry.

# 1. Introduction

Fourier-transform infrared spectroscopy is among the well-established methods for characterization of proteins and peptides. The repeated amino acid building blocks constructing the backbone of proteins and peptides give rise to multiple distinctive infrared absorption bands, containing both compositional and structural information. This inherent ability to monitor changes of proteins and peptides has consequently been widely utilized in the study of quality parameters linked to protein and peptide secondary structures. This includes parameters such as hydration and solvent effects, pH, peptide size distribution and degree of hydrolysis (DH%).[1-6] Currently, peptide size distribution and DH% are measured using laborious and time-consuming techniques. FTIR, on the other hand, represents a rapid alternative which is applicable in an industrial setting, and recent studies have shown that FTIR can be used to predict parameters such as weight-average molecular weight ($M_w$) derived from the peptide size destitution and DH% with high accuracy.[7-9]

In recent years, EPH has gained significant attention as a versatile processing technology of protein-rich raw materials. For instance, several value-added peptide products have been developed employing EPH on food-processing by-products. However, protein hydrolysates originating from different raw materials, e.g. by-products from poultry and fish production, and even different enzymes will display unique FTIR fingerprints.[7] The fingerprint differences are linked to the raw material composition and the proteolytic enzyme mode of action, giving rise to unique protein degradation patterns, observed in the FTIR spectra. The authors recently reported a generic FTIR-based method for monitoring the average molecular weights of proteins during enzymatic hydrolysis of by-products from the food industry.[7] In this study, the best predictions were obtained using a hierarchical regression approach. This method involved supervised classification of the FTIR spectra according to raw material

quality and enzyme used in the hydrolysis process, and subsequent local regression models tuned to specific enzyme-raw material combinations. In other words, if different raw material and enzyme combinations are used, separate models for each combination usually are required. This shows that the predictive ability of the FTIR model towards protein size distribution and DH% is indeed raw material specific, which has several practical implications. For instance, when the aim is to establish industrially feasible generic prediction models for protein degradation, larger spectral libraries will be needed. A natural question that arises from the previous study is therefore if there are ways to reduce the raw material-specific information in the FTIR spectra, while keeping the information related to protein and peptide size distributions.

TFA is frequently used for denaturation, precipitation and analysis of proteins and peptides in biological samples, and this acid is known for its ability to interact with and modify the structure of proteins and peptides.[10-12] TFA is often used in the synthesis and analysis of peptides and proteins and is therefore a common contaminant. This can be an issue in many applications e.g. when used in biological assays.[13,14] As a result of the extensive use of TFA in many techniques, multiple methods for removal of TFA exist ranging from acid treatment to chromatographic methods.[15,16] When TFA interacts with proteins and peptides, a complex is formed that can be dried into films.[10,17] Such trifluoroacetate-protein complexes have previously been characterized using FTIR spectroscopy.[11,18,19] One could thus anticipate that trifluoroacetate-protein complexes could be measured the same way as the standard dry-film approach frequently employed in FTIR spectroscopy. However, the use of dry-films treated with TFA for FTIR characterization of proteins and peptides has not been studied or evaluated in detail.

In the present study, the potential of using FTIR spectra of the trifluoroacetate-protein complex for gaining information related to enzymatic protein degradation and protein characterization is explored. FTIR measurements were performed on both TFA-treated and untreated dry-films obtained from different types of protein hydrolysates. Hydrolysis time-series were obtained of a pure protein and more complex raw materials such as by-products from chicken and turkey processing. A thorough evaluation of the approach and its possibilities, together with initial FTIR band assignments, was also performed. To the best of our knowledge this is the first time TFA is used for modifying protein samples for dry-film FTIR analysis in order to characterize protein hydrolysates during EPH reactions.

## 2. Results and discussion

A range of enzymatic protein hydrolysis reactions were carried out in order to study TFA-induced effects in the dry-film FTIR spectra of the hydrolysate products. The samples were prepared by hydrolysing bovine serum albumin (BSA) and poultry-based raw materials. An overview of the hydrolysate samples and all processing parameters is provided in Table 1. To confirm time-dependent enzymatic degradation, SDS-PAGE electrophoresis and size exclusion chromatography were carried out. These results are presented in the supporting information (SI) Fig. S-1 and Table S-1.

The idea of treating dry-films of proteins with TFA prior to FTIR analysis is based on the unique properties of TFA. By adding a large excess of TFA to protein hydrolysate samples, the possible protonation sites of the proteins and peptides will be saturated. This includes the oxygens of the secondary amides in the protein backbone.[20,21] The resulting protein denaturation and protonation will occur using any strong acid. However, TFA has a low

boiling point when protonated, and it also has complex forming abilities. When forming TFA-treated dry-films of hydrolysates, the excess TFA acid will evaporate and the deprotonated acid ($CF_3COO^-$) will interact as a counter ion with the positive charges of the sample material.[22,23] During an EPH reaction the ratio of N-terminals relative to peptide bonds (secondary amides) is increasing. This ratio will be proportional to the $CF_3COO^-$ complexes formed as a film after the excess TFA evaporation. Therefore, as the proteolytic reaction proceeds, systematic changes are expected in both the $CF_3COO^-$ and amide absorption bands of the denaturised proteins and peptides in the corresponding FTIR spectra.

### 2.1. Spectral changes in TFA-treated dry-films of BSA

Formation of TFA-treated dry-films of hydrolysates was optimized by investigating different concentrations of TFA and other solvent compositions. The simplest and best method was to deposit the sample on the well-plate and let it dry for a minimum of 30 minutes before FTIR measurements. After the first FTIR measurement, a TFA solution was added to the sample. The sample was gently mixed and allowed to dry for a minimum of 30 minutes and measured again.

The interpretation of FTIR spectra of proteins and peptides has been thoroughly studied and reported.[24] The second derivative FTIR spectrum of BSA hydrolysed with the enzyme MaxiPro AFP for 80 minutes is provided in Fig. 1. The spectrum is divided into regions marked *i*-*vii*, and a list with the specific bands and wavenumber values are presented Table 2. Some of these bands have been shown to be important for monitoring protein and peptide degradation. This includes bands such as the amide I and II, the N-terminal ($NH_3^+$) and the C-terminal ($COO^-$), but changes can also be observed in other IR bands. Fig. 1 also shows the FTIR spectrum of the same hydrolysate sample treated with TFA. The formation of the

trifluoroacetate-protein or peptide complexes is expected to significantly affect the FTIR spectrum and this was observed by changes in all the defined spectral regions *i-vii*. In the regions *i-iii*, for example, alterations are seen in all the bands listed above. When comparing the TFA treated spectrum to the untreated, original elements from amide I and II are preserved and the $COO^-$ stretches disappeared altogether due to protonation. New bands from the $CF_3COO^-$ of the trifluoroacetate-protein or peptide complexes are also clearly visible and the most dominant are the C=O (~1677 $cm^{-1}$), C-F (~1250-1100 $cm^{-1}$) and OCO (~950-700 $cm^{-1}$) stretching bands. Protonation of a secondary amides are known to affect the amide II IR band, as it change the C-N and N-H bond length of the corresponding amide group.[25] The C-N bond becomes shorter when the carbonyl oxygen is protonated.[20] This gives the C-N bond a stronger double bond nature and the nitrogen will have a positive charge which are stabilised by the $CF_3COO^-$ counter ion. The $CF_3COO^-$ interaction with secondary amide have been assigned to a band in the ~1620 $cm^{-1}$ region.[23]

Second derivative FTIR spectra of full time series of BSA hydrolysed with MaxiPro AFP are provided in Fig. 2. The untreated FTIR spectra (Fig. 2A), display features closely resembling previously published FTIR spectra of hydrolysates produced using the same protein.[24] By comparing Fig. 2A and 2B, differences can clearly be observed between the TFA-treated and untreated dry-film spectra. Most important, both time series shows systematic time-dependent changes. For instance, for both time series, although the spectral changes appearing are different, the C=O amide I is changing systematically with time. It is also important to note that the new bands appearing in the FTIR spectra of the TFA-treated spectra are also systematically changing with time. Examples can be seen by following the C=O (1677 $cm^{-1}$), $CF_3COO^-$ (1627 $cm^{-1}$), C–F (1205 $cm^{-1}$) and the OCO (802 $cm^{-1}$) bands. The bands assigned to the $CF_3COO^-$ acid group are expected to change as the ratio of N-terminal ends relative to

peptide bonds is increasing during the EPH reaction. Finally, the figure clearly reveals that for the major bands in the region 1100-900 $cm^{-1}$, the time-dependent behaviour of the spectra are similar for both untreated and TFA-treated samples. In this region, functional groups with lower polarity give rise to several bands (see Table 2). TFA treatment is therefore expected to provide minimal influence on these bands.

### 2.2. Spectral changes in TFA-treated dry-films of complex samples

The BSA hydrolysates are relatively simple as they originate from a pure protein. The protein is also fairly small and its structure is mostly composed of α-helixes and turns.[26] Time-series of samples obtained from poultry-based raw materials were therefore produced to study the effect of TFA on more complex structures, before and after TFA treatment. Fig. 3A and 3C show the second derivative FTIR spectra of two EPH time-series for poultry substrates. The spectra from the two time-series are visually different from one another and the largest differences are seen in regions *i-iii* where secondary structure attributes are located. This region also contains important IR bands for EPH monitoring e.g. amide I and II bands, the N-terminal ($NH_3^+$) and the C-terminal ($COO^-$). The complexity difference seen in the spectra derives from both specific raw materials and differences in the mode of action of the specific protease. For example, turkey raw material contains more collagen than the chicken raw material. Also, Flavourzyme (an exopeptidase) is less efficient when it comes to degradation of both muscle proteins and collagens as compared to Alcalase (with mostly endopeptidase activity).[27] All in all, differences like the ones shown in Fig 3A and 3C are typical for hydrolysates originating from different raw materials and enzymes. The spectra of these samples treated with TFA are shown in Fig. 3B and 3D. Compared to the untreated spectra the differences between the spectra of the TFA-treated spectra are dramatically reduced (i.e., Fig. 3A vs. 3C as compared to Fig. 3B vs. 3D) and the time-series become more similar to

one another. The new $CF_3COO^-$ bands are also directly comparable to the ones described for the TFA-treated BSA spectra in Fig. 2.

The TFA-treated spectra in Fig. 3 show four new bands, i.e. C=O, $CF_3COO^-$ (counter ion effect with secondary amide) C–F and OCO, that seem to change more with hydrolysis time than the others. In order to investigate if these changes contained quantitative information related to the EPH reactions, principal component analysis (PCA) was applied to the FTIR spectra of the EPH time-series. This was done following the description given by Böcker et al.[24] In that study, first component (PC1) scores of each raw material enzyme time-series were seen to depict the change in FTIR signature as a function of hydrolysis time, and it was concluded that the spectral fingerprint of the product is likely to be strongly connected to the composition and homogeneity of the raw material. In the present study each raw material-enzyme group consisting of two time-series was subjected to PCA analysis. The PC1 explained 73-93% of the variance for the six untreated groups. The explained variances (PC1) for all groups are presented in SI Table S-2. For the TFA-treated samples, 53-70% variance was explained by PC1. Thus, the explained variance was slightly lower for the TFA-treated samples. When plotting the scores of the first component against time as shown in Fig. 4, it can be seen that time-dependent EPH information is present in the spectra time-series of both TFA-treated and untreated samples.

From the PC1 loadings (Fig. 5) for the different EPH time-series in Fig. 4 it is clear that the FTIR bands important to explain the variation across PC1 changed upon addition of TFA. Fig. 5A (i.e. untreated samples) shows that the PC1 loadings of the selected raw materials-enzyme combinations are very different in the amide region. Here, information concerning protein secondary structure attributes is located. In the PC1 loadings for the TFA-treated

samples (Fig. 5B), on the other hand, one can see that even though there are clear intensity differences in the amide region between the two loadings, the shape of the loadings is very similar. It is also clear that three new peaks related to TFA interactions are important for the spectral variation. This includes C=O (~1677 $cm^{-1}$), $CF_3COO^-$ (counter ion effect with secondary amide), and the C–F (~1205 $cm^{-1}$) bands. Overall, these differences in the loadings indicate that whereas untreated FTIR samples reveal protein degradation patterns related to raw-material specific changes in secondary structures, TFA-treated samples reveal protein degradation patterns related to a more generic $CF_3COO^-$ counter ion effect with secondary amide relative to N-terminal interactions.

### 2.3. General discussion

The results of this study illustrate the potential of treating dry-films of protein hydrolysates with TFA to reduce spectral differences in FTIR spectra originating from secondary protein/peptide structures in protein hydrolysis samples from different origins. Currently, these spectral differences are the major factor responsible for the need of large data libraries when samples of different raw material origin are used for prediction of quality parameters such as average molecular weights from FTIR spectra.[7] This study also shows that information important for monitoring EPH reactions is preserved in the FTIR spectra of the TFA-treated samples. Thus, it is shown that FTIR spectra of TFA-treated samples contain semi-quantitative information related to protein degradation. The next step is then to show that FTIR spectra of TFA-treated samples can be used quantitatively.

A concern with the TFA treatment is that essential information in the amide I band would be lost under the C=O stretching band from $CF_3COO^-$.[7,18,28] This is not an uncommon issue in IR spectroscopy. For instance, it can be seen in a study by Poulsen et al. predicting DH% in

aqueous whey protein hydrolysate samples. Here, the water peak overlaps the amide I band and they concluded that information for monitoring the EPH reaction was partially lost.[29] The difference to this study is that the $CF_3COO^-$ peaks overlapping with the amide I bands is directly linked to the progression of the EPH reaction, as more and more N-terminals are formed. This strengthens the argument that TFA-induced reduction of spectral differences can be utilized for EPH monitoring applications, especially in processes where a variety of raw materials are included.

There are some challenges that have to be overcome in order to fully take advantage of the spectral changes induced by TFA. One of these challenges is dry-film quality. From our observations, the TFA-treated films have a less uniform surface as compared to the untreated ones. This can result in more noise in the FTIR spectra, which in turn can reduce the amount of information available in the spectra. From figures. 2 and 3 the amount of noise seems to be generally higher in the treated spectra compared to the untreated. To overcome this challenge, we will address several possible solutions to get uniform surfaces of TFA-treated samples in future studies. From printing biomolecules on surfaces for diagnostic applications it is well-known that an appropriate buffer composition and the addition of a surfactant/detergent to the sample may result in uniform surfaces.[30,31] Changing the hydrophobicity of the Si-plate or the relative humidity upon drying of the samples on the plate may be other ways to solve this issue.[32,33]

When performing dry-film FTIR spectroscopy, it is generally known that the film-thickness cannot be fully reproduced. Thicker films will lead to higher overall absorbance intensities than thinner films, and manual pipetting of small sample aliquots is very difficult to reproduce. Usually, some kind of normalization procedure (like EMSC used in the present

study) is used to account for this effect.[34] However, if the spectra are very similar, the normalization procedure will lead to correction of more than the thickness effect. Since TFA treatment could provide such similar spectra, this might be a challenge when quantitative studies are being made. Using internal bands as internal standards, or even adding an external standard, could be possible solutions to this challenge. Alternatively, solutions proposed above to improve spot uniformity may result in a more reproducible spot thickness.

EPH is a growing industrial segment related to valorisation of protein-rich co-streams from for example, the meat and fishery industry. For these sectors, there is a need for simple and fast industrially applicable tools for monitoring product quality development during processing to ensure a constant quality. The present study clearly shows that TFA-treated dry-film FTIR has the potential to simplify the spectra by reducing spectral differences. Consequently, large datasets to build robust quality prediction models based on FTIR spectroscopy will not be necessary anymore. Thus, the technique is a potential future tool for analysis of all types of proteolytic reactions in the laboratory as well as in the industry.

# 3. Conclusion

The present study shows the potential of using dry-films treated with TFA for monitoring changes in proteins and peptides during proteolytic reactions. When TFA interacts with proteins and peptides, a complex is formed that can be dried into films. In the study it is shown that this complex contains information important for monitoring EPH reactions. Compared to untreated dry-films, where protein degradation patterns are related to raw-material specific changes in secondary structures, TFA-treated samples reveal protein degradation patterns related to a more generic $CF_3COO^-$ counter ion effect. There are, however, challenges concerning the dry-film quality and pre-processing of FTIR spectra that have to be

investigated and improved in order to fully take advantage of the TFA interactions. When these challenges are solved, FTIR of TFA-treated EPH samples may enable the use of simpler and more generic models for the prediction of quality parameters of proteolytic reactions in laboratory as well as industrial applications.

# 4. Materials and methods

## 4.1. Materials

Protease from *Bacillus licheniformis* (Alcalase, 2.4 U/g), *Aspergillus oryzae* (Flavourzyme), bovine serum albumin pH 5.2, ≥96% and trifluoroacetic acid were purchased from Sigma-Aldrich (St. Louis, MO, USA). Corolase 2TS was provided by AB enzyme (Darmstadt, Germany) and MaxiPro AFP, by DSM (Delft, the Netherlands). Millex-HV PVDF 0.45 µm 33 mm filters were used for sample preparation (MilliporeSigma, Burlington, MA, USA).

## 4.2. Raw materials

Raw materials derived from chicken and turkey were hydrolysed by a selection of commercially available enzymes. The poultry raw materials, turkey mechanical deboning residue (TMDR) and Chicken carcass (CC) were supplied from a Norwegian slaughterhouse. All samples were minced and packed in bags shortly after arrival, before being stored at -20 °C.

## 4.3. Enzymatic hydrolysis and sampling

Hydrolysis samples produced using raw materials from poultry were hydrolysed according to the description by Wubshet et al.[8] The BSA hydrolysates were produced by dissolving 5%

306 (w/v) substrate of BSA in tap water. In a screw-cap glass bottle with a magnetic stirrer, 100
307 mL of the substrate solution was heated to 50 °C in a water bath. When the desired
308 temperature was attained, 4% (v/w) Maxipro AFP was added to start the enzymatic reaction.
309 From the reaction mixture, 5 mL of sample solution was collected at 0.5 min, 2.5 min, 5 min,
310 7.5 min, 10 min, 15 min, 20 min, 30 min, 40 min, 50 min, 60 min, and 80 min in 15 mL Falcon
311 tubes. The samples were immediately heated to 90 °C for 10 min to inactivate the enzyme.
312 After cooling to room temperature, the samples were centrifuged (4400 rpm, 15 min) and the
313 clear water phase was collected for freeze drying. The dried samples were dissolved in water
314 to a concentration of 25 mg/mL and filtrated through Millex-HV PVDF (MilliporeSigma,
315 Burlington, MA, USA) before being subjected to spectroscopic analysis. The samples were
316 analysed to verify that the proteins were sufficiently broken down during the EPH reactions
317 and an overview of the samples are given in Table 1.
318

319 **4.4. FTIR spectroscopy and TFA treatment**

320 From each of the filtered protein hydrolysates, aliquots (7.5 µL) were deposited on 96-well
321 IR-transparent Si-plates and dried at room temperature for at least 30 minutes to form dry
322 films as described by Böcker et al.[24] From each hydrolysate sample, five aliquots were
323 deposited to allow for replicate measurements. FTIR measurements were performed using a
324 High Throughput Screening eXTension (HTS-XT) unit coupled to a Tensor 27 spectrometer
325 (Bruker, Billerica, MA, USA). The spectra were recorded in the region between 4000 and 400
326 $cm^{-1}$ with a spectral resolution of 4 $cm^{-1}$ and an aperture of 5.0 mm. For each spectrum, 40
327 interferograms were collected and averaged. Data acquisition was controlled using Opus v6.5
328 (Bruker, Billerica, MA, USA). After the first measurements a large excess of TFA (10 µL,
329 0.25M TFA in water) was deposited onto each sample well and the slurry was gently stirred
330 with the pipette tip. The plates were then once again dried at room temperature for at least 30

minutes to allow the excess of TFA to evaporate and form dry films before being subjected to FTIR measurements ones again.

### 4.5. Data treatment

Pre-processing of FTIR spectra from 1800 to 400 $cm^{-1}$ was performed using Savitzky-Golay 2$^{nd}$ derivative (window size 13 points) followed by extended multiplicative signal correction (EMSC) with 2$^{nd}$ order polynomial and using the mean spectrum as reference. This pre-processing approach enabled removal of major scattering effects in the spectra while at the same time removing the effect of varying thickness of the dry-films.[34] For all subsequent data analysis, the region from 1800 $cm^{-1}$ to 400 $cm^{-1}$ was used. Principal component analysis (PCA) was used for data exploration. Data processing and analysis were carried out using The Unscrambler® X v10.3 (CAMO Software AS, Oslo, Norway).

## Acknowledgements

Financial support from the Norwegian Fund for Research Fees for Agricultural Products through the projects “FoodSMaCK” (no. 262308) and “SunnMat” (no. 262300), and from the Norwegian Research Council through the project “Notably” (no. 280709) is greatly acknowledged. Internal funding from Nofima through the project “PepTek” is also acknowledged.

# References


1 Barth, A. The infrared absorption of amino acid side chains. *Progress in Biophysics and Molecular Biology* **74**, 141-173, doi:https://doi.org/10.1016/S0079-6107(00)00021-3 (2000).

2 Perisic, N., Afseth, N. K., Ofstad, R. & Kohler, A. Monitoring Protein Structural Changes and Hydration in Bovine Meat Tissue Due to Salt Substitutes by Fourier Transform Infrared (FTIR) Microspectroscopy. *Journal of Agricultural and Food Chemistry* **59**, 10052-10061, doi:10.1021/jf201578b (2011).

3 Andersen, P. V., Veiseth-Kent, E. & Wold, J. P. Analyzing pH-induced changes in a myofibril model system with vibrational and fluorescence spectroscopy. *Meat Science* **125**, 1-9, doi:10.1016/j.meatsci.2016.11.005 (2017).

4 Arabi, S. H. *et al.* Serum albumin hydrogels in broad pH and temperature ranges: characterization of their self-assembled structures and nanoscopic and macroscopic properties. *Biomaterials science* **6**, 478-492, doi:10.1039/c7bm00820a (2018).

5 Okabayashi, H.-F., Kanbe, H.-H. & O'Connor, C. J. The role of an l-leucine residue on the conformations of glycyl-l-leucine oligomers and its N- or C-terminal dependence: infrared absorption and Raman scattering studies. *European Biophysics Journal* **45**, 23-34, doi:10.1007/s00249-015-1072-3 (2016).

6 Martra, G. *et al.* The Formation and Self-Assembly of Long Prebiotic Oligomers Produced by the Condensation of Unactivated Amino Acids on Oxide Surfaces. *Angewandte Chemie-International Edition* **53**, 4671-4674, doi:10.1002/anie.201311089 (2014).

7 Kristoffersen, K. A. *et al.* FTIR-based hierarchical modeling for prediction of average molecular weights of protein hydrolysates. *Talanta* **205**, 120084, doi:https://doi.org/10.1016/j.talanta.2019.06.084 (2019).

8 Wubshet, S. G. *et al.* FTIR as a rapid tool for monitoring molecular weight distribution during enzymatic protein hydrolysis of food processing by-products. *Analytical Methods* **9**, 4247-4254, doi:10.1039/C7AY00865A (2017).

9 Wubshet, S. G. *et al.* in *Proteins: Sustainable Source, Processing and Applications* (ed Charis M. Galanakis) 225-258 (Academic Press, 2019).

10 Rajkhowa, R. *et al.* Structure and properties of biomedical films prepared from aqueous and acidic silk fibroin solutions. *Journal of Biomedical Materials Research Part A* **97A**, 37-45, doi:10.1002/jbm.a.33021 (2011).

11 Valenti, L. E., Paci, M. B., De Pauli, C. P. & Giacomelli, C. E. Infrared study of trifluoroacetic acid unpurified synthetic peptides in aqueous solution: Trifluoroacetic acid removal and band assignment. *Analytical Biochemistry* **410**, 118-123, doi:https://doi.org/10.1016/j.ab.2010.11.006 (2011).

12 Ahmad, A., Madhusudanan, K. P. & Bhakuni, V. Trichloroacetic acid and trifluoroacetic acid-induced unfolding of cytochrome c: stabilization of a native-like folded intermediate11CDRI communication number 5877. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1480**, 201-210, doi:https://doi.org/10.1016/S0167-4838(00)00069-8 (2000).

13 Ma, T. G., Ling, Y. H., McClure, G. D. & Tseng, M. T. Effects of trifluoroacetic acid, a halothane metabolite, on C6 glioma cells. *Journal of Toxicology and Environmental Health* **31**, 147-158 (1990).

14 Tipps, M. E., Iyer, S. V. & Mihic, S. J. Trifluoroacetate is an allosteric modulator with selective actions at the glycine receptor. *Neuropharmacology* **63**, 368-373, doi:10.1016/j.neuropharm.2012.04.011 (2012).

15 Andrushchenko, V. V., Vogel, H. J. & Prenner, E. J. Optimization of the hydrochloric acid concentration used for trifluoroacetate removal from synthetic peptides. *Journal of Peptide Science* **13**, 37-43, doi:10.1002/psc.793 (2007).
16 Roux, S. *et al.* Elimination and exchange of trifluoroacetate counter-ion from cationic peptides: a critical evaluation of different approaches. *Journal of Peptide Science* **14**, 354-359, doi:10.1002/psc.951 (2008).
17 Åsberg, D. *et al.* The importance of ion-pairing in peptide purification by reversed-phase liquid chromatography. *Journal of Chromatography A* **1496**, 80-91, doi:https://doi.org/10.1016/j.chroma.2017.03.041 (2017).
18 Haris, P. I. & Chapman, D. The conformational-analysis of peptides using Fourier-transform IR spectroscopy. *Biopolymers* **37**, 251-263, doi:10.1002/bip.360370404 (1995).
19 Pastrana-Rios, B., Sosa, L. D. & Santiago, J. Trifluoroacetic acid as excipient destabilizes melittin causing the selective aggregation of melittin within the centrin-melittin-trifluoroacetic acid complex. *Structural Dynamics-Us* **2**, doi:10.1063/1.4921219 (2015).
20 Murthy, A. S. N., Rao, K. G. & Rao, C. N. R. Molecular orbital study of the configuration protonation, and hydrogen bonding of secondary amides. *Journal of the American Chemical Society* **92**, 3544-3548, doi:10.1021/ja00715a003 (1970).
21 Sudmeier, J. L. & Schwartz, K. E. Protonation of carbonyl groups in peptides and amino-acids. *Chemical Communications (London)*, 1646-1648, doi:10.1039/C19680001646 (1968).
22 Klotz, I. M., Russo, S. F., Hanlon, S. & Stake, M. A. Protonation of Amides in a Helix-Breaking Solvent. *Journal of the American Chemical Society* **86**, 4774-4778, doi:10.1021/ja01076a010 (1964).
23 Tam, J. W. O. & Klotz, I. M. Protonation of amides by trifluoroacetic acid: infrared and nuclear magnetic resonance studies. *Spectrochimica Acta Part A: Molecular Spectroscopy* **29**, 633-644, doi:https://doi.org/10.1016/0584-8539(73)80093-5 (1973).
24 Böcker, U., Wubshet, S. G., Lindberg, D. & Afseth, N. K. Fourier-transform infrared spectroscopy for characterization of protein chain reductions in enzymatic reactions. *Analyst* **142**, 2812-2818, doi:10.1039/C7AN00488E (2017).
25 Wu, R. & McMahon, T. B. Protonation Sites and Conformations of Peptides of Glycine (Gly1−5H+) by IRMPD Spectroscopy. *The Journal of Physical Chemistry B* **113**, 8767-8775, doi:10.1021/jp811468q (2009).
26 Carter, D. C. & Ho, J. X. in *Advances in Protein Chemistry* Vol. 45 (eds C. B. Anfinsen, John T. Edsall, Frederic M. Richards, & David S. Eisenberg) 153-203 (Academic Press, 1994).
27 Merz, M. *et al.* Characterization of commercially available peptidases in respect of the production of protein hydrolysates with defined compositions using a three-step methodology. *Journal of Molecular Catalysis B: Enzymatic* **127**, 1-10, doi:https://doi.org/10.1016/j.molcatb.2016.02.002 (2016).
28 Surewicz, W. K., Mantsch, H. H. & Chapman, D. Determination of protein secondary structure by Fourier-transform infrared-spectroscopy: A critical assessment. *Biochemistry* **32**, 389-394, doi:10.1021/bi00053a001 (1993).
29 Poulsen, N. A. *et al.* Predicting hydrolysis of whey protein by mid-infrared spectroscopy. *International Dairy Journal* **61**, 44-50, doi:https://doi.org/10.1016/j.idairyj.2016.04.002 (2016).

30 Mujawar, L. H., van Amerongen, A. & Norde, W. Influence of buffer composition on the distribution of inkjet printed protein molecules and the resulting spot morphology. *Talanta* **98**, 1-6, doi:10.1016/j.talanta.2012.06.006 (2012).
31 Mujawar, L. H., van Amerongen, A. & Norde, W. Influence of Pluronic F127 on the distribution and functionality of inkjet-printed biomolecules in porous nitrocellulose substrates. *Talanta* **131**, 541-547, doi:10.1016/j.talanta.2014.08.001 (2015).
32 Mujawar, L. H., Norde, W. & van Amerongen, A. Spot morphology of non-contact printed protein molecules on non-porous substrates with a range of hydrophobicities. *Analyst* **138**, 518-524, doi:10.1039/c2an36104c (2013).
33 Mujawar, L. H., Kuerten, J. G. M., Siregar, D. P., van Amerongen, A. & Norde, W. Influence of the relative humidity on the morphology of inkjet printed spots of IgG on a non-porous substrate. *RSC Advances* **4**, 19380-19388, doi:10.1039/c4ra01327a (2014).
34 Afseth, N. K. & Kohler, A. Extended multiplicative signal correction in vibrational spectroscopy, a tutorial. *Chemometrics and Intelligent Laboratory Systems* **117**, 92-99, doi:https://doi.org/10.1016/j.chemolab.2012.03.004 (2012).

# Tables

**Table 1.** An overview of samples and hydrolysis reaction conditions.

| Sample name[a] | Enzyme[b] | Enzyme loading (%)[c] | Reaction time (min) | Water (mL)[d] | Samples per hydrolysis[e] | Raw material (g)[f] | Number of samples |
|---|---|---|---|---|---|---|---|
| BSAM | MaxiPro AFP | 4[f] | 80 | 100 | 12 | 5 | 24 |
| CCA | Alcalase | 1[c] | 80 | 1000 | 12 | 500 | 24 |
| CCC | Corolase 2TS | 1[c] | 80 | 1000 | 12 | 500 | 24 |
| CCF | Flavourzyme | 1[c] | 80 | 1000 | 12 | 500 | 24 |
| TMDRA | Alcalase | 1[c] | 80 | 1000 | 12 | 500 | 24 |
| TMDRC | Corolase 2TS | 1[c] | 80 | 1000 | 12 | 500 | 24 |
| TMDRF | Flavourzyme | 1[c] | 80 | 1000 | 12 | 500 | 24 |
| Total | | | | | | | 168 |

a) Turkey mechanical deboning residue (TMDR) and Chicken carcass (CC), abbreviation for the enzymes used are added to the sample name. b) Alcalase (A), Flavourzyme (F), MaxiPro AFP (M), and Corolase 2TS (C). c) Enzyme loading to raw material. d) Water added to reaction mixture. e) Number of sampling time points for each reaction and f) raw material loading.

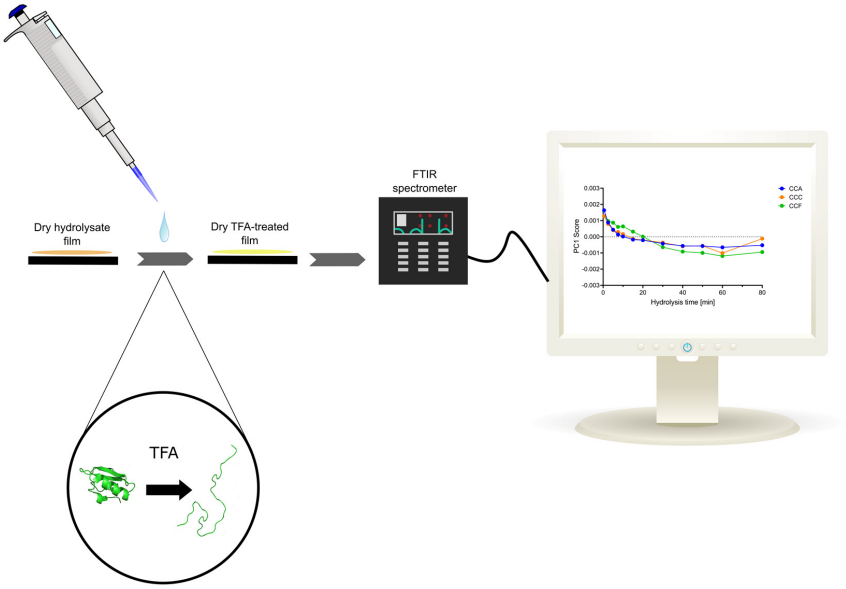
473 **Table 2.** 2nd derivative bands between 1800-400 $cm^{-1}$ for BSA samples.

| | | Band position $cm^{-1}$ | |
|---|---|---|---|
| Annotations | Region | BSA | TFA-treated BSA |
| C=O amide I: turns | *i* | 1687 | - |
| C=O (TFA) | | - | 1677 |
| C=O amide I: α-helix | | 1656 | 1656 |
| $CF_3COO^-$ | | - | 1627 |
| $COO^-$ (asym stretch) | *ii* | 1587 | - |
| Amide II: α-helix | | 1548 | 1548 |
| $NH_3^+$ (scissor) | *iii* | 1516 | 1517 |
| $CF_3COO^-$ | | - | 1471 |
| $CH_2$ (scissor) | | 1454 | 1454 |
| $COO^-$ (sym stretch) | *iv* | 1404 | - |
| Amide III, CH2 (def, rock), OH (def, bend) | | 1317 | - |
| Amide III, C–O (stretch) | | 1240 | - |
| C–F (TFA) | *v* | - | 1205 |
| C–F (TFA) | | - | 1182 |
| C–F (TFA) | | - | 1139 |
| $CNH_3$ (rock), $CH_2$ (wag) | *vi* | 1112 | 1116 |
| CO, CC, CN (stretch) | | 1045 | 1143 |
| CCOO (wagging) | | 995 | 995 |
| $CH_2$ (twist) | | 927 | 925 |
| C–C (TFA) | *vii* | - | 840 |
| OCO (TFA) | | - | 802 |
| OCO (TFA) | | - | 723 |

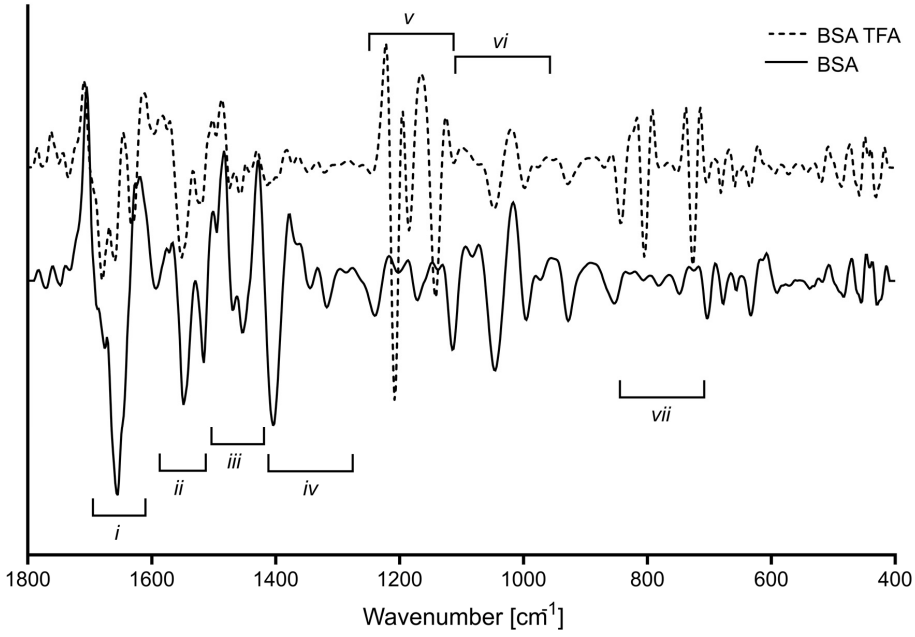
474

475

# Figures




**Graphical abstract**

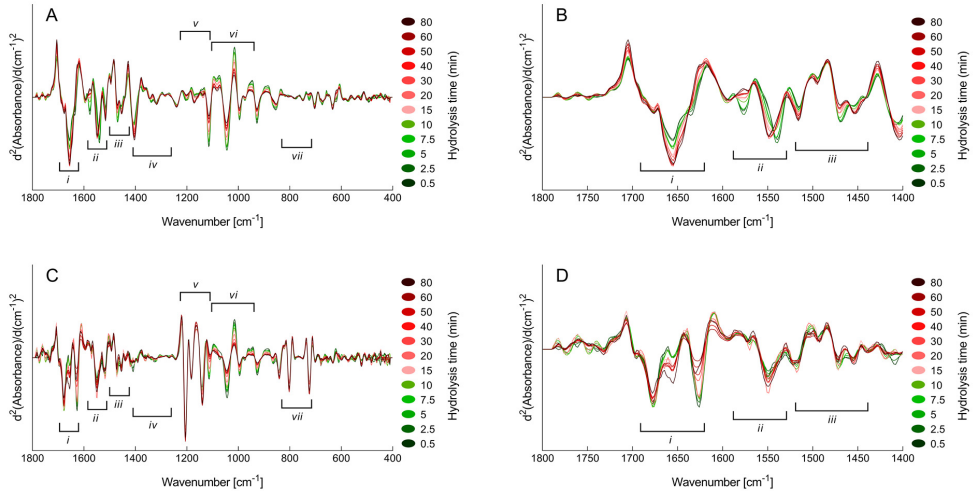



481

482 **Fig. 1**: TFA-treated dry-film spectra compared to non-treated spectra of BSA hydrolysed with

483 MaxiPro AFP for 80 minutes. Wavenumbers for assigned FTIR bands are listed in Table 2.

484




485

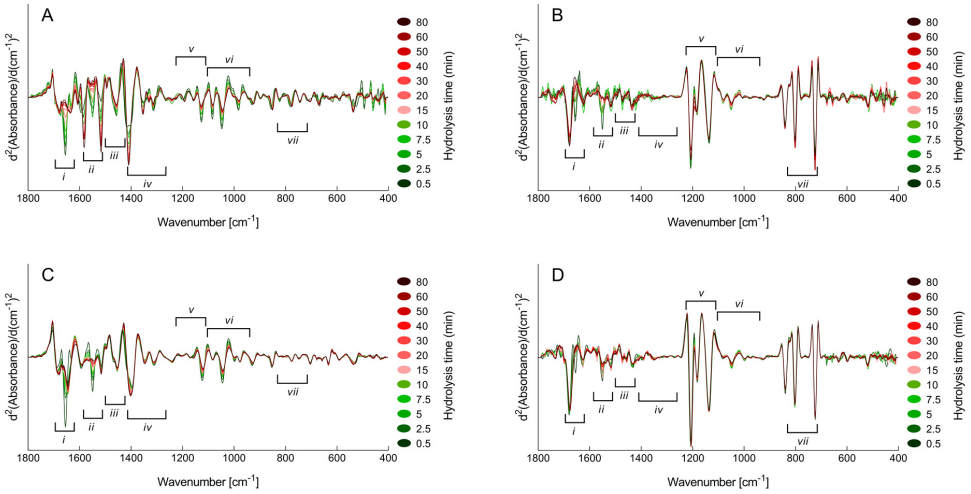
486 **Fig. 2**: Second derivative FTIR spectra of BSA hydrolysis time-series: A) The spectral region

487 1800-400 $cm^{-1}$ of the untreated dry-films. B) The spectral region 1800-1400 $cm^{-1}$ of the

488 untreated dry-films. C) The spectral region 1800-400 $cm^{-1}$ of the TFA-treated dry-films. D)

489 The spectral region 1800-1400 $cm^{-1}$ of the TFA-treated dry-films.

490




491

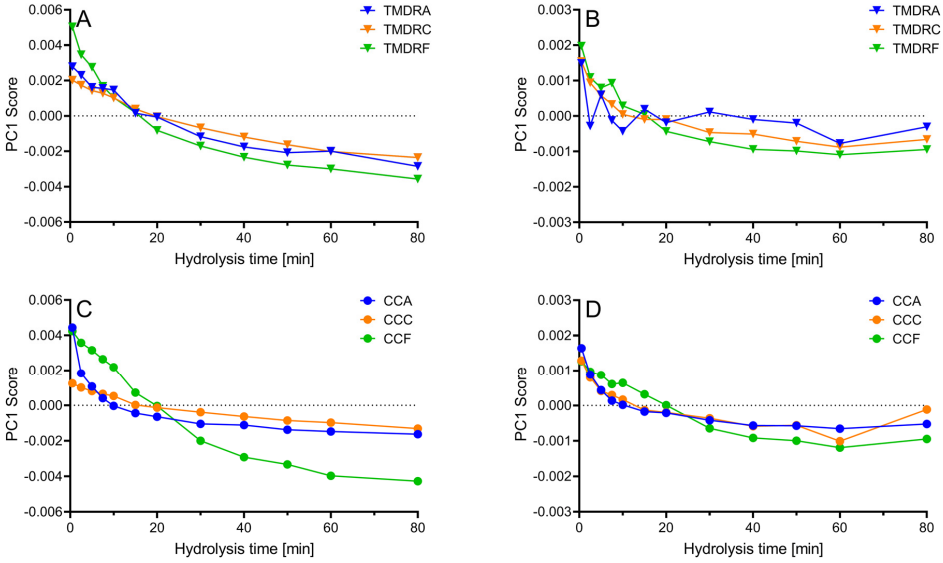
492 **Fig.3**: Second derivative FTIR spectra from 1800 to 400 $cm^{-1}$ of two different hydrolysis time-

493 series, one treated with TFA and one untreated: A and B) Turkey mechanically deboned

494 residue hydrolysed with Flavourzyme. C and D) Chicken carcass hydrolysed with Alcalase.

495 A and C) Untreated dry-film spectra. B and D) TFA-treated dry-film spectra.

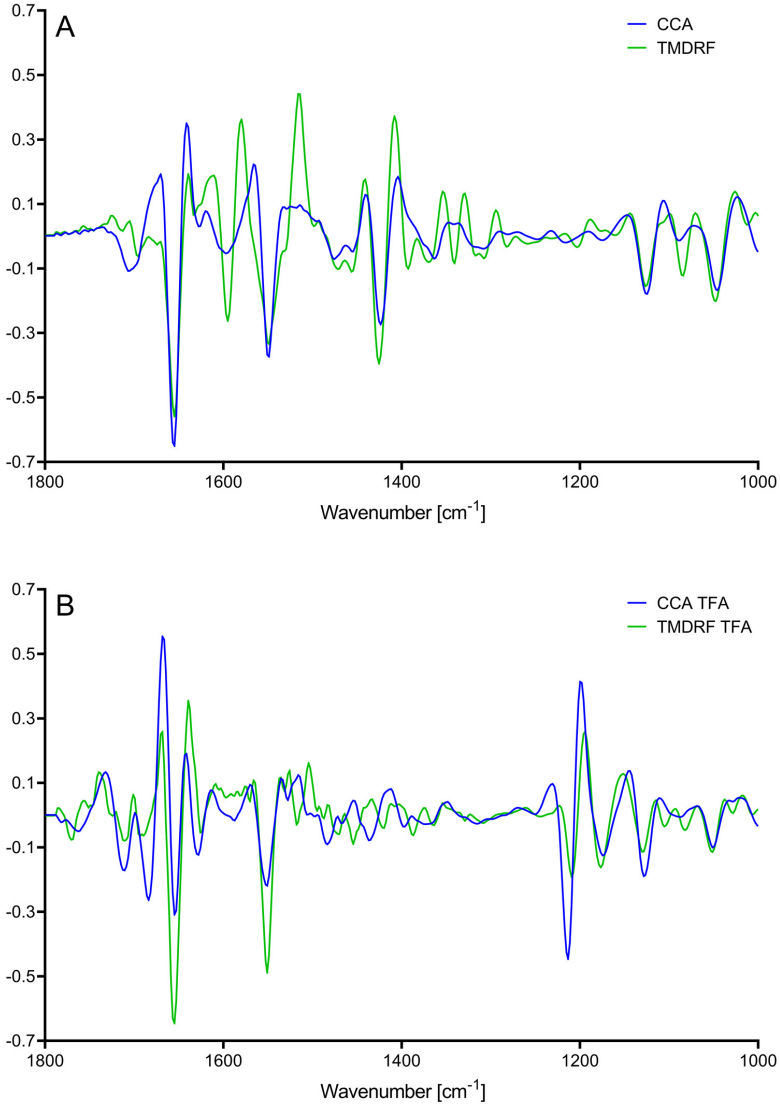
496




497

498 **Fig. 4**: Time-dependent development of PC1 of FTIR spectra during EPH of poultry samples,
499 TFA-treated vs. untreated samples. Sample names and enzyme abbreviations are defined in
500 Table 1. The data points are the average of two time-series. A and B) Turkey mechanically
501 deboned residue. C and D) Chicken carcass. A and C) Untreated FTIR spectra. B and D) TFA-
502 treated FTIR spectra.

503




504

505 **Fig. 5**: Range-normalized PC1 loadings from 1800-1000 $cm^{-1}$ for the EPH time-series
506 presented in Fig. 3, TFA-treated samples vs. untreated. A) Untreated samples. B) TFA-treated
507 samples.

508

509

# Supporting Information

# Fourier-transform infrared spectroscopy for protein hydrolysate characterization using dry-films treated with trifluoroacetic acid

Kenneth Aase Kristoffersen[a,b*], Aart van Amerongen[c], Ulrike Böcker[a], Diana Lindberg[a], Sileshi Gizachew Wubshet[a], Heleen de Vogel-van den Bosch[c], Svein Jarle Horn[b] and Nils Kristian Afseth[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: Kenneth Aase Kristoffersen, kenneth.kristoffersen@nofima.no*

This document contains the results of SDS-PAGE electrophoresis analysis of a selection of BSA time-series samples, the weight-average molecular weight of all the poultry samples studied and explained variance in PC1 from the PCA analysis.

For the SDS PAGE (Fig. S-1), samples representing a time series of the BSA hydrolysates (2.5, 7.5, 15, 30, 50, 80 min, all 25 mg/ml) was diluted 1:1 with a 2x sample buffer (0.125 M Tris pH 6.8, 4% SDS, 20% glycerol, 0.2 M DTT, 0.04% bromophenol blue) before heating at 10 min at 50ºC. After cooling, 10 µl of the samples was added to an Invitrogen NuPage Mini 10% Bis-Tris gel using a MOPS running buffer during electrophoresis, performed according to the manufacturer protocol using the XCell *SureLock* Mini-Cell system. The gels were stained with a Coomassie staining solution (1% Coomassie Brilliant Blue R, 40% MeOH, 7% acetic acid) for one hour, and destained overnight (40% MeOH, 7% acetic acid in distilled water). The weight-average molecular weights (Table S-1) were measured using the method described by Wubshet et al.[1]

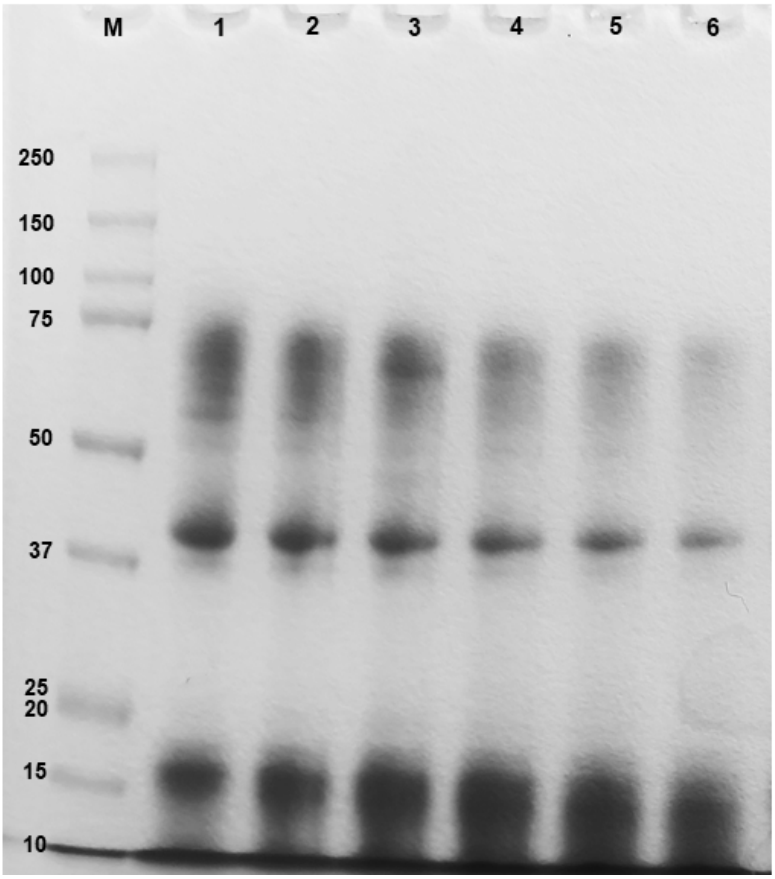



**Fig S-1**: The resulting SDS-PAGE gel. In lane M, the marker with the molecular weights defined in the BioRad Dual color prestained protein ladder product sheet, and to the left in the figure. In lane 1 to 6, the hydrolysates from 2.5, 7.5, 15, 30, 50, 80 min.

545 **Table S-1:** Weight-average molecular weight ($M_w$, g/mol) values for all poultry EPH samples.

| Time (min)[a] | 1) CCA | 1) CCC | 1) CCF | 1) TMDRA | 1) TMDRC | 1) TMDRF |
|---|---|---|---|---|---|---|
| **0.5** | 3246.3 | 2558.5 | 1617.1 | 3972.8 | 2595.9 | 3845.7 |
| **2.5** | 2796.8 | 2463.9 | 1653.5 | 5512.9 | 2780.4 | 3083.7 |
| **5** | 2714.1 | 2608.1 | 1821.6 | 3104.3 | 2816.5 | 6493.1 |
| **7.5** | 2702.4 | 2548.6 | 1621.1 | 5076.7 | 2706.8 | 2714.4 |
| **10** | 2624.4 | 2440.0 | 2629.5 | 2580.5 | 2845.3 | 2583.1 |
| **15** | 2288.1 | 2383.2 | 1557.8 | 4692.5 | 2658.2 | 4204.7 |
| **20** | 2208.3 | 2226.4 | 2669.8 | 2441.4 | 3001.2 | 2253.9 |
| **30** | 2042.7 | 2256.6 | 1381.6 | 2288.8 | 2649.1 | 2354.8 |
| **40** | 2038.5 | 2392.9 | 1310.6 | 2353.5 | 2649.1 | 1971.0 |
| **50** | 1872.3 | 2017.6 | 1247.8 | 2336.0 | 2366.5 | 1857.7 |
| **60** | 1889.7 | 2000.8 | 1206.2 | 2936.2 | 2383.3 | 1828.6 |
| **80** | 1742.7 | 1902.4 | 1139.7 | 1967.3 | 2320.2 | 1628.6 |
| **Time (min)[b]** | 2) CCA | 2) CCC | 2) CCF | 2) TMDRA | 2) TMDRC | 2) TMDRF |
| **0.5** | 5352.6 | 2394.3 | 1726.7 | 3940.4 | 2611.8 | 4806.0 |
| **2.5** | 2507.5 | 2266.1 | 1909.3 | 4968.4 | 2580.5 | 4129.6 |
| **5** | 2439.4 | 2142.5 | 1921.0 | 4175.9 | 3213.2 | 4574.5 |
| **7.5** | 3290.0 | 2563.8 | 1828.2 | 4897.5 | 2941.3 | 2944.2 |
| **10** | 2232.0 | 2220.6 | 1921.7 | 2503.3 | 2622.3 | 2514.2 |
| **15** | 2125.5 | 2566.7 | 2680.8 | 2291.3 | 2498.5 | 2365.3 |
| **20** | 1949.7 | 2007 | 3345.1 | 3966.9 | 2317.8 | 2201.5 |
| **30** | 1818.9 | 2175.1 | 1629.3 | 2131.1 | 2286.7 | 2065.7 |
| **40** | 2154.6 | 2173.7 | 1553.5 | 1977.1 | 2398.2 | 1934.8 |
| **50** | 1642.4 | 2169.5 | 1483.5 | 1881.7 | 2365.5 | 1810.4 |
| **60** | 1578.6 | 2342.7 | 1379.7 | 1823.1 | 2259.3 | 1757.7 |
| **80** | 1572.7 | 2016.6 | 1310.0 | 1699.2 | 2479.5 | 1640.6 |

546 a) First replicates b) Second replicates

547
548

549 **Table S-2:** Explained variance in PC1 from PCA analysis for all poultry time-series.

| PC1 | Untreated (%) | TFA-treated (%) |
|---|---|---|
| **CCA** | 85 | 62 |
| **CCC** | 74 | 70 |
| **CCF** | 93 | 60 |
| **TMDRA** | 73 | 59 |
| **TMDRC** | 81 | 64 |
| **TMDRF** | 80 | 53 |

550

# 551 References


552 1 Wubshet, S. G. *et al.* FTIR as a rapid tool for monitoring molecular weight distribution
553 during enzymatic protein hydrolysis of food processing by-products. *Analytical*
554 *Methods* **9**, 4247-4254, doi:10.1039/C7AY00865A (2017).
555


556

# Paper IV

# Effects of poultry raw material variation and choice of protease on protein hydrolysate quality

Diana Lindberg[a,*], Kenneth Aase Kristoffersen[a,b], Heleen de Vogel-van den Bosch[c], Sileshi Gizachew Wubshet[a], Ulrike Böcker[a], Anne Rieder[a], Enno Fricke[a] and Nils Kristian Afseth[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: Diana Lindberg, diana.lindberg@nofima.no*

# Abstract

The aim of this study was to investigate the role of industrially relevant variability of chicken and turkey by-products on the product characteristics after enzymatic protein hydrolysis (EPH). The variation was obtained using four by-products, namely chicken carcass, mechanical deboned chicken residue, turkey carcass and mechanical deboned turkey residue. All four materials are typically ingoing raw materials in the same industrial EPH process, and they can be regarded to represent "extremes" in the by-product variation, being materials from before and after mechanical deboning. Three different commercial proteases with different specificities were used to investigate to what extent protease choice influences product characteristics. Besides product protein yield, product characteristics such as nutrient and amino acid composition, size distribution of peptides, degree of hydrolysis (DH), rheological properties, and the hydrolysates ability to inhibit the angiotensin-I-converting enzyme (ACE), was investigated. The results reveal that both choice of enzymes and industrial relevant variations in poultry by-product composition have a major effect on product composition and qualities. When it comes to by-product variations, it is shown that particularly nutritional properties (provided as amino acid composition) and rheological properties are affected. Overall, this means that depending on the targeted market for a given EPH process, by-product variation should be handled and accounted for.

# 1. Introduction

Food by-products represent a viable source for recycling of nutrients into food or feed ingredients and for production of value-added materials and compounds by use of biorefineries [1, 2]. In this context, EPH is a well-established method for valorization of food industry by-products, utilized in production of proteins and lipids for the feed, pet food and human markets [2]. In EPH, by-products rich in proteins are digested by proteases into peptides and single amino acids. The final product quality is influenced by a multitude of factors, which can be grouped into process-specific parameters, such as substrate concentration, enzyme to substrate ratio, time, temperature, and pH; substrate-specific factors, such as origin, age, feed regimen, and complexity; and lastly, protease-specific factors, such as specificity, stability, and sensitivity to inhibitors [3].

The current EPH-based industry is mainly based on processing of by-products from the marine, aquaculture, or the milk industry, with by-products from the meat industry generally not being utilized to the same extent. One of the reasons for this is the regulations restraining the use of by-products from animals by the European Union, for example, the hygiene, the animal by-product, and the Transmissible Spongiform Encephalopathies (TSE) regulations [4-6]. In this context, poultry by-products are of extra interest for the EPH industry as poultry are not susceptible to TSE diseases. Furthermore, large extents of the poultry by-products can be used for human consumption, assuming the whole industrial waste area is included into the clean zone and by-products being treated with the same hygiene rules as food. Based on global production, poultry is one of the largest sources of by-products, and poultry is suggested to be the most sustainable source of meat protein as measured by green-house gas emissions on a per protein basis [7].

Many poultry slaughterhouses facilitate processing of several sources of poultry species (e.g. chicken and turkey) at the same facility, hence generating by-products from several species that could be used for EPH processing. To cut costs in storage and EPH processing, it would be of extra interest for EPH facilities to process the whole raw material without intermediate downtime in EPH processing for shift between species. Especially when aiming for products in pet foods and food, such as bioactive peptides, this assumes that use of the same processing parameters results in products falling within the production specifications, irrespective of species used.

A range of studies use EPH for utilisation of by-products from chicken and turkey, also addressing potential market possibilities [8-12]. Recently, Wubshet et al. studied the potential of using raw material characterisation and processing settings to predict process outcomes in EPH of by-products from poultry [13]. Here, large variations in by-product composition was revealed, affecting process yield and end-product qualities. To the authors knowledge, no studies has been targeted to understand how industrially relevant raw material variations in poultry affect end-product qualities and subsequent market possibilities. The aim of this study was thus to investigate the role of industrially relevant variability of chicken and turkey by-products on the product characteristics after EPH. Three different commercial proteases with different specificities were used to investigate to what extent protease choice influences product characteristics. Besides product protein yield, product characteristics such as nutrient and amino acid composition, size distribution of peptides, DH, rheological properties, and the hydrolysates ability to inhibit the (ACE), was investigated.

# 2. Materials and methods

## 2.1. Chemicals and hydrolysis materials

The poultry raw materials, chicken carcass (CC), mechanical deboned chicken residue (MDCR), turkey carcass (TC) and mechanical deboned turkey residue (MDTR) were supplied from a Norwegian slaughterhouse (Nortura, Hærland, Norway). On the day of collection, the fresh raw materials were ground using a Seydelmann SE130 meat grinder (Seydelmann, Stuttgart, Germany) and a mesh size of 1 cm, vacuum packed in individual packages, and frozen at -20°C until the day of hydrolysis. The raw materials were either used directly from the frozen packages.

All chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA). Water used for HPLC analyses was purified by deionization and 0.22 μm membrane filtration (MilliporeSigma, Burlington, MA, USA).

## 2.2. Enzymatic hydrolysis and sampling

Protein-rich by-products derived from chicken and turkey were hydrolyzed by commercially available food-grade enzyme preparations: Flavourzyme (Fl) and Alcalase 2.4 L (Al) (both from Novozymes A/S, Denmark), and Corolase 2TS (Co) (AB enzymes, Germany). All reactions were performed in a Reactor-Ready™ jacketed reaction vessel (Radleys, Saffron Walden, Essex, United Kingdom). Temperature was kept at 50 °C by water running in a jacketed vessel, circulated using a JULABO circulator pump (Julabo GmbH, Seelbach, Germany). 500g of the raw materials and 1L of water were added to the reactor. All reaction mixtures were thoroughly mixed until the suspension reached 50±1 °C, followed by addition of 1% enzyme preparation (w/w enzyme preparation: by-product). The total reaction time was

80 minutes, from addition of enzyme to termination. All reactions were performed in duplicates. During the hydrolysis experiments, aliquots of approximately 7 mL were collected at 12 time points (i.e., 0.5, 2.5, 5, 7.5, 10, 15, 20, 30, 40, 50, 60 and 80 minutes, respectively). The samples and the products after 80 minutes hydrolysis were thermally inactivated by fast heating using a Menumaster commercial microwave oven (ACP, IA, USA). The temperature was raised to a minimum of 90 °C and kept at this temperature for 15 minutes in a water bath. Following protease inactivation, the samples were centrifuged for 15 minutes (4400 rpm, 25 °C) to separate the mixtures into three phases: sediment, the peptide-containing stickwater, and fat. For the products, fat and stickwater separation were performed using a separator funnel after centrifugation. The weight of all three phases was recorded and the stickwater was aliquoted in 250 ml plastic containers and stored frozen at -40 °C until lyophilized (CHRIST 1-16 LSCplus, Germany). For size exclusion chromatography (SEC) and FTIR analytical measurements of stickwater, aliquots of all samples (hydrolysate time samples and controls) were filtered through a Millex-HV PVDF 0.45 µm 33 mm filter (MilliporeSigma, Billerica, MA, USA) to remove any remaining fat and particulate matter. For analytical measurements of the freeze-dried product no further modifications were done. Except for SEC analysis of the dried products, which were dissolved 2.5% w/v and filtrated through a Millex-HV PVDF 0.45 µm 33 mm filter (MilliporeSigma, Billerica, MA, USA) before analysis.

### 2.3. Nutrient composition

Protein, fat and ash content as well was amino acid composition of all raw materials and hydrolysates were determined. Total nitrogen of each sample was measured according to the Nordic Committee on Food Analysis (NMKL) method NMKL 6 (NMKL, 2003) based on the Kjeldahl method and converted to percentage protein using the standard conversion factor of 6.25. The results are presented Table 1 and in supplementary information (SI), Table S-2.

These results were used to calculate protein recovery. Fat content was determined according to NMKL 131 (NMKL, 1989) established on the principles of the SBR (Schmid-Bondzynski-Ratslaff) method and ash measurements were performed according to the NMKL 173 (NMKL, 2005), i.e. weight determination after incineration of the samples in a muffle furnace at 550°C. The results are presented Table 1 and in the SI, Table S-2. All raw materials and hydrolysates were analyzed for amino acid composition based on the method ISO 13903:2005 as stated by Commission Regulation EC 152/2009. The results are presented in Tables 2 and 3 and in the SI, Table S-1.

### 2.4. Size exclusion chromatography

Injection solutions of standards were prepared in water at a concentration of 2 mg/mL. Filtrates of the water phase collected from the hydrolysis were directly used as injection solution. Chromatographic separation of both standards (in triplicate) and the hydrolysates was performed with an Agilent 1200 series instrument (Agilent Technologies, Santa Clara, CA, USA) 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, separation was performed at 25 °C using BioSep-SEC-s2000 column (Phenomenex, 300 × 7.8 mm). The mobile phase consisted of 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 0.9 mL/min for 17.0 minutes. Between 17.0 and 17.1 minute the mobile phase was changed to $NaH_2PO_4$ (0.10 M) and maintained until 20.0 minute for column cleaning. Elution conditions were restored between 20.0 minute 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, Santa Clara, CA, USA). From chromatographic runs of both the standards and hydrolysates, a UV trace of 214 nm was used. For the analytical

standards, the retention times were obtained using the automatic peak-picking algorithm of OpenLAB CDS. The average of three retention times from triplicate measurements of the standards was used to construct a calibration curve (see the SI, Table S-3). Finally, for each of the chromatograms obtained from the hydrolysate samples were analyzed. Weight-average molecular weights ($M_w$) were calculated and the area under the curve was normalized and divided into four sections based on estimated peptide sizes. Calculations of the average weights and areas were performed automatically using PSS winGPC UniChrom V 8.00 (Polymer Standards Service, Mainz, Germany). The calculation from the software was based on a slicing method, similar to those previously used for analysis of protein hydrolysates [14]. The results are presented in the SI Tables S-4, S-5 and S-6.

### 2.5. FTIR spectroscopy

From each of the filtered protein hydrolysates 7-10 µl was deposited on a 96-well Si-microtiter plate (Bruker Optik GmbH, Germany) and dried at room temperature for at least 45 min to form dried films. From each hydrolysate sample, five aliquots were deposited to allow replicate measurements. FTIR measurements were performed using a High Throughput Screening eXTension (HTS-XT) unit coupled to a Tensor 27 spectrometer (both Bruker Optik). The spectra were recorded in the region between 4000 and 500 $cm^{-1}$ with a spectral resolution of 4 $cm^{-1}$ and an aperture of 5.0 mm and for each spectrum 40 interferograms were collected and averaged. Data acquisition was controlled using Opus v6.5 (Bruker Optik). To enhance spectral resolution the absorbance spectra were converted to 2$^{nd}$ derivative spectra by applying the Savitzky-Golay filter with 13 point smoothing in 2$^{nd}$ polynomial order.

## 2.6. Degree of hydrolysis measurements using the TNBS method

The DH% was measured using the TNBS method described by Satake et al. and Adler-Nissen with some adjustments [15, 16]. 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, Waltham, MA, USA). 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 ºC. 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 following Equation 1, where $h_{tot}$ was calculated from the amino acid composition and protein content.

$$\mathrm{DH\%} = \frac{h}{h_{tot}} \times 100\%$$

Equation 1

## 2.7. Bioactivity measurements

ACE inhibitory activity is based on the liberation of hippuric acid from Hip-His-Leu substrate catalyzed by ACE. The samples were prepared by diluting the protein hydrolysate to a concentration of 0.1% (w/v) in Milli-Q followed by a serially dilution to obtain a range of

sample concentrations (in duplo). In a microtiter plate 25 µL of the samples were mixed with 25 µL of 5 mM Hip-His-Leu (Bachem M-1485) and 25 µL of 25 mU/mL enzyme ACE (Sigma A6778) in 0.1 M borate buffer, pH 8.3 containing 300 mM NaCl. 25 µL 5 mM $Na_2EDTA$ solution was added to the duplo wells as a 100% ACE inhibition. After 2 hours incubation at 37ºC the reaction was stopped by adding 25 µL of 5 mM $Na_2EDTA$ solution. After stopping the reaction 15 µL of 0.25 M NaOH, 35 µL of 0.5 M Bicine, pH 9.1 containing 1 M NaCl and 25 µL of 50 mM TNBS in 0.1 M $Na_2HPO_4$ was added. After 10 minutes incubation at 37ºC the yellow color developed was measured at 405 nm in a spectrophotometer and the $IC_{50}$ (the concentration of inhibitor that inhibits 50% of the activity) was calculated.

### 2.8. Rheology measurements

Freeze dried hydrolysate samples were dissolved in phosphate buffered saline (PBS) at a protein concentration of 5% (approximately equivalent to final protein concentration in samples during hydrolysis) by stirring with a magnetic stirring bar in 50 mL glass bottles with lock. The bottles were placed in a water batch at 80°C for 30 min. Samples were cooled to 20 °C for 30 min and filtered through 0.8 µm syringe filters. Viscosity of the reconstituted hydrolysates was measured using a Physica MCR 301 rheometer (Anton Paar, Stuttgart, Germany) fitted with a double gap geometry (DG26.7). The DG geometry was used with overfill (approximately 10 mL) to avoid undesirable surface tension and capillary effects. After a temperature equilibration of 60 s, apparent viscosity was measured at 22°C at constant shear rate of 10 $s^{-1}$. All measurements were performed in duplicate with a time difference of one year between the replications.

# 3. Results

The by-product variation in the present study was obtained using four by-products, namely chicken carcass, mechanical deboned chicken residue, turkey carcass and mechanical deboned turkey residue. Carcasses are here defined as by-products after removal of head, feathers, neck, intestinal content, breast filet, legs (both thigh and drumstick) and tail. The mechanically deboned residues are the remains after mechanical deboning of the respective carcasses. All four materials are typically ingoing raw materials in the same industrial EPH process, and they can thus be regarded to represent "extremes" in the by-product variation.

## 3.1. Raw material composition

As shown in Table 1 and in the SI Table S-1, the by-products varied in gross chemical composition and amino acid composition, respectively. As expected, there is a relatively higher amount of bone and collagen tissue in the mechanically deboned residue fractions, as evident by the higher amount of ash for both chicken and turkey raw materials after mechanical deboning. The hydroxyproline (Hyp) content of the materials, indicative of the amount of collagen tissue, is also varying with by-products. The amount of collagen tissue in the different by-products can be estimated from the content of hydroxyproline (see the SI, Table S-1). Based on the assumption that collagen tissue consists of 13.5% Hyp, the collagen content in CC, MCDR, TC and MDTR were 9.4, 21, 26 and 38 %, respectively [17, 18]. The higher fat contents observed in turkey by-products is well in accordance with data from a previous study on Norwegian poultry by-products [13].

### 3.2. Product composition

In this study, Corolase 2TS (Corolase), Alcalase 2.4L (Alcalase), and Flavourzyme were used in EPH of all four by-products defined above. The hydrolysis reactions were all run at ambient pH using the same process settings irrespective of protease and by-product. These settings were selected to come as close as possible to those used in the current hydrolysis industry in Norway.

*Chemical composition of hydrolysates*. The gross chemical composition of the hydrolysates is provided in the SI, Table S-2. Hydrolysates produced by using Alcalase or Corolase contained higher amounts of protein than hydrolysates generated by Flavourzyme, irrespective of raw material. Carcass by-products result in higher protein contents than the mechanically deboned residues. Interestingly, the much higher amount of ash in the turkey by-products is not necessarily reflected in the hydrolysate products. Instead, hydrolysates derived from chicken by-products had higher ash contents. It can be postulated that this can be a result of a better hydrolysis of the collagenous bone material in chicken as compared to turkey.

*Protein recovery.* The protein yield of the different hydrolysates is presented in Fig. 1, alongside the protein content. When comparing protein yield between the different raw materials, turkey by-products seemingly result in higher protein yields than chicken by-products. However, it is known that measurements on fat-rich raw materials can lead to an underestimation of the Kjeldahl protein concentration due to incomplete digestion of the start sample [19]. Turkey raw materials contain higher fat than the chicken samples, this might lead to overestimation of the final protein yield in these hydrolysates. Alcalase generally

results in the highest protein yield, closely followed by Corolase, while Flavourzyme results in significantly lower protein yields than the other two enzymes.

A protein recovery of 91% was reported by Kurozawa et al. to be the maximum achievable protein recovery using Alcalase on chicken breast meat at optimum conditions, i.e. 52.5 ºC, 4.2% (w/w) enzyme preparation:substrate ratio, and a pH value of 8.00 [20]. It is noteworthy that Corolase with its two endoproteases results in almost as high protein yields as in Alcalase hydrolysis, irrespective of by-product. According to the manufacturers specifications, Corolase has quite a broad pH profile between pH 6 to pH 9, and these results prove that the enzyme preparation is well suited to hydrolyze these materials at the lower end of this range. The low yield seen with Flavourzyme is likely explained by the fact that in this enzyme preparation, a significant part of the total protease activity is derived from proteases only able to digest proteins and peptides from their respective N- and C-terminal ends.

*Degree of hydrolysis*. A low autolytic activity is expected for the kind of materials used in this study as they exclude the major protease-containing organs, such as viscera and intestines, where high autolytic activity is seen [8]. DH-values for all reactions are provided in Fig. 2. The figure reveals that the type of raw material hydrolyzed seems to have only a minor influence on the DH, but for Alcalase and Corolase there is a trend towards lower DH values for turkey by-products. Furthermore, as seen in the figure, the difference in DH values between Alcalase and Corolase hydrolysates are generally low compared to the Flavourzyme hydrolysates. The relatively higher DH of Flavourzyme compared to Alcalase has also been reported for hydrolysis of salmon by-products [21].

Corolase, Alcalase and Flavourzyme have previously been shown to yield products of low, medium and high degree of hydrolysis, respectively, using lupin proteins as substrate [22]. Merz and coworkers showed that Corolase, which resulted in the lowest DH of these three commercial protease products, contains two endopeptidases. Alcalase contains three endo-proteases and one exo-peptidase. Flavourzyme, resulting in the highest DH values, was shown to contain a total of 10 proteases, whereof 4 exhibit endo-protease, 4 exo- or carboxy-peptidase, and 2 dipeptidyl peptidase activity. The DH values resulting from any hydrolysis are dependent on both process settings such as enzyme concentration and time, as well as the amount of easily accessible proteins within the material. Also, the method (OPA or TNBS), which is used to determine the DH strongly affects the results [23]. The DH reported by Merz and coworkers can therefore not be directly compared with the DH obtained in the present study. It is still interesting to note that the observed trend in DH between proteases from the Merz study is seen also in this study for all materials except for the mechanically deboned chicken residues.

*Protein-size distributions*. In order to gain insight into protein size distribution all hydrolysate samples were subjected to SEC and FTIR analysis. The SEC profiles in Fig. 3 A) to D) clearly show that samples derived from all four raw materials that were treated with Flavourzyme consist of fewer large (F1) and medium size peptides (F2) while containing higher shares of small peptides (F3) and free amino acids (F4). At the same time, they also display a smaller $M_w$ as compared to the Alcalase and Corolase derived hydrolysates (see SI, the Table S-6). The information from the SEC profiles is consistent with the corresponding FTIR 2$^{nd}$ derivative spectra in the region from 1800-1200 cm$^{-1}$. Average spectra for the raw materials and three enzyme preparations are presented in Fig. 4 A) to D). For all 4 raw materials there is a clear larger share of free carboxylate and free amino groups for Flavourzyme hydrolysates compared to Alcalase- and Corolase-derived hydrolysates. This is underlined by larger

negative peaks at 1587 $cm^{-1}$ and 1407 $cm^{-1}$ corresponding to $COO^-$ asymmetric and symmetric stretching vibrations, respectively as well as a larger contribution from the $NH_3^+$ scissoring vibration at 1515 $cm^{-1}$ for Flavourzyme [24]. Additionally, there is less contribution from secondary structures from medium and large proteins and peptides in the Amide I (1600-1700 $cm^{-1}$) and Amide II around 1550 $cm^{-1}$ in the spectral signatures from Flavourzyme hydrolysates.

*Rheology.* The majority of hydrolysates had a very low viscosity in the range of 1-3 mPas (viscosity of water at 20 °C is 1.002 mPas) when re-dissolved at 5% protein content in PBS. This is shown in Fig. 5. Hydrolysates prepared with Flavourzyme showed generally higher viscosities than hydrolysates prepared with Alcalase or Corolase. Especially the collagen-rich turkey by-products resulted in hydrolysates with a higher viscosity in combination with Flavourzyme. This was most pronounced for MDTR. When MDTR was dissolved at higher protein concentrations (10 and 20%), the liquid gelled during cooling to room temperature (results not shown). This indicates that MDTR, and to a lesser extent TC-hydrolysates, prepared with Flavourzyme contain large enough collagen peptides to potentially form a three-dimensional gel-network like it is found in gelatin gels. This is reflected in the higher content of collagen amino acids (Gly, Pro, Hyp) in the turkey hydrolysates (Table 2). However, the hydrolysates prepared with Flavourzyme had lower contents of collagen amino acids than those prepared with Alcalase or Corolase (Table 3) but might contain a small portion of non-degraded collagen with functional properties even though they had a higher DH as the hydrolysates prepared with Alcalase or Corolase. The SEC chromatograms of the different hydrolysates did in general not indicate a higher relative proportion of large peptides in the MDTR and TC samples hydrolysed with Flavourzyme compared to the other samples.

However, there was a small peak at 6 min, only seen in these two samples, which might consist of little or undegraded dissolved collagen molecules.

The viscosity values in Fig. 5 are average values of two separately dissolved sample duplicates. The values obtained for TC and MDTR hydrolysates prepared with Flavourzyme were significantly lower at the second measurement after one year storage of the freeze dried hydrolysate powders at -20°C. Frozen storage has been reported to decrease the solubility of macromolecules such as beta-glucan, presumably due to the formation of inter-molecular aggregates [25, 26]. The lower viscosity values after storage may therefore be the result of a reduced protein/peptide solubility.

*Amino acid composition*. As seen in Table 2, the amino acid composition of the hydrolysates clearly varies depending on both by-products and proteases used. The largest differences in amino acid composition based on by-product differences are the relative higher amount of charged (His, Lys, Asp, Glu) and polar uncharged amino acids (Ser, Thr) seen in chicken carcass hydrolysates as compared to the rest of the materials. Another single amino acid with large concentration differences is Trp, with turkey carcass hydrolysates resulting in the highest amount. Surprisingly, there is also a tendency of hydrolysis with Flavourzyme resulting in the highest amounts of Trp, irrespective of raw material. There is a generally higher amount of the collagen amino acids (Gly, Pro, Hyp) in the turkey hydrolysates as compared to chicken hydrolysates, with the highest amounts in the mechanically deboned residues. One example of large differences seen in amount of collagen peptides between hydrolysates is the one seen between Alcalase-treated mechanically deboned residues, which

contains 6.7 times more Hyp than Flavourzyme-treated chicken carcass, 4.8 against 0.72 g/100g, respectively.

*Nutritional properties.* For the EPH industry aiming at marketing their poultry hydrolysates for general nutritional purposes, e.g. as a protein supplement hindering muscle deterioration in elderly, the sum of essential amino acids is of extra interest. For athletes and bodybuilders on the other hand, the amount of branch-chained amino acids (BCAA), comprising the three essential amino acids Leu, Val, Ile, is of extra interest. For the latter markets, the sum of BCAA is commonly used for marketing purposes in these supplements arguing that these aid with muscle growth and performance [27]. For industry that focus on food supplements containing high amounts of collagen peptides, the sum of collagen peptides is in focus. In Table 3, the sum of amino acids for these three nutritionally important groups of amino acids in the hydrolysates are presented.

For the sum of essential amino acids, the highest values are seen in hydrolysis of chicken carcass, and the lowest in hydrolysis of the MDTR raw material, varying 25% between the highest and the lowest values. The difference between the highest (chicken carcass) and lowest (turkey by-products) values for the BCAAs is almost as large. The reverse relationship is valid for the collagen peptides, with the difference between the highest and lowest value being 35%, and the preferred substrate for high concentrations being mechanically deboned turkey. The values in Tables 2-3 and S-1 in the IS also show that Flavourzyme with its mostly exo-peptidase activity is least effective protease for extraction of collagen peptides from these materials.

### 3.3. Bioactivity

One of the markets with the highest potential for peptides from hydrolysates are as functional foods or when used as pharmaceuticals [28, 29]. Probably the most frequently reported bioactivity assays used for screening of bioactive peptides is the ACE assays, based on inhibition of the angiotensin-I-converting enzyme [30]. Alcalase has been previously reported to exhibit ACE inhibition in hydrolysis of chicken leg as well as several other by-products [31]. As presented in Fig. 6, Alcalase also exhibit a moderate ACE inhibition on the currently investigated by-products, with the highest inhibition seen on by-products with the highest content of meat, favoring chicken before turkey by-products for the lowest $IC_{50}$ value. The best ACE inhibition is seen in hydrolysates originating from proteolysis using Corolase. The $IC_{50}$ values originating from hydrolysis using either Flavourzyme or control samples are above 10, which suggests these samples show no ACE inhibition. Although no investigations have been made in finding the peptides responsible for the presented inhibitions, these results imply that when aiming for finding ACE inhibition candidates, both choice of enzyme preparation and by-products are important parameters influencing the resulting hydrolysate inhibitory properties.

## 4. General discussion

The results of this study show that choice of enzymes has a major effect on product composition and product quality. The results also reveal that industrial relevant variations in poultry by-product composition have a significant extent effect on product composition. This has implications both related to processing and for potential markets. As turkey by-products in general have higher contents of collagen proteins, care should be taken in order to choose enzymes for EPH with the sufficient efficiency of degrading such materials. This is shown in

the SEC and the rheology results, where selected turkey and enzyme combinations provided hydrolysates with gelatin-like behavior. The study also shows that both by-product and enzyme type affect the nutritional properties of the hydrolysates, as revealed by the amino acid composition of the hydrolysates. Thus, in an ideal world, the choice of enzyme and by-product would be guided by the targeted market of the hydrolysate being produced.

The present study evaluated four poultry by-products as "extreme" samples spanning all industrially relevant chemical variations in the by-products. In industrial processing, these by-products will to a varying degree be mixed, based on day-to-day variations such as type of raw materials processed and processing settings used, et cetera. Thus, the properties of the hydrolysates produced could be varying significantly from day to day, and even from hour to hour, depending on this by-product variation. As shown in this study, the need to adjust processing settings related to by-product variation will depend on the targeted market for the given process. The question is then how to account for such by-product variability going into an EPH process. One strategy to handle this would involve better "sorting" of the by-products, for instance not allowing by-products of different species to be mixed. This could involve severe practical challenges related to storage and oxidation. Wubshet et al. provided another solution for this where vibrational spectroscopy of by-products and processing settings were used to predict product properties of the hydrolysates [13]. It is a question if the analytical tools studied have the chemical specificity needed to address the key quality parameters of the by-products being responsible for the relevant product properties. The present study presents a range of complimentary analytical techniques giving a good overview of the physical and chemical properties of the hydrolysates and the results could give important indications on which relevant parameters should be analyzed in an industrial setting. Detailed analytical tools like NMR and LCMS would be needed to complement this picture even

442 further and would most likely reveal even greater differences in product qualities related to
443 by-product variation. Overall, such characterization would be needed in order to fully
444 understand industrial EPH and potential markets given the relevant by-product variability.
445

## 446 5. Conclusion

447 The results of this study show that both choice of enzymes and industrial relevant variations
448 in poultry by-product composition have a major effect on product composition and product
449 qualities in EPH. This means that depending on the targeted product from a given EPH
450 process, by-product variation should be handled and accounted for. Only then the true
451 potential of using enzymes for tailor-making EPH products with specific properties could be
452 realized.
453

## 454 Acknowledgements

455 Silje Kristine Aas and Nina Christine Helle (students from OsloMet University) are greatly
456 acknowledged for their contributions to the experimental work. Financial support from the
457 Norwegian Fund for Research Fees for Agricultural Products through the projects
458 “FoodSMaCk” (no. 262308) and “SunnMat” (no. 262300) and internal financing from
459 Nofima through the project “PepTek” is also greatly acknowledged.
460

# References


[1] J. Esteban, M. Ladero, Food waste as a source of value-added chemicals and materials: a biorefinery perspective, International Journal of Food Science & Technology 53(5) (2018) 1095-1108.

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

[3] S.G. Wubshet, D. Lindberg, E. Veiseth-Kent, K.A. Kristoffersen, U. Böcker, K.E. Washburn, N.K. Afseth, Chapter 8 - Bioanalytical Aspects in Enzymatic Protein Hydrolysis of By-Products, in: C.M. Galanakis (Ed.), Proteins: Sustainable Source, Processing and Applications, Academic Press2019, pp. 225-258.

[4] The European Parliament and the Council of the European Union, REGULATION (EC) No 853/2004 OF THE EUROPEAN PARLIAMENTAND OF THE COUNCIL of 29 April 2004: laying down specific hygiene rules for on the hygiene of foodstuffs, in: T.E. Union (Ed.) 2004.

[5] The European Parliament and the Council of the European Union, REGULATION (EC) No 1069/2009 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 21 October 2009: laying down health rules as regards animal by-products and derived products not intended for human consumption and repealing Regulation (EC) No 1774/2002 (Animal by-products Regulation), in: T.E. Union (Ed.) 2009.

[6] The European Parliament and the Council of the European Union, REGULATION (EC) No 999/2001 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 22 May 2001: laying down rules for the prevention, control and eradication of certain transmissible spongiform encephalopathies, 2001.

[7] Food and Agriculture Organization of the United Nations, Global Livestock Environmental Assessment Model (GLEAM), 2019. http://www.fao.org/gleam/results/en/. (Accessed 2 October 2019).

[8] D. Lapeña, K.S. Vuoristo, G. Kosa, S.J. Horn, V.G.H. Eijsink, Comparative Assessment of Enzymatic Hydrolysis for Valorization of Different Protein-Rich Industrial Byproducts, Journal of Agricultural and Food Chemistry 66(37) (2018) 9738-9749.

[9] A. Brandelli, L. Sala, S.J. Kalil, Microbial enzymes for bioconversion of poultry waste into added-value products, Food Research International 73 (2015) 3-12.

[10] K. Jayathilakan, K. Sultana, K. Radhakrishna, A.S. Bawa, Utilization of byproducts and waste materials from meat, poultry and fish processing industries: a review, J Food Sci Technol 49(3) (2012) 278-293.

[11] A. Lasekan, F. Abu Bakar, D. Hashim, Potential of chicken by-products as sources of useful biological resources, Waste Management 33(3) (2013) 552-565.

[12] M.C. Gómez-Guillén, B. Giménez, M.E. López-Caballero, M.P. Montero, Functional and bioactive properties of collagen and gelatin from alternative sources: A review, Food Hydrocolloids 25(8) (2011) 1813-1827.

[13] S.G. Wubshet, J.P. Wold, N.K. Afseth, U. Böcker, D. Lindberg, F.N. Ihunegbo, I. Måge, Feed-Forward Prediction of Product Qualities in Enzymatic Protein Hydrolysis of Poultry By-products: a Spectroscopic Approach, Food and Bioprocess Technology 11(11) (2018) 2032-2043.

[14] D.S.T. Hsieh, C. Lin, E.R. Lang, N. Catsimpoolas, C.K. Rha, Molecular-weight distribution of soybean globulin peptides produced by peptic hydrolysis, Cereal Chemistry 56(4) (1979) 227-231.

[15] J. Adler-nissen, Determination of the degree of hydrolysis of food protein hydrolysates by trinitrobenzenesulfonic acid, Journal of Agricultural and Food Chemistry 27(6) (1979) 1256-1262.
[16] K. Satake, T. Okuyama, M. Ohashi, T. Shinoda, The spectrophotometric determination of amine, amino acid and peptide with 2,4,6-trinitrobenzene 1-sulfonic acid, Journal of Biochemistry 47(5) (1960) 654-660.
[17] I. Stoilov, B.C. Starcher, R.P. Mecham, T.J. Broekelmann, Chapter 7 - Measurement of elastin, collagen, and total protein levels in tissues, in: R.P. Mecham (Ed.), Methods in Cell Biology, Academic Press2018, pp. 133-146.
[18] T. Sadat, C. Volle, Integration of a linear accelerator into a production line of mechanically deboned separated poultry meat, Radiation Physics and Chemistry 57(3) (2000) 613-617.
[19] P. Saez-Plaza, M.J. Navas, S. Wybraniec, T. Michalowski, A.G. Asuero, An Overview of the Kjeldahl Method of Nitrogen Determination. Part II. Sample Preparation, Working Scale, Instrumental Finish, and Quality Control, Critical Reviews in Analytical Chemistry 43(4) (2013) 224-272.
[20] L.E. Kurozawa, K.J. Park, M.D. Hubinger, Optimization of the Enzymatic Hydrolysis of Chicken Meat Using Response Surface Methodology, Journal of Food Science 73(5) (2008) C405-C412.
[21] H.G. Kristinsson, B.A. Rasco, Fish Protein Hydrolysates: Production, Biochemical, and Functional Properties, Critical Reviews in Food Science and Nutrition 40(1) (2000) 43-81.
[22] M. Merz, W. Claaßen, D. Appel, P. Berends, S. Rabe, I. Blank, T. Stressler, L. Fischer, Characterization of commercially available peptidases in respect of the production of protein hydrolysates with defined compositions using a three-step methodology, Journal of Molecular Catalysis B: Enzymatic 127 (2016) 1-10.
[23] S.M. Rutherfurd, Methodology for Determining Degree of Hydrolysis of Proteins in Hydrolysates: A Review, Journal of Aoac International 93(5) (2010) 1515-1522.
[24] U. Böcker, S.G. Wubshet, D. Lindberg, N.K. Afseth, Fourier-transform infrared spectroscopy for characterization of protein chain reductions in enzymatic reactions, Analyst 142(15) (2017) 2812-2818.
[25] M.U. Beer, P.J. Wood, J. Weisz, N. Fillion, Effect of Cooking and Storage on the Amount and Molecular Weight of (1→3)(1→4)-β-d-Glucan Extracted from Oat Products by an In Vitro Digestion System, Cereal Chemistry 74(6) (1997) 705-709.
[26] T.H. Gamel, K. Badali, S.M. Tosh, Changes of β-glucan physicochemical characteristics in frozen and freeze dried oat bran bread and porridge, Journal of Cereal Science 58(1) (2013) 104-109.
[27] D.-H. Kim, S.-H. Kim, W.-S. Jeong, H.-Y. Lee, Effect of BCAA intake during endurance exercises on fatigue substances, muscle damage substances, and energy metabolism substances, J Exerc Nutrition Biochem 17(4) (2013) 169-180.
[28] J. Zamora-Sillero, A. Gharsallaoui, C. Prentice, Peptides from Fish By-product Protein Hydrolysates and Its Functional Properties: an Overview, Marine Biotechnology 20(2) (2018) 118-130.
[29] T. Lafarga, M. Hayes, Bioactive protein hydrolysates in the functional food ingredient industry: Overcoming current challenges, Food Reviews International 33(3) (2017) 217-246.
[30] Y.-M. Lee, T. Skurk, M. Hennig, H. Hauner, Effect of a milk drink supplemented with whey peptides on blood pressure in patients with mild hypertension, European Journal of Nutrition 46(1) (2007) 21-27.
[31] F.-Y. Cheng, T.-C. Wan, Y.-T. Liu, C.-M. Chen, L.-C. Lin, R. Sakata, Determination of angiotensin-I converting enzyme inhibitory peptides in chicken leg bone protein hydrolysate with alcalase, Animal Science Journal 80(1) (2009) 91-97.

# Tables

**Table 1.** By-product proximate analysis data (g/100g raw material)

| | CC | MDCR | TC | MDTR |
|---|---|---|---|---|
| Protein | 19 | 19 | 17 | 18 |
| Fat | 9.9 | 8.4 | 19 | 15 |
| Ash | 3.8 | 6.8 | 6.4 | 8.8 |
| Water | 65 | 64 | 56 | 55 |

**Table 2.** The resulting amino acid profiles of the product hydrolysates.

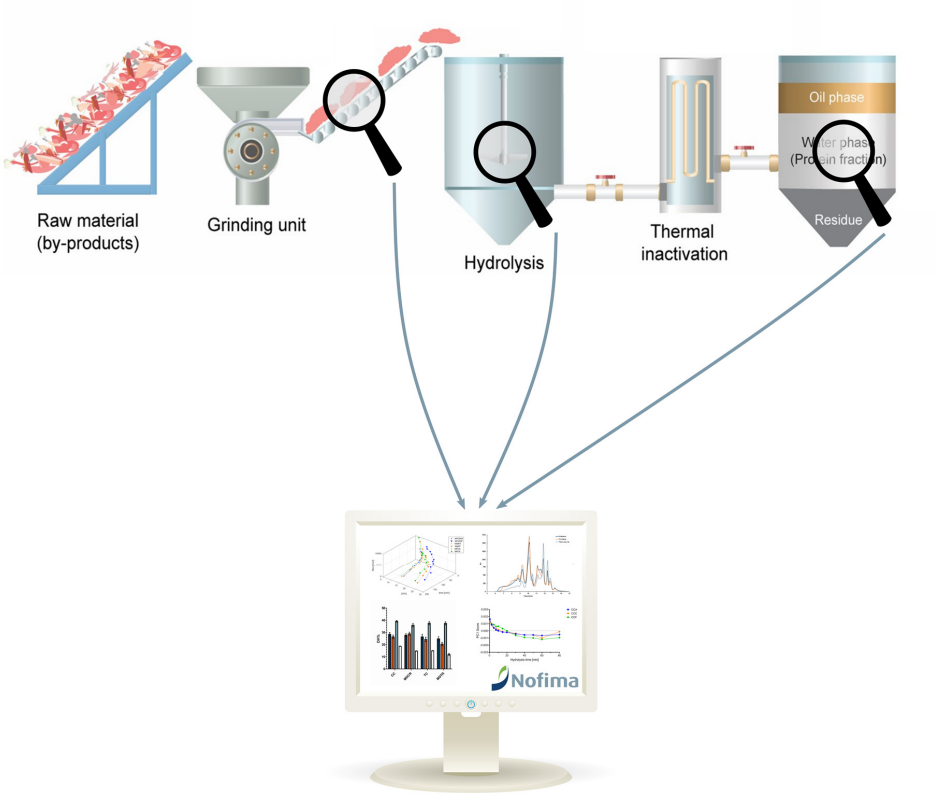
| Raw material | CC | | | MDCR | | | TC | | | MDTR | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Enzyme | Al | Co | Fl | Al | Co | Fl | Al | Co | Fl | Al | Co | Fl |
| Arg | 5.2 | 5.4 | 4.5 | 5.2 | 5.2 | 4.1 | 4.6 | 5.2 | 4.6 | 5.6 | 5.6 | 4.5 |
| His | 2.2 | 2.2 | 2.1 | 1.8 | 1.9 | 1.6 | 1.6 | 1.8 | 1.9 | 1.6 | 1.6 | 1.6 |
| Lys | 6.7 | 6.9 | 6.6 | 5.5 | 5.7 | 5.3 | 5.1 | 5.6 | 5.9 | 4.9 | 4.8 | 5.0 |
| Asp | 7.3 | 7.6 | 6.6 | 6.5 | 6.7 | 5.8 | 6.0 | 6.5 | 6.4 | 6.3 | 6.2 | 5.9 |
| Glu | 12 | 12 | 12 | 11 | 11 | 10 | 10 | 11 | 12 | 11 | 10 | 11 |
| Ser | 3.2 | 3.2 | 2.9 | 2.9 | 2.9 | 2.6 | 2.7 | 3.0 | 2.9 | 2.9 | 2.9 | 2.7 |
| Thr | 3.5 | 3.6 | 3.1 | 3.1 | 3.2 | 2.8 | 2.7 | 3.0 | 2.9 | 2.8 | 2.8 | 2.7 |
| Cys + Cys ox | 0.88 | 0.82 | 0.79 | 0.81 | 0.83 | 0.65 | 0.85 | 0.82 | 0.73 | 0.72 | 0.66 | 0.72 |
| Gly | 5.7 | 6.1 | 4.1 | 8.2 | 7.8 | 5.3 | 7.1 | 7.4 | 6.0 | 11.0 | 10.6 | 7.0 |
| Pro | 3.9 | 4.2 | 3.0 | 5.1 | 4.9 | 3.4 | 4.6 | 4.8 | 4.2 | 6.6 | 6.4 | 4.5 |
| Hyp | 1.5 | 1.6 | 0.7 | 3.0 | 2.7 | 1.5 | 2.8 | 2.6 | 2.0 | 4.8 | 4.6 | 2.4 |
| Ala | 5.3 | 5.5 | 4.7 | 5.6 | 5.5 | 4.5 | 4.9 | 5.3 | 5.0 | 6.2 | 6.0 | 5.0 |
| Val | 3.8 | 3.9 | 3.4 | 3.4 | 3.5 | 3.0 | 3.0 | 3.2 | 3.2 | 3.2 | 3.0 | 3.1 |
| Ile | 3.4 | 3.4 | 3.1 | 2.9 | 2.9 | 2.7 | 2.7 | 2.9 | 3.0 | 2.6 | 2.5 | 2.7 |
| Leu | 5.9 | 5.9 | 5.5 | 5.1 | 5.2 | 4.6 | 4.6 | 5.1 | 5.2 | 4.7 | 4.6 | 4.7 |
| Met | 2.1 | 1.9 | 1.8 | 1.9 | 2.0 | 1.6 | 2.0 | 2.0 | 1.8 | 1.7 | 1.5 | 1.6 |
| Phe | 3.4 | 3.5 | 3.1 | 3.1 | 3.1 | 2.8 | 2.7 | 2.9 | 2.9 | 2.8 | 2.8 | 2.7 |
| Tyr | 2.4 | 2.3 | 1.9 | 1.9 | 2.0 | 1.6 | 1.8 | 2.0 | 1.9 | 1.7 | 1.7 | 1.8 |
| Trp | 0.10 | 0.11 | 0.13 | 0.09 | 0.06 | 0.19 | 0.26 | 0.21 | 0.30 | 0.10 | 0.03 | 0.22 |

565 **Table 3.** The sum of BCAA and essential amino acids in the turkey and chicken hydrolysates.
566 Figures rounded to closest two-digit number after summarizing the individual amounts of
567 amino acids.

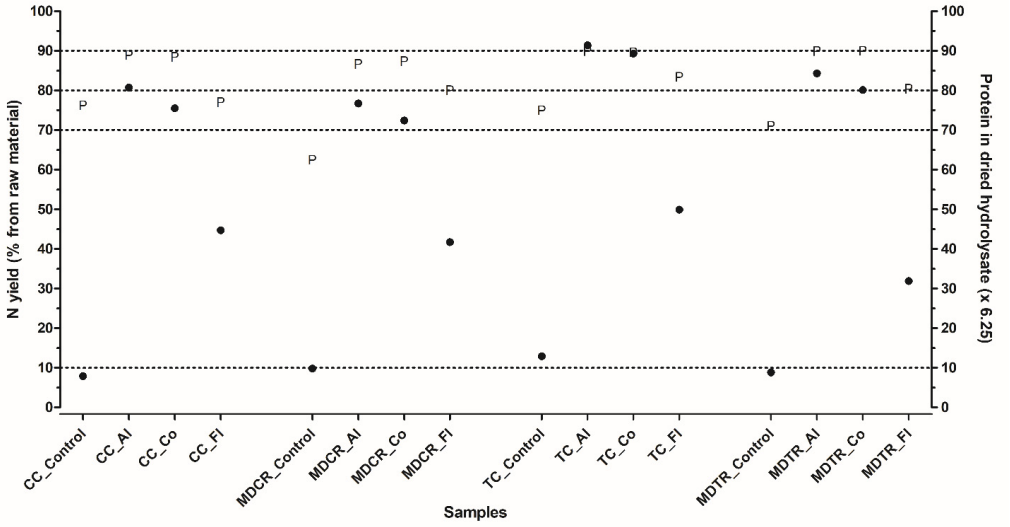
| | CC | | | MDCR | | | TC | | | MDTR | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sum of amino acids (g/100g) | Al | Co | Fl | Al | Co | Fl | Al | Co | Fl | Al | Co | Fl |
| Essential amino acids | 31 | 31 | 29 | 27 | 27 | 24 | 25 | 26 | 27 | 24 | 24 | 24 |
| BCAAs | 13 | 13 | 12 | 11 | 12 | 10 | 10 | 11 | 11 | 11 | 10 | 11 |
| Collagen peptides | 11 | 12 | 7.8 | 16 | 16 | 10 | 14 | 15 | 12 | 22 | 22 | 14 |

568
569
570

# Figures




**Graphical abstract**




575

576 **Fig. 1.** The protein concentration (%) in the powder fraction after freeze drying (P) and the

577 final protein yield (•) (Kjeldahl protein, g/100 g dry weight) for all samples that has been

578 hydrolysed.

579

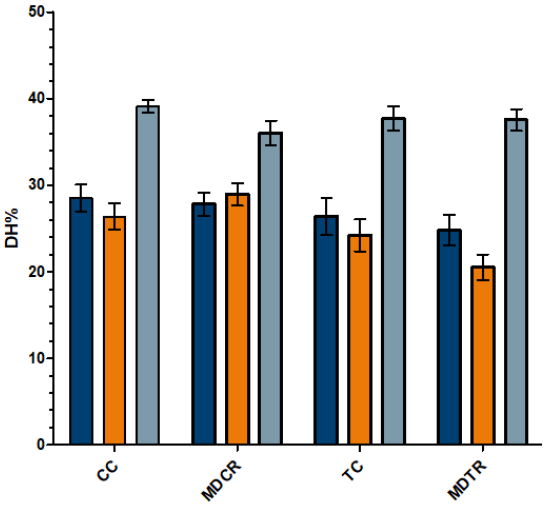



**Fig. 2.** The DH values of the hydrolysates as measured in triplicate by the TNBS method (error bars: SD). The Alcalase hydrolysates are visualized in dark blue, Corolase hydrolysates in orange, Flavourzyme hydrolysates in light blue.

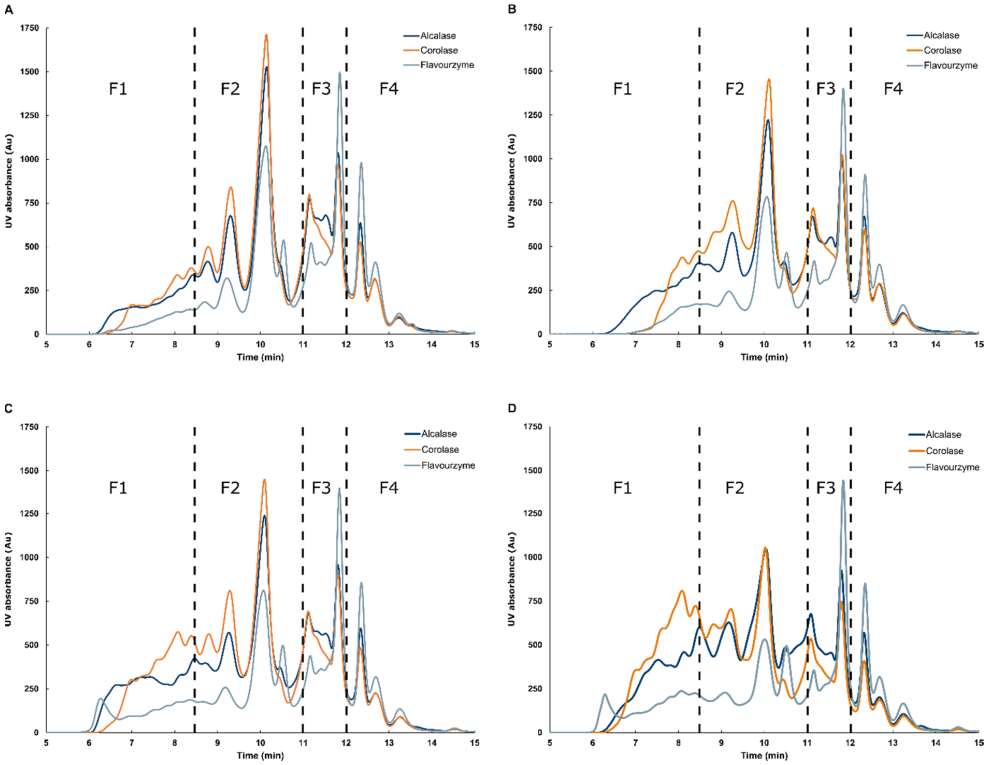
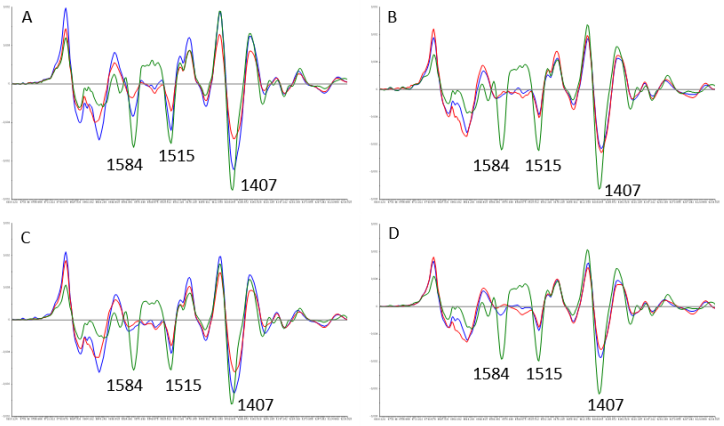



**Fig. 3.** The resulting size exclusion chromatograms for the hydrolysates after treatment with the three proteases on all four materials: A) chicken carcass, B) MDCR, C) turkey carcass, and D) MDTR.




590

591 **Fig. 4.** FTIR 2nd derivative spectra 1800-1200 $cm^{-1}$; Alcalase, Corolase and Flavourzyme in

592 blue, red and green, respectively. A) chicken carcass, B) MDCR, C) turkey carcass, and D)

593 MDTR.

594

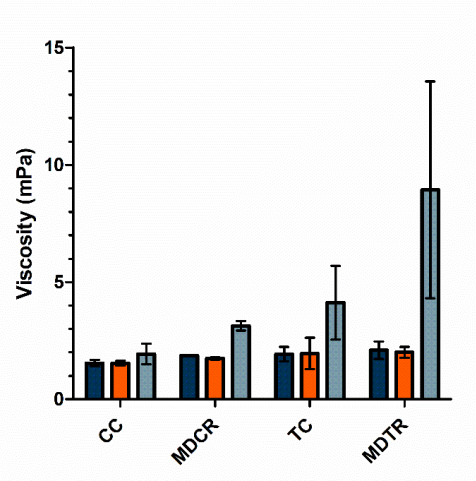
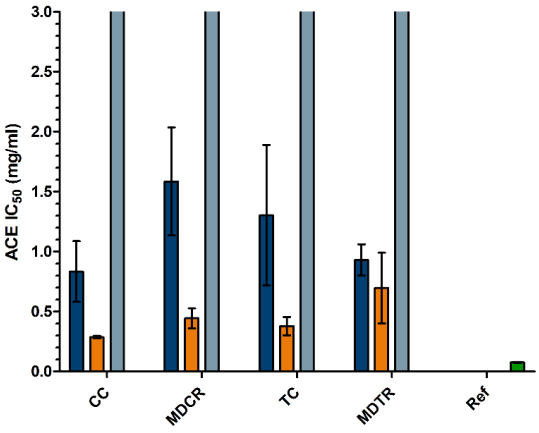



**Fig. 5.** The resulting viscosity, divided by raw material, after 5% of each hydrolysate were mixed with water at 80ºC and stirred for 30 minutes, followed by cooling for 30 minutes at 20ºC.The Alcalase hydrolysates are visualized in dark blue, Corolase hydrolysates in orange, and Flavourzyme hydrolysates in light blue.




601

602 **Fig. 6.** The inhibitory concentration needed the study hydrolysates to achieve a 50% reduction
603 of porcine ACE activity. The same colour-scheme as previously is utilized for the proteases
604 (Alcalase, dark blue; Corolase, orange; Flavourzyme, light blue). Also included in green, to
605 the right, are the $IC_{50}$ value for a reference hydrolysate obtained from digestion of lysozyme.

606

# Supplementary Information

## Effects of poultry raw material and protease variations on protein hydrolysate quality

Diana Lindberg[a,*], Kenneth Aase Kristoffersen[a,b], Heleen de Vogel-van den Bosch[c], Sileshi Gizachew Wubshet[a], Ulrike Böcker[a], Anne Rieder[a], Enno Fricke[a] and Nils Kristian Afseth[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: Diana Lindberg, diana.lindberg@nofima.no*

620 **Table S-4:** Protein, fat and ash content (g/100g) in lyophilized hydrolysates.
621

| Amino acids | CC | MDCR | TC | MDTR |
|---|---|---|---|---|
| Arg | 1.1 | 1.2 | 1.0 | 1.1 |
| His | 0.48 | 0.42 | 0.32 | 0.30 |
| Lys | 1.5 | 1.3 | 0.98 | 0.89 |
| Asp | 1.6 | 1.5 | 1.2 | 1.1 |
| Glu | 2.5 | 2.4 | 1.9 | 1.9 |
| Ser | 0.70 | 0.68 | 0.53 | 0.56 |
| Thr | 0.77 | 0.71 | 0.53 | 0.51 |
| Cys + Cys ox | 0.17 | 0.13 | 0.14 | 1.2 |
| Gly | 1.2 | 1.8 | 1.6 | 2.2 |
| Pro | 0.81 | 1.2 | 1.0 | 1.3 |
| Hyp | 0.28 | 0.65 | 0.64 | 1.1 |
| Ala | 1.1 | 1.2 | 1.0 | 1.2 |
| Val | 0.84 | 0.79 | 0.59 | 0.55 |
| Ile | 0.76 | 0.67 | 0.50 | 0.45 |
| Leu | 1.3 | 1.2 | 0.91 | 0.85 |
| Met | 0.39 | 0.30 | 0.34 | 0.29 |
| Phe | 0.78 | 0.74 | 0.58 | 0.53 |
| Tyr | 0.56 | 0.48 | 0.39 | 0.35 |
| Trp | 0.21 | 0.17 | 0.14 | 0.11 |
| | | | | |
| Total aa (g/100 g) | 17 | 18 | 14 | 17 |
| Hyp % of total aa's | 1.6 | 3.7 | 4.5 | 6.7 |
| Collagen* | 9.4 | 21 | 26 | 38 |

622

623

624 **Table S-2:** Protein, fat and ash content (g/100g) in lyophilized hydrolysates.

| Raw material | CC | | | MDCR | | | TC | | | MDTR | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Enzyme | Al | Co | Fl | Al | Co | Fl | Al | Co | Fl | Al | Co | Fl |
| Protein | 85 | 87 | 82 | 84 | 85 | 78 | 88 | 88 | 83 | 88 | 85 | 79 |
| Fat | 0.49 | 0.18 | 0.10 | 0.28 | 0.17 | 0.3 | 0.31 | 0.25 | 0.33 | 0.76 | 1.3 | 1.4 |
| Ash | 5.8 | 5.8 | 8.3 | 5.9 | 8.8 | 8.7 | 4.7 | 5.1 | 7.8 | 4.7 | 5.0 | 8.6 |

625

626

627

**Table S-3:** The molecular weights (Mw), average retention times (RT) and standard deviation (SD) for the analytical standards for one BioSep-SEC-s2000 column calibration used.

| Compound name | Mw ($g\ mol^{-1}$) | Average RT[a] (min.) | SD RT (min.) |
|---|---|---|---|
| Bovine albumin | 66000 | 6.039 | 0.004 |
| Albumin from chicken egg white | 44287 | 6.108 | 0.001 |
| Carbonic anhydrase | 29000 | 6.130 | 0.004 |
| Lysozyme | 14300 | 6.703 | 0.001 |
| Cytochrome c from bovine heart | 12327 | 6.431 | 0.001 |
| Aprotinin from bovine lung | 6511 | 7.313 | 0.004 |
| Insulin Chain B Oxidized from bovine pancreas | 3496 | 9.041 | 0.007 |
| Angiotensin II human | 1046 | 8.693 | 0.004 |
| Bradykinin Fragment 1-7 | 757 | 9.300 | 0.006 |
| Val-Tyr-Val | 379 | 9.728 | 0.004 |
| Tryptophan | 204 | 11.785 | 0.007 |

a) The average of three measurements of each analytical standard were used to construct a linear fitted calibration curve.

**Table S-4:** Weight-average molecular weight ($M_w$, g/mol) values for all chicken EPH samples.

| Time (min)[a] | 1) CC Al | 1) CC Co | 1) CC Fl | 1) MDCR Al | 1) MDCR Co | 1) MDCR Fl |
|---|---|---|---|---|---|---|
| 0.5 | 3246.3 | 2558.5 | 1617.1 | 3100.3 | 2684.5 | 1634.0 |
| 2.5 | 2796.8 | 2463.9 | 1653.5 | 2551.4 | 2506.6 | 1915.3 |
| 5 | 2714.1 | 2608.1 | 1821.6 | 2249.5 | 2445.3 | 2865.2 |
| 7.5 | 2702.4 | 2548.6 | 1621.1 | 2155.5 | 2427.3 | 1922.0 |
| 10 | 2624.4 | 2440.0 | 2629.5 | 2067.2 | 2544.4 | 1849.3 |
| 15 | 2288.1 | 2383.2 | 1557.8 | 1999.7 | 2434.4 | 1779.1 |
| 20 | 2208.3 | 2226.4 | 2669.8 | 1894.2 | 2371.4 | 1699 |
| 30 | 2042.7 | 2256.6 | 1381.6 | 1811.8 | 2098.6 | 1614.8 |
| 40 | 2038.5 | 2392.9 | 1310.6 | 1751.6 | 2450.2 | 1540.9 |
| 50 | 1872.3 | 2017.6 | 1247.8 | 1718.9 | 2164.2 | 1482.5 |
| 60 | 1889.7 | 2000.8 | 1206.2 | 1615.1 | 2071.5 | 1452.9 |
| 80 | 1742.7 | 1902.4 | 1139.7 | 1502.4 | 2030.3 | 1404.3 |
| Time (min)[b] | 2) CC Al | 2) CC Co | 2) CC Fl | 2) MDCR Al | 2) MDCR Co | 2) MDCR Fl |
| 0.5 | 5352.6 | 2394.3 | 1726.7 | 3001.2 | 2312.3 | 1895.0 |
| 2.5 | 2507.5 | 2266.1 | 1909.3 | 3064.0 | 2386.2 | 2069.4 |
| 5 | 2439.4 | 2142.5 | 1921.0 | 2475.5 | 2309.7 | 2012.5 |
| 7.5 | 3290.0 | 2563.8 | 1828.2 | 2348.4 | 2186.9 | 1983.6 |
| 10 | 2232.0 | 2220.6 | 1921.7 | 2300.8 | 2087.1 | 1880.1 |
| 15 | 2125.5 | 2566.7 | 2680.8 | 2186.1 | 2112.7 | 1923.5 |
| 20 | 1949.7 | 2007 | 3345.1 | 2076.9 | 2132.9 | 1798.2 |
| 30 | 1818.9 | 2175.1 | 1629.3 | 1945.8 | 1936.7 | 1679.9 |
| 40 | 2154.6 | 2173.7 | 1553.5 | 1869.6 | 2358.1 | 1580.8 |
| 50 | 1642.4 | 2169.5 | 1483.5 | 1825.1 | 1986.1 | 1513.9 |
| 60 | 1578.6 | 2342.7 | 1379.7 | 1643.7 | 1792.2 | 1452.7 |
| 80 | 1572.7 | 2016.6 | 1310.0 | 1599.8 | 1829.2 | 1380.7 |

a) First replicates b) Second replicates

**Table S-5:** Weight-average molecular weight ($M_w$, g/mol) values for all turkey EPH samples.

| Time (min)[a] | 1) TC Al | 1) TC Co | 1) TC Fl | 1) MDTR Al | 1) MTDR Co | 1) MTDR Fl |
|---|---|---|---|---|---|---|
| 0.5 | 2623.5 | 4753.9 | 2639.1 | 3972.8 | 2595.9 | 3845.7 |
| 2.5 | 3060.8 | 3869.5 | 2976.4 | 5512.9 | 2780.4 | 3083.7 |
| 5 | 2437.9 | 3432.4 | 3004.4 | 3104.3 | 2816.5 | 6493.1 |
| 7.5 | 2552.2 | 4313.5 | 2322.6 | 5076.7 | 2706.8 | 2714.4 |
| 10 | 2647.7 | 3376.0 | 2610.1 | 2580.5 | 2845.3 | 2583.1 |
| 15 | 2388.2 | 2983.3 | 3843.6 | 4692.5 | 2658.2 | 4204.7 |
| 20 | 2055.8 | 4116.9 | 2061.0 | 2441.4 | 3001.2 | 2253.9 |
| 30 | 1927.1 | 3137.5 | 1835.6 | 2288.8 | 2649.1 | 2354.8 |
| 40 | 2053.8 | 3330.3 | 2932.0 | 2353.5 | 2649.1 | 1971.0 |
| 50 | 2015.6 | 2554.8 | 1550.7 | 2336.0 | 2366.5 | 1857.7 |
| 60 | 1908.4 | 2364.0 | 2176.5 | 2936.2 | 2383.3 | 1828.6 |
| 80 | 1601.6 | 2398.4 | 1437.0 | 1967.3 | 2320.2 | 1628.6 |
| Time (min)[b] | 2) TC Al | 2) TC Co | 2) TC Fl | 2) MDTR Al | 2) MDTR Co | 2) MDTR Fl |
| 0.5 | 3033.9 | 2438.1 | 3721.7 | 3940.4 | 2611.8 | 4806.0 |
| 2.5 | 2656.4 | 2395.5 | 2526.8 | 4968.4 | 2580.5 | 4129.6 |
| 5 | 3829.4 | 2204.0 | 2128.3 | 4175.9 | 3213.2 | 4574.5 |
| 7.5 | 2420.7 | 2358.7 | 2501.5 | 4897.5 | 2941.3 | 2944.2 |
| 10 | 3482.6 | 2351.0 | 2336.1 | 2503.3 | 2622.3 | 2514.2 |
| 15 | 2095.0 | 3234.2 | 2155.3 | 2291.3 | 2498.5 | 2365.3 |
| 20 | 1959.2 | 2021.3 | 2007.3 | 3966.9 | 2317.8 | 2201.5 |
| 30 | 2709.9 | 2044.6 | 1899.4 | 2131.1 | 2286.7 | 2065.7 |
| 40 | 3566.0 | 1968.6 | 2149.2 | 1977.1 | 2398.2 | 1934.8 |
| 50 | 2060.4 | 1913.6 | 1710.8 | 1881.7 | 2365.5 | 1810.4 |
| 60 | 2847.6 | 1901.1 | 1843.2 | 1823.1 | 2259.3 | 1757.7 |
| 80 | 1651.5 | 1928.8 | 1590.6 | 1699.2 | 2479.5 | 1640.6 |

a) First replicates b) Second replicates

**Table S-6:** Weight-average molecular weight ($M_w$, g/mol) values for combined EPH produces and relative size regions based on the size exclusion chromatograms.

| Product name[a] | $M_w$[b] | F1 (>15 amino acids)[c] | F2 (5-15 amino acids)[c] | F3 (2-5 amino acids)[c] | F4 (free amino acids)[c] |
|---|---|---|---|---|---|
| CC Al | 2483.3 | 15.18 | 47.06 | 24.90 | 12.86 |
| CC Co | 2281.7 | 14.96 | 51.33 | 22.30 | 11.42 |
| CC Fl | 1459.9 | 7.90 | 42.25 | 28.18 | 21.67 |
| MDCR Al | 2492 | 17.39 | 46.76 | 22.47 | 13.40 |
| MDCR Co | 1958.1 | 13.23 | 54.17 | 21.01 | 11.59 |
| MDCR Fl | 1333.8 | 8.20 | 40.07 | 27.69 | 24.03 |
| TC Al | 3699.9 | 24.02 | 43.48 | 21.62 | 10.88 |
| TC Co | 3276.8 | 25.45 | 46.80 | 18.70 | 9.05 |
| TC Fl | 2866.6 | 16.37 | 37.63 | 26.58 | 19.43 |
| MDTR Al | 3172.7 | 23.50 | 48.27 | 18.32 | 9.91 |
| MDTR Co | 3888.3 | 33.61 | 44.11 | 14.36 | 7.92 |
| MDR Fl | 3158.1 | 20.15 | 34.22 | 24.81 | 20.82 |

a) The combined product from the two hydrolysis replicates. b) $M_w$ of the combined replicates. c) The relative size region area based on the size exclusion chromatograms for the combined products.

ISBN: 978-82-575-1647-5
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no

105574 / ANDVORDGRAFISK.NO