

DOCTORAL THESIS

Development and Implementation of High Throughput Peptidomics for Microbial Studies

Georg Arju

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
72/2022

Development and Implementation of High Throughput Peptidomics for Microbial Studies

GEORG ARJU



TALLINN UNIVERSITY OF TECHNOLOGY
School of Science
Department of Chemistry and Biotechnology

This dissertation was accepted for the defense of the degree of Doctor of Philosophy in Chemistry on 25/11/2022

Supervisor: Prof. Raivo Vilu
Department of Chemistry and Biotechnology
Tallinn University of Technology
Tallinn, Estonia

Co-supervisor: Dr. Ildar Nisamedtinov
Department of Chemistry and Biotechnology
Tallinn University of Technology
Tallinn, Estonia

Opponents: Prof. Erich Leitner
Department of Analytical Chemistry and Food Chemistry
Graz University of Technology
Graz, Austria

Associate Prof. Koit Herodes
Faculty of Science and Technology
University of Tartu
Tartu, Estonia

Defense of the thesis: 19/12/2022, Tallinn

Declaration:

Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been submitted for doctoral or equivalent academic degree.

Georg Arju

Signature



European Union
European Regional
Development Fund



Investing
in your future

Copyright: Georg Arju, 2022
ISSN 2585-6898 (publication)
ISBN 978-9949-83-943-8 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9949-83-944-5 (PDF)
Printed by Auratrükk

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
72/2022

**Suure läbilaskevõimega peptidoomika
meetodite arendamine ja juurutamine
mikrobioloogilisteks uuringuteks**

GEORG ARJU



Contents

List of Publications	6
Author's Contribution to the Publications	7
Introduction	8
1 Literature Review	9
1.1 Peptides as a Chemical Class	10
1.2 Scientific and Technological Interest in the Analysis of Peptides	11
1.3 Experimental Design for Untargeted Peptidomics Experiments	12
1.4 Pre-Analytical Steps for Untargeted Peptidome Analysis	13
1.5 Instrumental Steps for Untargeted Peptide Analysis	14
1.5.1 Sample Introduction in LC-MS Analysis of Peptides	14
1.5.2 Peptide Separation by Liquid Chromatography	15
1.5.3 Peptide Analysis by Mass Spectrometry	19
1.6 Post Instrumental Steps	23
1.6.1 Data Pre-Processing	23
1.6.2 Statistical Analysis	23
1.6.3 Peptide Identification in Complex Matrices	24
1.6.4 Peptide Quantification	25
1.6.5 Peptide Derivatisation	25
2 Aims of the Dissertation	27
3 Materials and Methods	28
3.1 Sample Matrices	28
3.2 Sample Preparation	28
3.3 Instrumental	29
3.3.1 Liquid Chromatography	29
3.3.2 Mass Spectrometry	29
3.4 Data Processing and Analysis	30
4 Results and Discussion	31
4.1 Case Study I: Peptide Profiling and Screening During Cheese Ripening (Publications I and II)	31
4.2 Case Study II: Peptide Mapping and Screening of Protein Hydrolysate Consumption During Alcoholic Fermentation (Supplementary Study I and Publication III)	34
4.3 Case Study III: Untargeted Peptide Profiling During Wort Production (Supplementary Study II)	35
5 Conclusions	40
References	41
Acknowledgements	47
Abstract	48
Lühikokkuvõte	49
Appendix 1	51
Appendix 2	65
Appendix 3	81
Curriculum vitae	103
Elulookirjeldus	104

List of Publications

Main Publications

These publications form the basis of this thesis and are reproduced in the appendices with permission from the publishers:

- I **Arju, G.**, Taivosalo, A., Pismennoi, D., Lints, T., Vilu, R., Daneberga, Z., Vorslova, S., Renkonen, R., Joenvaara, S. (2020) Application of the UHPLC-DIA-HRMS Method for Determination of Cheese Peptides. *Foods* 9, 979. doi:10.3390/foods9080979

- II Xia, X., Arju, G., Taivosalo, A., Lints, T., Kriščiunaite, T., Vilu, R., Corrigan, B.M., Gai, Nan., Fenelon, M.A., Toin, J.T., Kilcawley, K., Kelly, A.L., McSweeney, P.L.H., Sheehan, J.J. (2023) Effect of β -casein reduction and high heat treatment of micellar casein concentrate on proteolysis, texture and the volatile profile of Emmental cheese during ripening. *International Dairy Journal* 138, 105540. doi.org/10.1016/j.idairyj.2022.105540

- III **Arju, G.**, Berg, H.Y., Lints, T., Nisamedtinov, I. (2022) Methodology for Analysis of Peptide Consumption by Yeast during Fermentation of Enzymatic Protein Hydrolysate Supplemented Synthetic Medium Using UPLC-IMS-HRMS. *Fermentation* 8, 145. doi.org/10.3390/fermentation8040145

Author's Contribution to the Publications

Contributions to the papers in this thesis are:

- I The author was involved in the study design, developed, and conducted the LCMS experiments and analyses, processed the data, interpreted the results, and wrote the manuscript.
- II The author was involved in the study design, developed, and conducted the LCMS experiments and analyses, and participated in processing the data, interpreting the results, and writing the manuscript.
- III The author was involved in the study design, developed, and conducted the LCMS experiments and analyses, and participated in processing the data, interpreting the results, and writing the manuscript.

Introduction

Advancements in big-data capture are required to aid scientists in better understanding processes in systems biology and the design of industrial bioprocesses. The requirement for high-quality big-data is driving the necessity for increased chromatographic and mass spectrometric selectivity, as well as detector sensitivity, dynamic range, and acquisition speed, which should be achieved along with increasing analytical throughput. As maximising analyte coverage increases the cost and complexity of associated targeted analysis, untargeted analysis plays an important role in the discovery and relative quantification of metabolites in biological systems. Thus, untargeted analysis acts as an important prerequisite step for the subsequent quantitative targeted analysis.

Peptides are a class of compounds that are analysed in both metabolomics, where they occur as native metabolites, and proteomics, where they are analysed as intermediaries for protein identification and quantification. Most of the recent peptidomics methodologies are affected by lower throughput and sample preparation, which results in diminished profiling of the lower molecular weight peptidome.

This thesis aims to provide an overview of the application of liquid chromatography, coupled with mass spectrometry for peptidome profiling and semi-quantitative screening in different fermentation matrices, such as cheese, synthetic grape must, and beer wort. Special focus was put on improving the throughput, ease of use and flexibility of peptidome profiling and screening in complex biological matrices. The summary of theoretical and practical knowledge given in the literature review, published articles and supplementary materials, present a basis and progression from a very sensitive nano-UHPLC¹ with data-independent acquisition workflow toward a higher-throughput workflow utilising low-flow UHPLC¹-coupled IMS²-enabled data-independent acquisition and simplified sample preparation. As a result, the reader is provided with a better understanding of critical correlations between chromatographic performance, data acquisition strategies and data processing workflows.

¹ Ultrahigh pressure liquid chromatography

² Ion mobility separation

1 Literature Review

Metabolomics and proteomics are fields of analytical chemistry that provide complementary information to other omics techniques, such as genomics and transcriptomics. While metabolomics focuses on smaller molecular weight compounds (typically up to 1500 Da), the advancement in genome sequencing has sparked the interest in the functional analysis of gene products, resulting in the establishment of proteomics (typically over 10 kDa) (Martini et al. 2021). However, the increasing interest in the analysis of peptides in their native form has resulted in the subdivision of peptidomics as a separate compound-class-centric field of analytics. Due to various practical reasons, discussed later, the analysis of small peptides has been recently attracting particular interest across many industries.

The two approaches of omics include targeted and untargeted workflows. According to Gerstman and Barshop (2018), targeted workflows aim to accurately quantify specific, known compounds within an established range of concentrations while untargeted workflows focus on all detectable analytes to statistically establish what compounds (if any) significantly differentiate samples or groups of samples. Subsequent analysis of the data generated by the untargeted analysis typically allows for the identification of the unknown compounds, but it can also provide relative quantitative information across sample groups. This may be useful for further targeted analysis of compounds of interest (Schrimpe-Rutledge et al., 2016). As the core principles of both metabolomics and proteomics significantly overlap with peptidomics workflows, they also introduce the associated benefits and inherent challenges of both into the peptide analysis (discussed in the dedicated sections) (Figure 1).

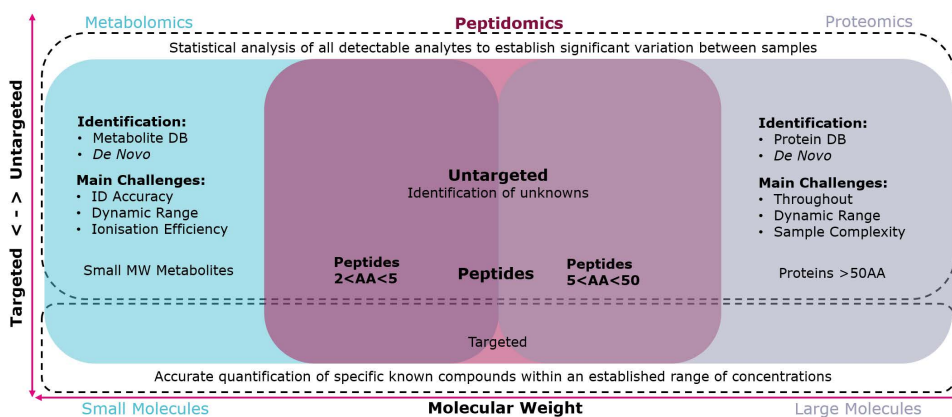


Figure 1. Overview of LC-MS-based omics and respective challenges.

Although the ultimate goal of metabolomics is to analyse the entire chemical space of a biological system, the current state-of-the-art technology is still far from reaching that goal (Pinu et al., 2019). While optimising the peptidomics workflows, this thesis focuses on highlighting and overcoming some of the challenges associated with the untargeted analysis of peptides in complex matrices.

1.1 Peptides as a Chemical Class

Peptides are considered one of the major classes of polymeric biomolecules consisting of peptide bond-linked amino acids (AA). Amino acids consist of amino (-NH₂), carboxyl (-COOH) and side chain (R) groups that are connected by an intervening carbon atom (Figure 2).

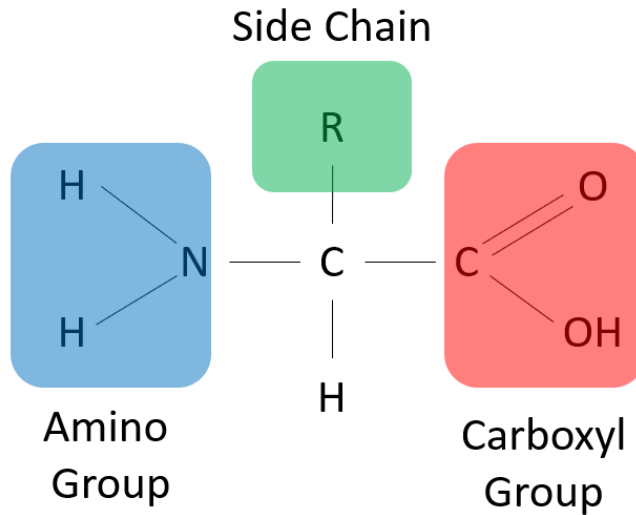


Figure 2. Structure of amino acid. Adapted from “ α -amino acid structure” by Aouniti et al., 2017.

The peptide bond is an amide-type covalent chemical bond linking two consecutive amino acids from the carbon of one amino acid (C-terminus) and the nitrogen of another (N-terminus), along a peptide or a protein chain (Figure 3).

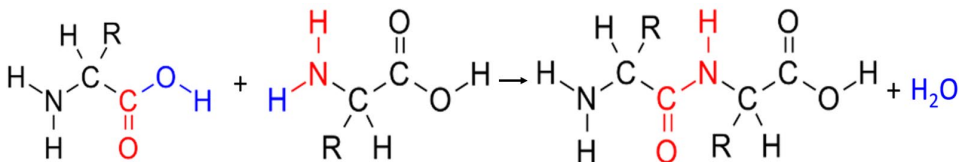


Figure 3. Formation of a peptide bond from two amino acids. Adapted from “peptide bond formation by a condensation reaction” by H. Panayiotou, 2004.

The diversity of amino acid properties, attributed to their side chains (Figure 4), and the very high degree of combinations possible result in peptides' diverse physicochemical properties and the ability to form various structural conformations, including the formation of cyclic peptides which makes their analytical determination and identification a challenge.

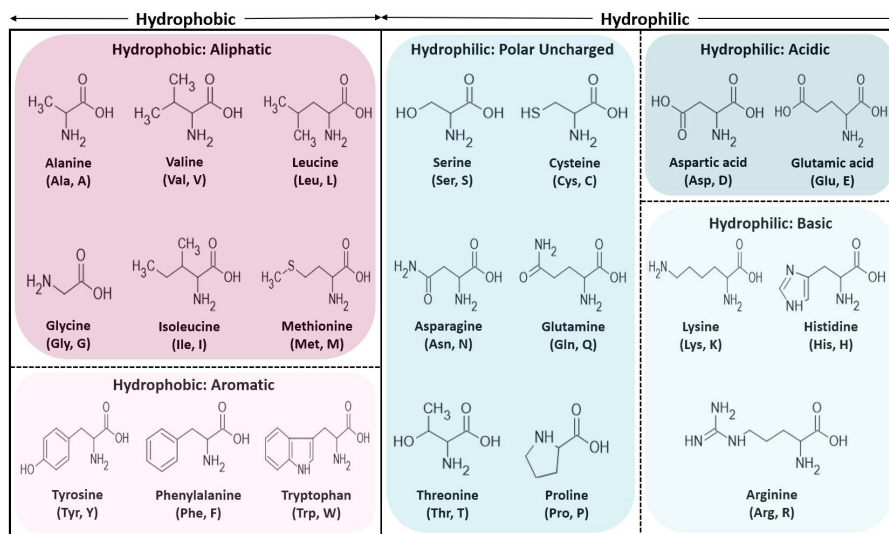


Figure 4. Amino acid classification and their physicochemical properties. Adapted from "amino acid chart" by K. Steward, 2019.

Finally, further complicating the peptide chemical class, there are over 300 types of modifications that may be associated with peptides including, oxidation, methylation, acetylation, phosphorylation, glycosylation and lipidation that frequently play important biological or technological roles (Zhao and Jensen, 2009).

1.2 Scientific and Technological Interest in the Analysis of Peptides

Peptides play many important roles in the processes and products of various industries. Among others, these include the pharmaceutical, food, feed, and biotechnology industries. Some examples which further justify the requirement for (untargeted) peptide analysis are given below.

Peptides as flavour compounds. Naturally and artificially occurring proteolytic mechanisms in foods and beverages (such as cheese, yoghurt, soy sauce etc.) produce taste-active peptides that are often attributed to the organoleptic properties of the final product. Although five taste categories (sour, salty, sweet, bitter and umami) are recognised, sweetness, bitterness and umami are considered the most important for food acceptance or rejection of the product (Temussi., 2011). Temussi (2011) summarised the classification of taste-active peptides as sour peptides (acidic AA rich), bitter peptides (hydrophobic AA rich), and peptides with little to no taste (balanced AA composition). Although there are no naturally occurring sweet peptides, synthesised short peptides (for example aspartame) play an important role as artificial sweeteners. It is important to note that the taste of peptides is not directly related to the taste of amino acids, and it is often affected by the conformation of the peptide. In addition to the hydrolytic enzyme-driven peptide-producing mechanisms, the heat- or acid-induced peptide bond cleavage of proteins during food processing can also result in the formation of peptides (Bikaki et al., 2021). For example, hydrophobic peptides, formed during coffee bean roasting and brewing, contribute to the distinct taste associated with coffee beverages.

Bioactive peptides. Peptides that can carry out specific biological functions, next to being an important source of amino acids, are referred to as bioactive peptides. One of

the mechanisms of the formation of the bioactive peptides is the proteolytic effect of food processing and namely by the activity of probiotic bacteria (Romero-Luna et al., 2022). Agyei and Danquah (2011) claim that the advantages of bioactive peptides, such as a wide action spectrum and low toxicity as well as high biospecificity, made them one of the prime research foci of the biopharmaceutical industry. Food-derived bioactive peptides, discovered in fermented foods as well as in both animal and human intestinal digesta, are reported to have a positive influence on digestive, endocrine, immune, nervous as well as cardiovascular systems (Moughan et al., 2014). For example, plant-derived peptides with an angiotensin-converting enzyme and pancreatic lipase inhibitory properties are reported to help modulate hypertension and diabetes (Urbizo-Reyes et al., 2021).

Peptides as a nitrogen source during fermentation. Alongside free amino acids, peptides can also provide an important assimilable organic nitrogen that, unlike inorganic nitrogen, can be utilised by a broad range of species of microorganisms. Peptones and yeast extracts, depending on the raw material and proteolytic enzymes used, can contain a wide variety of peptides which can be easily assimilated by industrially important microorganisms and thus, the study of peptides in such complex fermentation ingredients or of the proteolytic activity of various proteases used to produce them is yet another important application of peptidomics (Proust et al., 2019). In addition to industrial microbial cultivations, certain naturally occurring peptides in many food and beverage fermentations (beer wort, grape must, whisky mash etc.) also act as a valuable nitrogen source and thus, making peptide analysis even more important as mostly smaller peptides can be assimilated by yeast (Kevvai et al., 2016).

While overlapping with proteomics and metabolomics, the diversity of chemical properties of peptides as well as the wide scope of their applications, has justified peptidomics emerging as a separate subdivision of omics to tackle the challenges unique to the analysis of this class of compounds. The major challenges of peptidomics include the identification of short (2-4 AA) peptides as well as their quantification (discussed below), especially in complex natural matrices.

1.3 Experimental Design for Untargeted Peptidomics Experiments

Despite the proposals for various approaches in untargeted omics, a universal workflow to account for all challenges associated with the untargeted peptidomics pipeline is yet to be fully established and widely recognised (Sumner et al., 2007). Any untargeted omics experiment that aims to provide meaningful and statistically significant results consists of several critical steps that play an important role in the overall success of a study. Those steps can be generalised into four main parts: I) experimental design, II) pre-analytical, III) analytical, and IV) post-analytical stages. Proficiency in untargeted peptidomics comes from understanding and balancing the trade-offs intrinsic to any analytical methodology associated with each of the four steps.

The main goal of an efficient experimental design is to minimise the required study input while maximising the output. Higher throughput can be partially achieved by limiting types of experiments that would inherently provide little additional information or discriminatory power. Power analysis is a statistical tool that can provide valuable information on the number of samples or replicates required for analysis to achieve group differentiation at a desired significance level (Dicker et al., 2010; Nakayasu et al., 2021). Power analysis is especially useful in the case of large cohort studies, where running even one additional replicate per sample can result in days of extra instrument

usage time. Moreover, preliminary knowledge about the sample (e.g., protein sequence, chemical modification, or protease cleavage specificity) can also often provide valuable insight.

Although statistical analysis is typically one of the last steps in the untargeted workflow it is greatly affected by the experimental design. Thus, the background knowledge for a sample and a suitable number of biological and technical replicates are vital to ensure the best data quality possible.

1.4 Pre-Analytical Steps for Untargeted Peptidome Analysis

Even though numerous analytical methods, such as nuclear magnetic resonance, capillary electrophoresis, and various types of chromatographic techniques coupled to mass spectrometry (MS) are commonly utilised in untargeted omics, the most used technique to study peptides is liquid chromatography (LC) coupled to mass spectrometry (LC-MS). Hence, all pre-analytical steps discussed next are concerned with LC-MS-based analysis.

Sampling and storage. As with any other analytical technique, the untargeted peptidome analysis process starts with the collection of a sample and its preparation for the subsequent instrumental analysis. That includes representativeness of sampling, fixation, and storage of the sample before downstream sample preparation. Lyophilised peptides (or samples containing peptides) are generally considered stable when stored at - 20 °C in dry, anaerobic, and light-free conditions. However, additional precautions should be considered when storing peptides in solution or as a part of a complex biological sample. Namely, the number of freeze-thaw cycles should be minimised and thus, for repeated use samples should be aliquoted. Further, additional concerns for long-term storage of peptides are oxidation and residual proteolytic activity in many biological matrices, which might require storage at even lower temperatures ($\leq - 80$ °C) for full inactivation (Morris and Marchesi, 2016). Finally, non-specific binding of the peptides to various container surfaces (plastic, glass, and metal) can occur and thus, when possible, strategies to minimise sample loss should be in place to maximise peptide recovery (DeLano et al., 2021).

Sample preparation. The main goal of sample preparation for analysis is to ensure the sample's compatibility with downstream sample manipulations. As with any untargeted omics study, peptidomics typically aims to characterise as broad a peptide profile in a sample as possible, and thus, sample preparation might present a unique set of challenges. Due to the broad range of peptides' physicochemical properties (polarity, electrical charge, and molecular size) in the biological samples, an increased sample preparation selectivity for peptides with specific chemical properties will typically result in a diminishing global profiling coverage (Finoult et al., 2011). For example, commonly used peptide desalting and clean-up (solid-phase extraction, SPE) after tryptic digestion of proteins (used in proteomics) might not be well suitable to analyse smaller naturally occurring peptides due to differences in hydrophobicity and thus, potentially resulting in their loss during reversed-phase SPE. Peptides with up to 6 amino acid residues are typically fully dissolved in water unless the entire sequence of the peptide consists of hydrophobic amino acids, such as W, L, I, F, M, V and Y. However, longer peptides might require adjustment of pH or the addition of an organic solvent to ensure full solubility. Therefore, for untargeted peptidomics, the least invasive and balanced sample preparation that ensures sample solubility and compatibility with subsequent analysis is often preferred due to the least adverse effect on the sample composition.

Protein precipitation is a frequently used technique to remove peptides and proteins that, under reversed-phase (RP) and hydrophilic interaction chromatography (HILIC) conditions used for peptidomics, might otherwise precipitate during instrumental analysis, potentially resulting in loss of the sample or even damage to the hardware (Fic et al., 2010). However, the precipitation solvent and protocol should be carefully designed to avoid diminishing recovery of peptides due to sugaring-out, associated with precipitation of biological matrices rich in sugars (Wang et al., 2008).

Another sample preparation method that provides easy sample clean-up with minimal undesired side effects is ultrafiltration (Boukil et al., 2018). Ultrafiltration allows the analysis of specific molecular weight ranges, unobstructed by compounds outside of the desired molecular weight range. However, ultrafiltration should be used with some caution as it might not remove all compounds, precipitation of which might occur under the aforementioned conditions.

Lastly, chromatographic sample fractionation is often considered part of sample preparation, while it is an equally valid part of two-dimensional liquid chromatography (2DLC). In this thesis, chromatographic fractionation will be further discussed in the section on peptide separation by liquid chromatography.

1.5 Instrumental Steps for Untargeted Peptide Analysis

This thesis primarily focuses on the hyphenation of liquid chromatography to electrospray ionisation-based (ESI) high-resolution mass spectrometry (HRMS). Even after many decades of continuous research, improvement, and application of both liquid chromatography and mass spectrometry, there is still significant room for further improvement of sensitivity, linear and dynamic ranges, ion mobility resolution, and identification accuracy. The technicalities of each will be discussed in subsequent sections.

1.5.1 Sample Introduction in LC-MS Analysis of Peptides

Sample introduction is an intermediate step between sample preparation and instrumental analysis. Although there are two main ways of introducing the sample into the system, namely infusion and injection, the latter is used in combination with LC. Sample introduction should facilitate the transfer of the prepared sample into the LC system without causing sample/analyte alternations or loss.

As mentioned in the sample preparation section, one of the most prominent issues encountered during sample introduction (but which also applies to many subsequent sections) are non-specific interactions of analytes with the sample-wettable metal-based surfaces, such as storage containers, sample needles, transfer tubing as well as analytical columns (DeLano et al., 2021; Liu et al., 2022). For example, phosphorylated compounds, including but not limited to phosphopeptides, are known to non-specifically adsorb onto sample-wettable surfaces of the regular LC systems (Guimaraes et al., 2022). A common counteraction against this is the replacement of commonly used stainless-steel components with those from alternative materials, e.g., titanium-based and polyether ether ketone (PEEK)-lined components. Although titanium-based hardware shows improvement compared to stainless steel, titanium is still a metal and does not fully mitigate the issue. PEEK-lined hardware greatly minimises secondary interactions but suffers from reduced structural integrity compared to metal-based hardware and thus cannot be fully deployed in ultrahigh pressure applications. One of the most promising developments is the proprietary surface treatments that mitigate the undesired

interactions. One such treatment is MaxPeak high-performance surface (HPS) treatment by Waters (Milford, MA, USA). Surface treatment can not only be applied to the LC hardware but also to the inner surface of analytical columns showing great recovery of the intensity of such challenging compounds such as phosphopeptides (Isaac et al., 2022). HPS can also be applied to the sample container surfaces greatly improving the long storage recovery and sensitivity of “sticky” compounds (Liu et al, 2022).

Although not peptide-specific, another important aspect that is often overlooked is the specifics of various injection mechanisms. The main injection mechanism types employed in LC are either a fixed loop (FL) or a flow-through needle (FTN) (Gritti et al., 2022). Unlike FTN, FL is characterised by the sample needle not being part of the main mobile phase flow path. A sample is aspirated into a loop, that is isolated from the flow path during aspiration, which is then incorporated into the flow part via a fluidics valve position switch. In the FTN configuration, the needle is a part of the main flow path and acts as a sample loop, facilitating aspiration, storage, and transfer of the sample. Although serving the same purpose, both systems have notable differences. The first difference is the volume of sample needed with sample aspiration, which might be crucial for applications with a limited sample quantity. In FTN-based systems, the injected volume is typically the same or very close to the consumed sample volume. However, in FL-based systems, the consumed sample volume needs to be adjusted for the sample needle volume (typically 10-30 μL) which facilitates the transfer of the sample from the sample vial to the loop. The second difference is the increased system volume of FTN-based compared to FL-based and consequently, increased system dispersion. Minimisation of system dispersion typically results in increased chromatographic performance.

1.5.2 Peptide Separation by Liquid Chromatography

Three main types of LC-based applications in combination with MS can be used in peptidomics. First, flow injection analysis (FIA) is an application of LC-MS without any chromatographic separation and relies purely on the resolving power of the mass spectrometer (Nanita and Kaldon, 2016). Although this type of experiment is typically characterised by very high throughput with analytical runs often lasting as little as a few seconds per sample, it greatly suffers reduced peak capacity (maximum number of chromatographically resolvable peaks) and selectivity due to the lack of chromatographic separation (Sarvin et al., 2020). The second and most common type is single-dimensional chromatography which relies on the selectivity of a single column to separate analytes that are detected by the MS upon elution. Last, the most complex but showing the highest potential is multidimensional chromatography which combines multiple complementary separation mechanisms into one analysis.

Chromatographic selectivity considerations for peptide analysis. Similar to sample preparation for liquid chromatography, the three primary characteristics of peptides that determine their separation in LC are polarity, electrical charge, and molecular size (Bagdett et al., 2018). For untargeted peptidomics of small molecular weight peptides, the most often used separation methods are reversed-phase (Figure 5), which are often characterised by poor retention of polar analytes, and HILIC (Figure 6). Both reverse-phase and HILIC are used to separate analytes primarily based on their polarity.

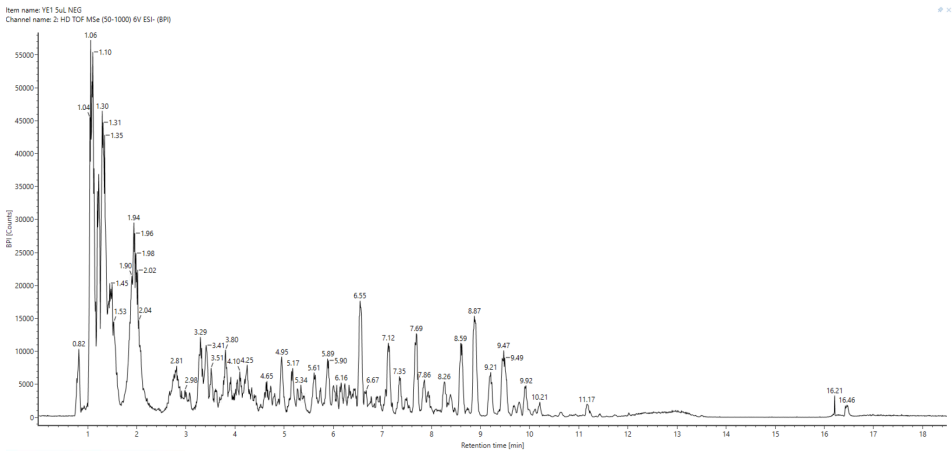


Figure 5. RP chromatogram of yeast extract. Poor retention (coelution and peak broadening) of polar analytes can be seen within the first 2.5 minutes of the chromatogram.

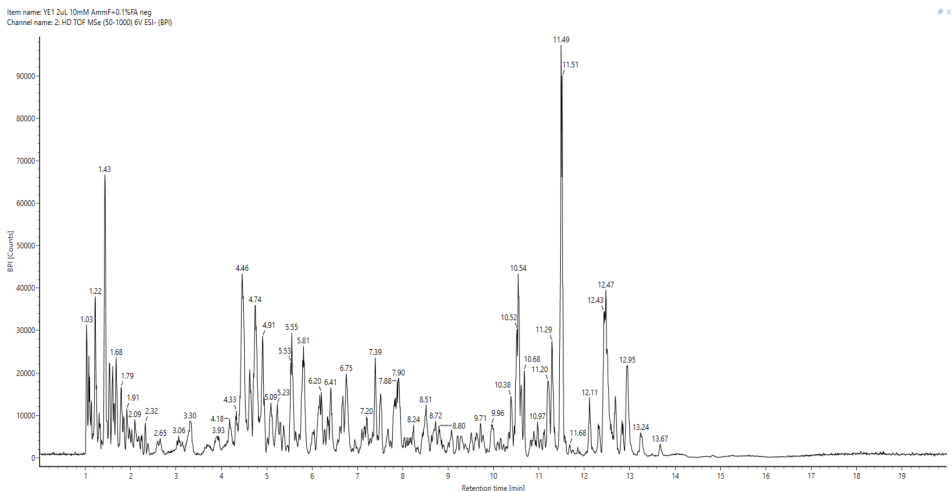


Figure 6. HILIC chromatogram of yeast extract.

Unlike RP, HILIC often benefits from the improved signal intensity of eluting analytes due to a higher ratio of organic solvent in the initial mobile phase (easier to evaporate compared to the highly aqueous mobile phase used for RP), which results in higher ionisation efficiency of analytes.

Although typically with lesser efficiency, the size exclusion chromatography (Figure 7) (Heusel et al., 2019) and ion exchange/exclusion chromatography can also be used for specific peptidomics applications (Valeja et al., 2015).

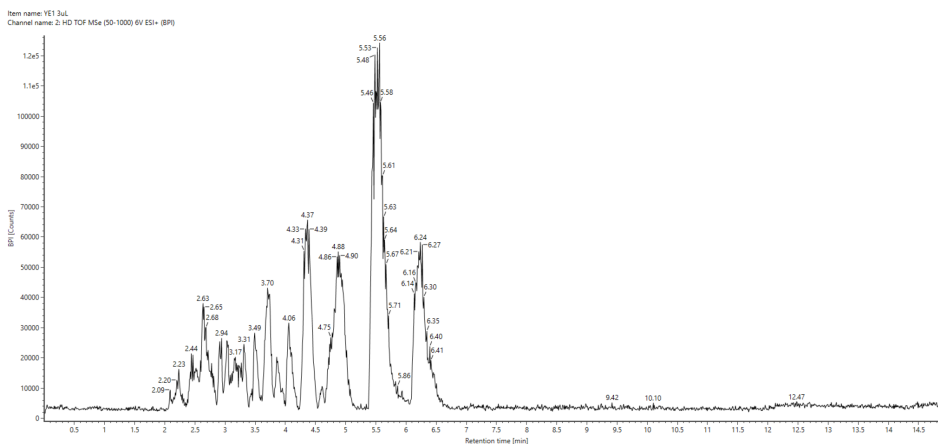


Figure 7. Size exclusion chromatogram of yeast extract.

Most frequently applied RP chromatography is based on C18 column chemistry that is often modified to withstand higher concentration aqueous mobile phases and provide additional retention of highly polar compounds (especially useful for smaller polar peptides) compared to unmodified (bare silica) C18. Among HILIC-type columns, the most common ones are unmodified HILIC, amide, amine, and sulfobetaine-based column chemistries. A wide range of column selectivity options coupled with specific conditions, such as mobile phase modifiers, provides scientists with a great toolkit to tackle the most challenging of samples. However, the complexity of biological samples still often exceeds the selectivity and peak capacity of a single column and thus, two-dimensional chromatography (2DLC) has a great potential for further selectivity, sample concentration and dynamic range improvement (Zhu et al., 2021).

Two-dimensional applications for peptide analysis. There are two main types of 2DLC: online and offline. First, online 2DLC implies the use of both columns within the same analytical run, i.e., peptides separated by the first column are then further separated by the second column. However, unlike two-dimensional gas chromatography (GCxGC) (Ralston-Hooper et al., 2008), where analyte trap/release is efficiently implemented by a thermal modulator (as a major compound elution mechanism primarily depends on the temperature), trap/release on online 2DLC can be achieved under very specific conditions that are typically not very applicable for untargeted analysis of smaller peptides (Liu et al., 2007). For example, coupling complementary RP and HILIC into online 2DLC is not feasible as the mobile phase with weak elution strength for the first dimension acts as a mobile phase with high elution strength for the second dimension, making a separation in the second phase ineffective. Offline 2DLC (introduced in the sample preparation section), on the other hand, benefits from intermediate sample manipulation, such as eluent removal and reconstitution with a suitable solvent for the second dimension, before further separating analytes in the second column. Even though such sample handling requires more steps and can potentially lead to increased variability, it is highly scalable to multiwell plate formats, remarkably increasing the method throughput. Moreover, stacked fractionation injections can be used to significantly concentrate compounds present in low concentrations.

Column dimensions considerations for peptide analysis. Continuous improvement of analytical columns' properties, as well as diversification of column chemistries and

dimensions, allows scientists to find solutions that are specifically tailored for peptide applications. Reduction in particle size to sub 2 micrometres resulted in the transition of LC-based methods from conventional high-pressure liquid chromatography (HPLC) to the so-called ultrahigh pressure liquid chromatography (UHPLC), that in comparison to HPLC is characterised by increased resolving power (theoretical plate number) even at higher linear velocities and increased column backpressure for the column with the same dimensions.

Particle size, column length, and column internal diameter (ID) are the major parameters when it comes to balancing between the sample throughput and the analyte coverage depth, as well as the amount of sample that can be efficiently loaded onto a column. Columns as short as 30 and 50 mm are typically used for applications where sample throughput is more important (e.g., for samples with low complexity) than the depth of the analyte coverage. Alternatively, columns upwards of 500 mm in length are frequently applied for metaproteomics to study samples with higher complexity to achieve the deepest profiling coverage possible (Roberg-Larsen et al., 2021). However, such applications are limited by extremely low throughput, where analytical runs can sometimes last over 10 hours. As a result, columns 100-150 mm in length are most frequently found in the methodological sections of untargeted metabolomics studies, providing the best balance between peak profile, sample capacity, resolution, and throughput.

In addition to particle size reduction, the improvement of manufacturing precision and accuracy resulted in the ability to produce and pack columns of smaller and smaller IDs. The transition from HPLC to UHPLC resulted in a reduction of analytical column ID from typically used 4.6-7.8 mm to 2.1 mm or narrower. The main positive change associated with the reduction of column ID is the reduction of mobile phase flow rate to achieve the same performance, which is highly beneficial for applications using MS, as ionisation efficiency is highly dependent on the ability to dissolve eluent in addition to the ability to sample the resulting ion plume. The typical UHPLC mobile phase flow rates used with 1-2.1 mm ID analytical columns are 100-1000 μL per minute. In contrast, nanoUHPLC-based applications are characterised by columns with a typical ID between 75 and 150 μm , resulting in flow rates as little as 100 nL per minute. An intermediate solution bridging the two is microUHPLC, which is typically represented by capillary columns with an ID of 300 to 500 μm and flow rates of 1-30 μL per minute.

For higher throughput applications, the higher mobile phase flow rate required to achieve a higher linear velocity with analytical columns results in not only a significant solvent consumption of expensive LC-MS-grade solvents, but also in a significant overspray as mass spectrometers' sampling efficiency is typically limited at around 100 μL per minute. Although benefiting from significantly increased sensitivity (Figure 8), the application of nanoUHPLC columns can often result in the requirement for longer equilibration times between runs and dedicated hardware for low-flow applications which are prone to leaks that are hard to detect. The nano emitters or sprayers, used for low-flow applications, due to their narrower inner diameter (1-10 μm), are also highly sensitive to a column overload and can result in a blockage due to too concentrated sample injected and may require additional sample load normalisation steps during sample preparation.

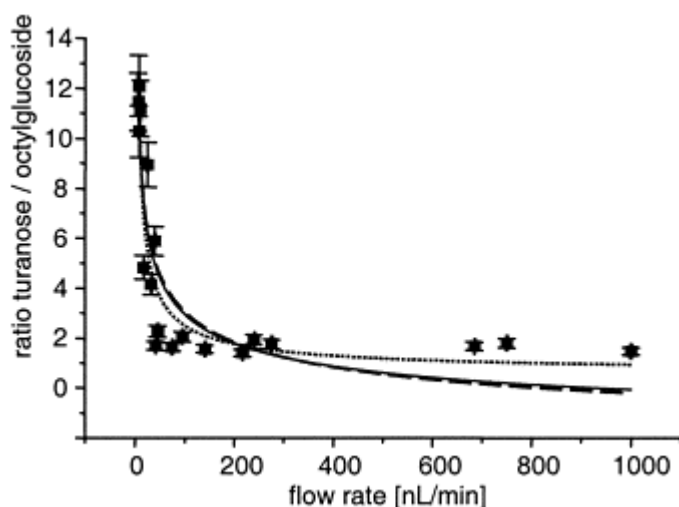


Figure 8. Signal response increase with a decrease in flow rate (Reprinted with permission from Schmidt et al., 2003. Copyright 2022 American Chemical Society).

MicroUHPLC applications, despite also requiring specialised hardware to optimise system volume and low flow rate stability, coupled with column length of 100-150mm, strike the best balance in terms of sample throughput, chromatographic and mass spectrometric performance.

Analysis of peptides is often characterised by shallow analytical gradients that start with nearly 100% aqueous mobile phase and reach 30%-40% organic mobile phase. A binary solvent manager suits this kind of application best, as it provides the highest mobile phase mixing accuracy in addition to minimal system volume (compared to a quaternary system), especially when coupled to a fixed loop sample manager.

All available technical options described above and variability in column chemistries dictate the importance of understanding the advantages and limitations of each configuration to be able to tailor a system for specific application needs.

1.5.3 Peptide Analysis by Mass Spectrometry

The mass spectrometer is a mass selective detector, suggesting that it does not only record a signal corresponding to the concentration of an analyte and its intensity but also records resolving mass spectra (MS^1) (Ramachandran and Thomas, 2021). The most common types of mass spectrometers used in untargeted analysis are high-resolution mass spectrometers that are based on time of flight (TOF) or Fourier transform (FT) technologies (orbitrap and Fourier transform ion cyclotron resonance). Like in chromatography, due to the ever-increasing requirement for chemical space profiling depth, in mass spectrometry, there is never enough selectivity, sensitivity and dynamic range (Liu et al., 2007). High-resolution mass spectrometers are characterised by higher resolving power, but also by the ability to measure accurate mass. The resulting accurately recorded and isotopically resolved ion species lay the foundation for the quantitation and identification process, namely deconvolution and generation of the compound's molecular formula. However, this alone is insufficient to identify a compound.

Common peptide fragmentation mechanisms. A combination of various mass analysers allowed for not only improvement of selectivity, but also sensitivity and dynamic range

(accurate detection and measurement of intensity within a broader range of concentrations or same spectrum) of many instruments (Valletta et al., 2021). The most frequent combination of HRMS is with a quadrupole (Q) mass analyser and a collision cell, allowing for various types of additional experiments, the most useful of which is the fragmentation of quadrupole isolated precursor ions, the fragment ions of which are resolved and detected by the high-resolution mass analyser, commonly referred as MS/MS or MS² (daughter scan) experiment. Recorded MS² spectra are used for interpretation of the molecular structure of a precursor ion or, in targeted analysis, also for quantitative purposes. Despite seeming similar, it must be pointed out that an MS² experiment does not equate to an MS/MS experiment. Unlike the MS² spectrum, which is simply a result of fragmentation (which may not involve precursor ion isolation), an MS/MS experiment refers to precursor ions isolation before fragmentation. In the case of specific instrument configurations (ion trap, FT or systems with multiple collision cells) multiple fragmentation experiments MS/MS/MS or MS³/MSⁿ can be carried out to further improve identification accuracy (Olsen and Mann, 2004).

There are multiple options to fragment an ion. When it comes to the analysis of peptides, the most common fragmentation methods are collision-induced dissociation (CID), higher energy collision dissociation (HCD, used in FT instruments), and electron transfer dissociation (ETD) (Ramachandran and Thomas, 2021). During CID and HCD, the kinetic energy of ions is increased via collisions with collision gas, which results in an internal conversion of kinetic energy into vibrational activation (Michalski et al., 2012). During ETD, positively charged ions are fragmented by electron transfer and, unlike CID and HCD, efficient application of ETD is limited to multiply charged species only (Kim et al., 2010). Although the application of ETD is limited for very small peptides (that commonly ionise as singly charged species), it still finds uses for the detection of labile post-translational modification in native oligopeptides (Huang et al., 2019) as well as complementary to the CID/HCD fragmentation mechanism for longer peptides. Instruments capable of MSⁿ can even be configured to combine ETD and CID/HCD. When it comes to differences between CID and HCD, it must be highlighted that for FT instruments under similar conditions, HCD tends to result in a capture of a higher number of smaller fragment ions. For instruments with a linear pass-through collision cell (such as QTOF), using a collision energy ramp with CID results in a wide range of fragments that would be characteristic of multiple collision energy states and thus, is preferential for high throughput applications, as it reduces the number of MS² experiments required for broader profiling.

The differences in the fragmentation mechanisms result in different peptide backbone fragmentation patterns. During peptide backbone cleavage up to 6 distinct site-specific fragment types can form (Figure 9). While CID and HCD typically result in the formation of b- and y-type fragments, a-type fragments, as well as the loss of ammonia (in RKNQ amino acids) and loss of water (in STED amino acids), can also be sometimes observed. ETD typically results in c- and z-type fragments.

Similar to CID, in-source fragmentation is referred to as any ion fragmentation which takes place before the collision cell. Although typically preferred to be avoided, in-source fragmentation is often applied on instruments lacking a collision cell to produce pseudo-MS² or increase the number of subsequent fragmentations using a collision cell.

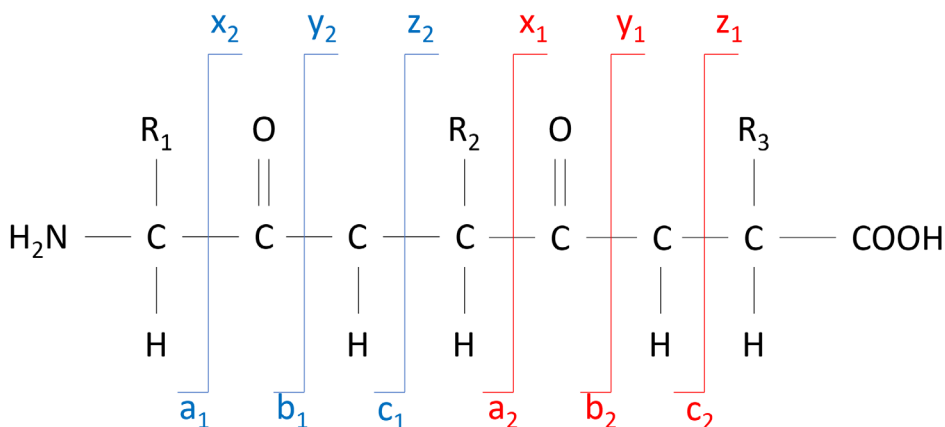


Figure 9. Peptide backbone fragmentation pattern. Adapted from “peptide fragmentation nomenclature” by C. Rath, 2011

Data acquisition strategies for untargeted analysis of peptides. For complex samples with unknown composition, automation and parallelisation of MS² spectra generation and processing were required to achieve a higher number of identified compounds per sample. An acquisition mode known as data-dependent acquisition (DDA) was proposed to help with automation (Bateman et al., 2014). DDA operates under the premise of automated triggering MS/MS experiment upon reaching the precursor ion (e.g., peptide) intensity threshold in MS¹ and hence, is called data dependent. State-of-the-art mass spectrometers usually allow for rapid acquisition of data from multiple MS/MS experiments within one cycle. Although this strategy provides clearer fragmentation information (due to active narrow isolation of precursor ions with quadrupole before fragmentation), it often cannot keep up with sample complexity and modern separation techniques. Every additional MS/MS event requires additional time that extends the overall cycle time thus reducing acquisition speed. Although the required data point number per chromatographic peak in the untargeted analysis is less strict than in targeted analysis, an accurate and reproducible peak integration must still be ensured. Reduction of acquisition speed limits the high throughput separation applications that are typically characterised by faster separation and thus, narrower chromatographic peaks. It is not only challenging to combine high throughput with DDA, but also the possibility of missing an MS/MS experiment due to limited time within the chromatographic peak complicates the matter even further. Moreover, within biological systems, there is a great likelihood of variability in the analyte concentration that leads to a potential situation where the concentration of peptides might remain below the MS/MS-triggering threshold resulting in a false negative result in some of the samples. Additionally, unlike TOF, FT-based systems’ resolving power depends on the scanning speed, i.e., the higher acquisition speed, the lower instrument’s resolution, thus diminishing the highest advantage of FT-based systems (resolving power).

Although HRMS is characterised by the higher resolution of the main analyser, the precursor ion isolation is still carried out by a typically low-resolution analyser, such as a quadrupole or ion trap. This can often result in coeluting compounds with similar near-isobaric mass-to-charge ratios (m/z) being co-fragmented and interfering with each other’s identification (chimeric spectra). Moreover, after a successful MS/MS experiment the precursor ion is typically excluded for the following 30-60 seconds after acquisition

resulting in an increased chance of not triggering the MS/MS experiment for closely or co-eluting peptides with similar or identical m/z .

Considering the limitations of DDA, data-independent acquisition (DIA) has been introduced as an alternative (Jiang et al., 2022). Unlike DDA, native DIA does not rely on the threshold-based triggering of quadrupole-based isolation and hence, is not dependent. DIA operates by sequentially alternating between non-fragmenting collision energy (low energy) and fragmenting collision energy (high energy). In contrast to DDA, this approach ensures no missed data and benefits from a wider dynamic range, but it suffers from inaccuracies during the alignment of fragment and precursor ions as DIA relies exclusively on the accurate integration of the peak's lift-off, apex, and touch-down points. Further improvements to the DIA experiment include increased selectivity for complex samples to reduce dependency on chromatography. Two main strategies such as stepping/scanning quadrupole and ion mobility separation (IMS) were introduced and discussed next (Puyvelde et al., 2022).

DIA selectivity enhancement strategies. Quadrupole-enhanced DIA is a hybrid between DDA and native DIA (Chen et al., 2021). Instead of isolating a narrow m/z window (1-4 m/z) that has been selected by software based on MS^1 scan precursors, the quadrupole indiscriminately steps or scans over the user-defined m/z range with an application optimised transmission window (typically 10-50 m/z) providing higher MS^2 spectral clarity than native DIA. It is important, however, to highlight the difference between stepping (SWATH, Sciex AB) and scanning quadrupole (SONAR, Waters) DIA approaches. Although quadrupole-enhanced DIA modes decrease the overall duty cycle on an instrument (number of ions entering the MS vs number of ions reaching the detector), as stepping and scanning quadrupoles affect both MS^2 and MS^1 scans, it greatly improves spectral clarity and identification accuracy, thus profiling depth for application without sample quantity limitations. In the case of complex samples, where the coeluting analytes can often fall within the same quadrupole transmission window, SONAR provides improved discrimination power due to the scanning of the transmission window. In practice, during quadrupole scanning, unless isotopic patterns of two coeluting compounds overlap, there is a higher likelihood of the quadrupole transmission window reaching the analyte with lower m/z before it reaches any subsequent analytes of higher m/z .

The ion mobility (IM) presents a complementary separation mechanism, which unlike quadrupole-based selectivity enhancement approaches, utilises alternative to mass-to-charge ratio ion properties (Puyvelde et al., 2022). Ion mobility generates and measures the effect that an electric field has on the ions moving through a gas phase. According to Gongyu et al. (2020), the ion mobility separation (IMS) mechanisms that are frequently applied to HRMS can be subdivided into 3 main categories:

- Temporally dispersive
 - Drift tube IMS (DTIMS)
 - Travelling wave IMS (TWIMS)
- Field dispersive
 - Trapped IMS (TIMS)
- Spatially dispersive
 - Field asymmetric IMS (FAIMS); cylindrical geometry
 - Differential mobility separation (DMS); planar geometry

Unlike the spatially dispersive ion mobility separation mechanisms, both temporally and field dispersive ion mobility separations that use a trap/eject mechanism can be calibrated to measure the ions' collision cross section (CCS). Not only are DT-/TW-/TIMS used to increase the peak capacity and spectral clarity while maintaining or improving the duty cycle (unlike the quadrupole enhanced DIA and spatially dispersive IMS), but the CCS has also started gaining additional value as an extra identification qualifier (Tejada-Casado et al., 2018). In some cases, the use of IMS also results in the improvement of the method's sensitivity, due to background noise reduction after the application of the IMS filter.

All IMS-MS mechanisms to a lesser or greater degree can be tuned to allow the transmission of ions of a particular charge state. This is frequently applied for the analysis of oligopeptides with great success by excluding the transmission of singly charged ions and thus, eliminating the interfering background. Recent studies demonstrate the application of FAIMS combined with DIA for longer peptide identification, resulting in a higher number of proteins identified (Bekker-Jensen et al., 2020). However, it must be pointed out that the same studies also show that the operation of FAIMS in combination with DDA results in a reduction of peptide and protein identifications.

The attempts to combine IMS with quadrupole-enhanced DIA have been made. Most notable was the recently introduced parallel accumulation serial fragmentation combined with the data-independent acquisition (diaPASEF) by Matthias Mann's group in collaboration with Bruker (Meier et al. 2020). The method capitalises on the correlation between mass to charge ratio and ion's mobility to further improve the duty cycle of the instrument while proving an even greater increase for the precursor identification specificity. However, despite all the recent advancements, there is still a significant room for further improvement in selectivity and acquisition performance.

1.6 Post Instrumental Steps

1.6.1 Data Pre-Processing

Pre-processing of HRMS data plays an important role to correct mass axis deviations and prepare data for subsequent analysis. Pre-processed peak intensities are often normalised between different analytical runs (either using spiked standards or other kinds of housekeeping ions) and chromatographic profiles are aligned based on retention time (RT). The resulting data is then used to co-detect peaks based on m/z , RT and, if available, CCS or drift time across the study sample set. As DDA is known to suffer from data completeness issues, "match between runs" algorithms were developed to try to mitigate the problem by extrapolating the data from the peaks that triggered MS/MS to the peaks that were below the triggering intensity threshold.

With a high chemical complexity of the sample comes an even higher complexity of the data. This is especially true for data recorded as profile (opposed to centroid) and containing IMS data. Working with such information-rich data requires high computational resources in addition to application-specific workflows for both alignment and identification.

1.6.2 Statistical Analysis

Depending on the sample, the data generated from untargeted peptidomics analysis can be often very complex, for example, resulting from a high number of proteins with unknown sequences hydrolysed by proteases with unknown specificities. Steps to reduce

the data complexity which in turn reduces processing time should be taken. Statistical analysis and subsequent data filtration can help with multiple aspects associated with untargeted workflows. First, statistical methods such as ANOVA are often used to exclude the differences between the mean values of different groups that are not statistically significant or are random (Nakayasu et al., 2021). The use of principal component analysis and orthogonal/partial least squares-discriminant analysis (O/PLS-DA) allows feature clustering and visualisation of grouping differences (Windarsih et al., 2022). Not only does it help to visualise the differences and grouping, but also, in the case of using pooled QC sample, ensures the validity of the experimental design. Blank samples are invaluable assets in eliminating background from the analysis. As previously mentioned, power analysis is used to calculate the minimum number of samples required to detect a statistically significant difference between groups.

1.6.3 Peptide Identification in Complex Matrices

As the ultimate goal of peptide profiling (i.e., peptide identification) is to elucidate its amino acid sequence, three main peptide identification strategies are commonly used. Firstly, peptide mapping or database search is effectively used when peptides originate from a protein the sequence of which is known. Considering protease/peptidase specificity (if known), unknown peptides are matched against all theoretical peptides that can be a result of the hydrolysis of a given protein(s). Secondly, in case of unknown protein sequence or unelucidated protease/peptidase specificity *de novo* sequencing can be carried out that attempts to predict oligopeptide composition purely based on the fragment ions of particular precursor ions. Thirdly, in the case of small peptides, typically 2-5 AA, identification can often be carried out via the existing *in-silico* databases (ChemSpider Peptides and METLIN) that have been generated before the analysis.

One of the biggest challenges in peptide identification is insufficient identification accuracy for unambiguous identification assignments. Unlike higher molecular weight oligopeptides, which due to a higher number of peptide bonds tend to ionise with a higher charge state ($\geq 2+$), smaller molecular weight peptides (2-5 AA) most commonly ionise as singly charged ions. The singly charged species produce fewer selective fragments, which leads to a reduced identification accuracy, which unlike longer peptides, renders short peptides not highly applicable to *de novo* sequencing. Fragmentation of quadrupole isolated precursor ions at various fixed collision energies can provide defining fragment ion ratios, resulting in a higher degree of confidence in identification results, but such experiments are ultimately aimed at identification confirmation of a small number of molecules against standards of peptide candidates and are not very applicable to large studies where identification of hundreds or thousands of peptides is a priority.

Ion mobility separations capable of measuring the ion's CCS provide an extra qualifier to reduce false positive, potentially separate isobaric peptides and result in higher confidence in identification results. While CCS measurement is a relatively novel application and empirical CCS libraries are being constantly updated with new entries (including different adducts), a computational approach provides a more cost-efficient alternative to profiling thousands of compounds (Zhou, Tu and Zhu., 2018).

Finally, another useful tool for validation of identification is the adducts (including neutral loss) formed during ionisation (Liu et al., 2015). For example, in positive electrospray ionisation adducts would commonly include protonation of various degrees, and formation of sodium, potassium, and ammonia adducts. Neutral loss, such as loss of water, is frequently observed and can be either result of an analyte-specific ionisation

mechanism or a result of unoptimised voltages leading to in-source fragmentation. In some cases, two conflicting adducts can result in misidentification. For example, even though the loss of water is a likely event during the ionisation of some organic and amino acids, it is not very common during peptide ionisation. For example, a protonated tripeptide Ala-Pro-Leu and protonated with a loss of water tripeptide Ser-Val-Leu will result in the same deconvoluted molecular formula of $C_{14}H_{27}N_3O_5$ and practically indistinguishable fragmentation patterns. Hence, loss of water should be considered for peptide identification with precaution.

1.6.4 Peptide Quantification

Quantification presents the last, but, perhaps, one of the most challenging processes in the untargeted peptide analysis of biological samples. The analyte concentration ranges are often greater than the dynamic range of any modern MS instruments (Beri et al., 2015). Accurately measuring the concentration of analytes in one analysis is very challenging as the concentrations may differ up to several orders of magnitude.

Mass spectrometry is intrinsically not quantitative. However, using internal (including isotopically labelled) and external standards, it can be calibrated to produce an absolute concentration value for a particular analyte of interest. This forms the basis of targeted analysis, where all compounds of interest are measured against a calibration curve of response ratios between the analyte and an isotopically labelled version of the analyte (Feteisi et al., 2015). However, this approach is not the most cost-efficient and with an increasing number of analytes increases the associated cost of analysis. Relative quantification or screening (for previously identified/profiled compounds) typically uses a normalised sample load combined with a spike with a set of non-endogenous to the sample standards to account for inter- and intra-sample variation (Kultima et al., 2009; Graw et al., 2020). Analyte response values normalised against spike standards provide relative amounts of the same analyte that can be used for comparison between samples. The quantitative accuracy of this approach is highly limited compared to the targeted workflow, but the use of relative abundances still allows for making decisions on the statistical importance of the chemical space of the sample.

Label-free quantification (LFQ) used in proteomics relies on the quantification of post-digest oligopeptides, ionisation efficiency of which levels out with an increase in peptide length (>6 AA). Shorter peptides have been demonstrated to have a vast difference in ionisation efficiency and are not applicable for LFQ (Liigand et al., 2018).

In recent years, computational approaches to estimate ionisation efficiency are being proposed. The studies are, however, frequently limited by the experimental design of the studies and a lot of further research is required to utilise similar approaches to mainstream applications (Liigand et al., 2018).

1.6.5 Peptide Derivatisation

Derivatisation of short peptides could potentially allow tackling some of the shortcomings of identification and quantification workflows. Qiao et al. (2012) claim that derivatisation of unmodified peptides with hydrophobic and/or basic groups will result in not only improved reversed-phase chromatographic performance (increased hydrophobicity resulting in higher column retention of polar peptides) but also increased ionisation efficiency (hydrophobic peptides undergo gas-phase protonation more easily). Moreover, neutralisation/derivatisation of the negative charge of the carboxyl group via

derivatisation could further improve ionisation efficiency as well as increase the peptide's charge state (Frey et al., 2013).

Selective derivatisation of the C- and/or N-terminus of a peptide could result in the generation of specific fragments that would further improve identification assignment accuracy (Wang et al., 2009). However, this approach would require extended sample preparation to account for kinetics experiments with time series, where derivatisation would be required after each time point. Moreover, a dedicated software workflow would need to be created to account for peptide modifications caused by derivatisation.

2 Aims of the Dissertation

The main objective of this thesis was to investigate and further develop methodologies for the analysis of peptides in various biological samples to improve the profiling coverage and throughput of untargeted peptides screening.

The developed workflows were tested for the following applications and categorised into three case studies:

- Case Study I: Profiling of peptide composition during cheese ripening (Publication I and Publication II)
- Case Study II: Characterisation of protein hydrolysate composition produced with the protease of unknown specificity and screening of peptide consumption by yeast during alcoholic fermentation (Publication III, Supplementary Study I)
- Case Study III: Untargeted peptide analysis to study the effects of different brewers' malts and mashing regimes on wort peptide profile (Supplementary Study II)

3 Materials and Methods

More specific information regarding the methods used in each published study can be found in the corresponding publications (Appendix 1, Appendix 2, and Appendix 3). The following sections are presented as a consolidated part of the materials and methods, providing a broader overview, as well as outlining important aspects of supplementary studies.

3.1 Sample Matrices

In Publications I and II, peptide profiling of two different kinds of cheese (Gouda and Emmental) was developed and carried out in response to different ripening conditions.

In Publication III and Supplementary Study I, a methodology for analysis of the composition of small peptides in protein hydrolysates prepared with protease(s) of unknown specificity was developed and used to monitor their consumption by yeast during fermentation of hydrolysate spiked synthetic grape must.

In Supplementary Study II, the effect of different mashing regimes of malt was explored on peptides composition in the resulting worts.

3.2 Sample Preparation

In Case Study I (Publications I and II), sample preparation of cheese water-soluble extracts (pH 4.6) for LC-MS analysis was accompanied by protein precipitation with acetonitrile and the following removal by centrifugation. Unlike in Publication I, in Publication II the requirement to normalise the sample load on the column was mitigated by using a larger internal diameter analytical column with higher load tolerance as well as a regular ESI source and thus, sample load normalisation was omitted.

In Case Study II (Publication III and Supplementary Study I), sample preparation of synthetic grape must and bovine serum albumin hydrolysate for LC-MS analysis was carried out by ultrafiltration and protein precipitation with acetonitrile and centrifugation. In Supplementary Study I, the enzymatic hydrolysis of bovine serum albumin (40 g/L) was performed in 10 mL potassium phosphate buffer (0.5M, pH 7) at 45°C for 24 hours. Three commercial proteases: Corolase® 7089, Corolase® 8000 and Corolase® APC (AB Enzymes, Darmstadt, Germany), and combinations thereof were studied. Samples preparation of BSA hydrolysate consisted of ultrafiltration and protein precipitation, as used in Publication III.

In Case Study III (Supplementary Study II), different malts (Château Pilsen 2RS, Château Pale Ale and Château Distilling) were acquired from La Malterie du Château SA (Verviers, Belgium). The worts were produced by mashing 440 grams of milled malt in 2 litres of distilled water. For each malt, three different mashing regimes were applied (Figure 10): mash 1 (67°C for 60 min); mash 2 (20 minutes at 50°C followed by 60 minutes at 67°C); mash 3 (20 minutes at 40°C, then 20 minutes at 50°C, followed by 60 minutes at 67°C). After mashing, the worts were boiled for 60 minutes and then cooled down to room temperature on ice. The worts were subsequently centrifuged at 4667 g to remove insoluble debris.

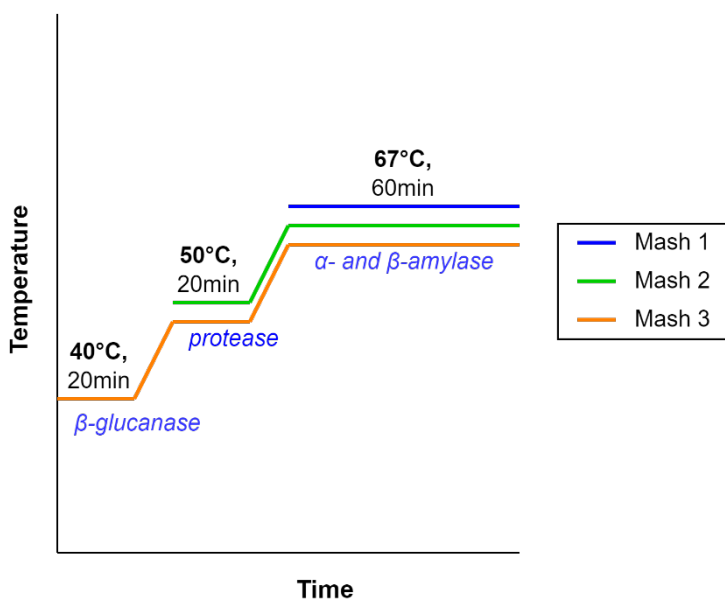


Figure 10. Temperature profiles of the different mashing regimes used in Supplementary Study II.

Due to the low peptide concentrations in the obtained worts, the samples required concentration before further analysis by HRMS. For concentration via evaporation, viscosity-increasing fermentable sugars were removed by fermentation of worts with a yeast strain in which the genes encoding for peptide transporters were knocked out (Becerra-Rodríguez et al., 2021). To promote dextrin hydrolysis, glucoamylase (0.1% w/w, MEGA PACIFIC TECHNOLOGY INC, Arcadia CA, USA) was also added.

3.3 Instrumental

3.3.1 Liquid Chromatography

In Publication I, a nanoUHPLC-FL (NanoAcquity, Waters Corporation, Milford, MA, USA) system was used. The LC system operated in forward-trap mode. A short column with a larger diameter and particle size (Acquity UPLC[®] Symmetry C18 Nanoacquity 10 k 2 g V/M, 100A, 5 μ m, 180 μ m \times 20 mm, Waters Corporation, Milford, MA, USA) was used to trap peptides that were then analysed with a nanoUHPLC column (Acquity UPLC[®] M-Class HSS T3, 1.8 μ m, 75 μ m \times 150 mm, Waters Corporation).

In Publication II and III, as well as Supplementary Study I and II, an analytical UHPLC-FL (I-Class Plus, Waters Corporation, Milford, MA, USA) using Acquity UPLC[®] HSS T3 (1.8 μ m, 1 \times 150 mm, Waters Corporation) was used. Chromatographic conditions used in Supplementary Study I and II were identical to Publication III.

3.3.2 Mass Spectrometry

For Publication I, a MALDI SYNAPT G2-Si mass spectrometer (Waters Corporation, Milford, MA, USA) with NanoLockSpray exact mass ionisation source was used. The instrument was operated in MS^E (DIA mode) as the travelling wave ion mobility separation option was not enabled.

For the remaining publications and supplementary studies, Vion IMS QToF mass spectrometer (Waters Corporation, Milford, MA, USA) was used and operated in HDMS^E, TWIMS-enabled DIA.

3.4 Data Processing and Analysis

Data processing was carried out using various software packages and workflows.

In Publication I, MassLynx (Waters Corporation, Milford, MA, USA) and Progenesis QI for Proteomics (Nonlinear Dynamics, Newcastle, UK) were used to acquire, process, and analyse the raw data.

In Publication II, UNIFI large molecule package (Waters Corporation, Milford, MA, USA) in conjunction with in-house data analysis and visualisation scripts written in the Python™ programming language (Python Software Foundation, Wilmington, USA) were used to acquire, process, analyse and visualise the raw data.

In Publication III and Supplementary Studies I and II, UNIFI Large molecule package (Waters Corporation, Milford, MA, USA) and Progenesis QI (Nonlinear Dynamics, Newcastle, UK) in conjunction with in-house data analysis and visualisation scripts written in the Python™ programming language (Python Software Foundation, Wilmington, USA) were used to acquire, process, analyse and visualise the raw data.

4 Results and Discussion

The results presented below are based on three publications and two supplementary studies. Considering the variability of the aims of the studies and the experiment design used, a summary of the main outcomes and the scientific/technological impact concerning methodology and respective outcomes are given in Table 1. In Case Study I (Publications I (P1) and II (P2)), the methodology development and optimisation of the analytical throughput for the cheese peptide profiling during ripening were carried out. In Case Study II (Publication III (P3) and Supplementary Study I (SS1)), peptide profiling of BSA hydrolysates prepared with different proteolytic enzymes was carried out. In Publication III, peptide profiling of BSA proteolytic digest prepared with the industrial protease of unknown specificity was carried out, including an assessment of consumption of these peptides by yeast during fermentation in synthetic must. In Case Study III (Supplementary Study II (SS2)) we explored the effect of different malt types and mashing regimes on the resulting peptide composition in the wort.

Table 1. Outcome and impact of the studies.

	Case Study I (Cheese) P1 and P2	Case Study II (Synthetic Wine) P3 and SS1	Case Study III (Wort) SS2
Study Aim	Peptide profiling during cheese ripening	Peptide consumption screening by yeasts	Untargeted peptide profiling in brewer's wort
Instrumental Steps	P1: 70-min nanoUHPLC-DIA-HRMS	P3 + SS1: 18.5-min UHPLC-IMS-DIA-HRMS	SS2: 18.5-min UHPLC-IMS-DIA-HRMS
	P2: 48.5-min UHPLC-IMS-DIA-HRMS		
Post-Instrumental Steps	P1: Progenesis QI for Proteomics	P3 + SS1: UNIFI LMP/Progenesis QI and Python Script	SS2: Progenesis QI Fully Untargeted Workflow
	P2: UNIFI LMP and Python Script		
Study Outcome	P1: Increased throughput of nanoUHPLC peptide profiling	P3: Increased throughput of peptide consumption screening	SS2: Developed methodology for differentiation of worts prepared from different malts and mashing regimes
	P2: Increased throughput of UHPLC peptide profiling	SS1: Increased throughput of protease specificity profiling	

4.1 Case Study I: Peptide Profiling and Screening During Cheese Ripening (Publications I and II)

Case Study I aimed at investigating the improvement of the analytical throughput of cheese ripening characterisation using peptide profiling via different approaches: application of nanoUHPLC (75 μm ID, flow rate of 0.3 $\mu\text{L}/\text{min}$) combined with data-independent acquisition mode (Publication I) and a narrow bore analytical column on UHPLC (1 mm ID, flow rate of 100 $\mu\text{L}/\text{min}$) was used in combination with ion mobility enabled DIA mode to characterise cheese proteolysis (Publication II).

Chromatographic conditions relative to sensitivity and sample preparation. Although the column load on the nanoUHPLC column (200 ng) was significantly lower than on the analytical column, the resulting base peak intensity (BPI) profile from nanoUHPLC (4.5e6) was more than 25-fold higher than the BPI profiles acquired using the analytical column

(1.8e5) (Figure 11). However, a significantly higher loading capacity of the analytical flow system resulted in simplification of the sample preparation that did not require sample concentration normalisation and thus, no further back-calculation to the original sample concentrations.

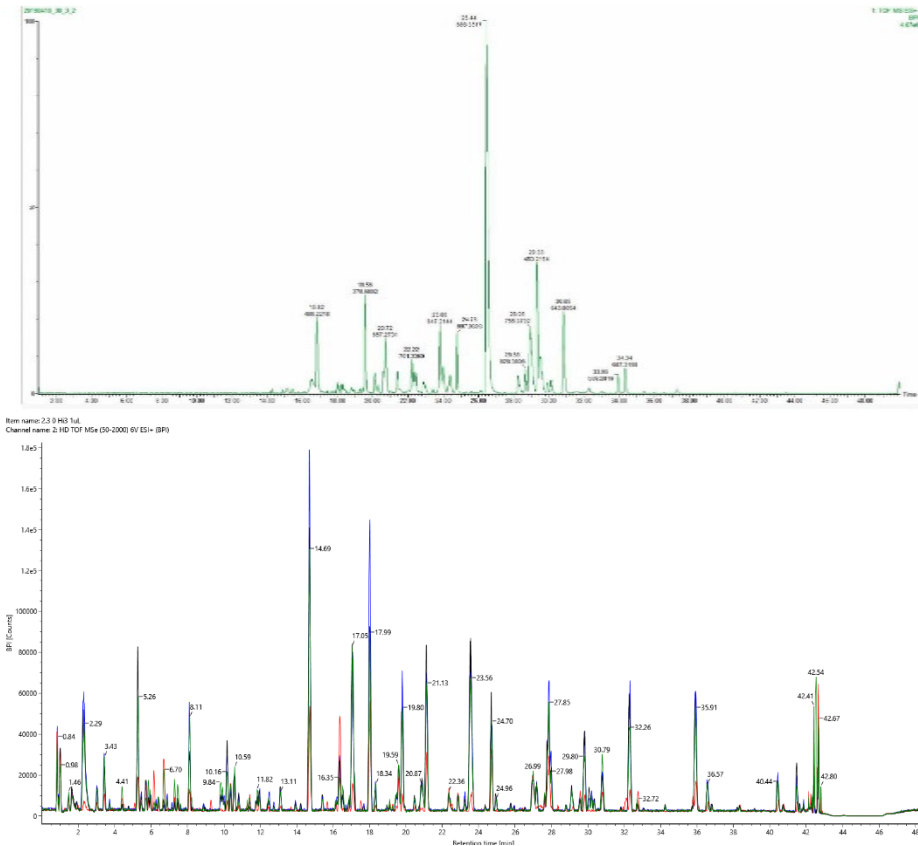


Figure 11. Comparison of cheese BPI profiles (retention time vs intensity): Top- nanoUHPLC-HRMS (Publication I); Bottom- UHPLC-IMS-HRMS (Publication II).

Peptide retention profiles. The overlay of total ion current (TIC) chromatograms of peptide profiles acquired in Publication II (Figure 12) better illustrates not only early eluting species that might have been lost with nanoUHPLC (due to loading into a trap column, Figure 11 Top) but also far better gradient utilisation to analyse the samples.

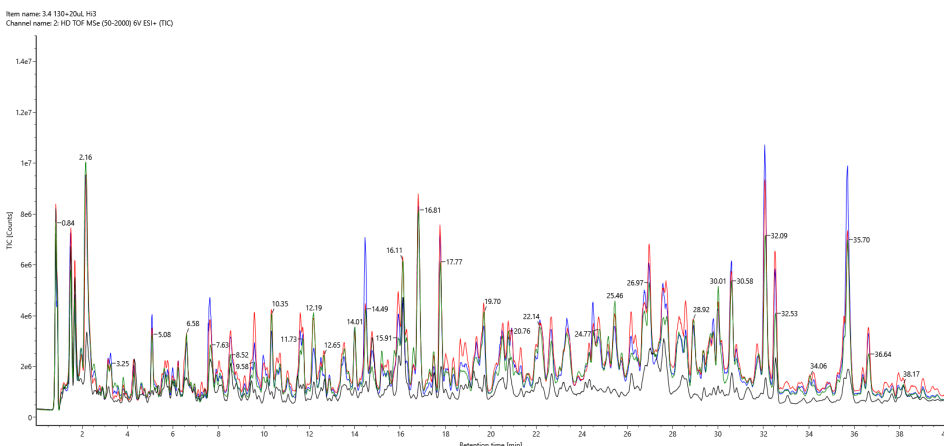


Figure 12. Cheese starting point sample TIC profiles (retention time vs intensity) UHPLC-IMS-HRMS overlay (Publication II).

MS sensitivity relative to the instrumentation used. In Publication II, the overall sensitivity loss, resulting from using travelling wave IMS-DIA (estimated at 20-30% versus native DIA) was mitigated by the higher sensitivity of the Vion IMS QToF (Water Corporation) compared to the SYNAPT G2-Si (Waters Corporation), used in Publication I. Considering the sensitivity performance characteristics of both systems, the nanoUHPLC demonstrated superior sensitivity results compared to analytical UHPLC.

The number of peptides identified. Both methods allowed the identification of over 400 peptides arising from casein hydrolysis. However, comparing BPI profiles from both publications, a significant delay in the elution profile from nanoUHPLC can be observed (Figure 11). The main reason for such delay was the requirement of the trap-elute step prior to separation on the analytical column. Moreover, sample complexity observed in the TIC profile of UHPLC (Figure 12) suggests that further improvement to the elution profile should have been made.

Although other studies have demonstrated higher numbers of identifiable peptides in similar studies (that are typically achieved with 2-4 hour analytical gradients), the final sample comparison is often narrowed down to the most significant peptides (abundance or fold change; typically less than 500, Taivosalo et al., 2018), thus making increased throughput (with no significant detrimental effect on the statistical model) a higher priority than deeper profiling capabilities.

The software played an important role in the analysis of data. PQI for Proteomics (Nonlinear Dynamics) used in the Publication I allowed for automated peak alignment, normalisation and processing, all of which had to be done manually or with help of Python script in the former study described in Publication II.

The addition of TWIMS in Publication II revealed the presence of peptides with the same sequence, CCS values of which were not identical. This allowed for even finer alignment and filtering as many peptides were either coeluting or eluting very closely. One of the possible explanations is that the peptides can originate from different protein isoforms, the structural conformation of which might be different. These peptides would have been fully missed with DDA mode due to dynamic exclusion, typically for 30-60 s, of species of the same m/z and therefore would have not been selected for fragmentation.

Finally, reflecting on the benefits and drawbacks of the methods tested in both publications, a combination of 300 μ m ID column and IMS-enabled DIA is highly recommended for peptide analysis in cheese to achieve high throughput while maintaining high sensitivity.

4.2 Case Study II: Peptide Mapping and Screening of Protein Hydrolysate Consumption During Alcoholic Fermentation (Supplementary Study I and Publication III)

In Case Study II, the use of UNIFI-based peptide mapping in combination with Progenesis QI (Nonlinear Dynamics) for profiling and screening of smaller molecular weight peptides was investigated.

First, the proteolytic activity of three commercial proteases or a combination of those was characterised to find a suitable protease for BSA hydrolysis (Supplementary Study I). As the goal of Publication III was to develop a methodology for assessment of peptides assimilation by yeast, the main criterium was to find a protease or combination of proteases that produces a peptide mixture from BSA with the maximum number of small MW (2-5 AA) peptides assimilable to yeast, with the least release of free amino acids. Based on this criterium, Corolase® 7089 (AB Enzymes, Darmstadt, Germany) was chosen for its ability to generate a significant number (123 peptides chosen for relative quantification) of small MW (di- to penta-) peptides while leaving the lowest concentration of free AA (Table 2).

Table 2. Comparison of performances of different industrial proteases and their combinations.

	Protease Combination						
	COROLASE® 7089	COROLASE® 8000	COROLASE® APC	COROLASE® 7089 & 8000	COROLASE® 7089 & APC	COROLASE® 8000 & APC	COROLASE® 7089 & 8000 & APC
	Free Amino Acid						
Concentration (mg/L)	33.42	436.92	571.45	392.67	658.47	774.55	693.41
Number identified peptides							
Dipeptides	23	38	45	45	43	50	49
Tripeptides	56	89	105	115	126	128	133
Tetrapeptides	47	76	73	88	91	81	95
Pentapeptides	45	62	60	61	57	52	57
Hexapeptides	41	40	48	38	39	34	41
Heptapeptides	22	26	29	35	25	19	17
Octapeptides	19	21	23	25	16	18	20
Nonapeptides	16	10	23	16	15	8	13
Decapeptides	15	16	10	13	9	9	8
Number of summed peptides							
Di-pentapeptides	171	265	283	309	317	311	334
Di-decapeptides	284	378	416	436	421	399	433

Although small peptide sequence identification using this method remains a challenge (due to uncertainty in absolute identification assignment of short peptides), the tentative peptide identification still provides useful information about peptides' chain length and amino acid composition. The proposed methodology, relying on the reduced sample complexity, benefits from a significantly shortened time required for sample analysis, namely 18.5 min (compared to 70 and 48.5 min for Publications I and II respectively). For large cohort studies, even sub-20-minute analytical runs can result in weeks-long sample sets. Therefore, further improvement of the method throughput should be investigated.

4.3 Case Study III: Untargeted Peptide Profiling During Wort Production (Supplementary Study II)

Based on the results of Publication III and namely, the correlation of identification assignment accuracy between peptide mapping and untargeted analysis, Supplementary Study II was carried out to evaluate the effects of different malt types and mashing regimes on the resulting peptide composition in the respective worts. Figure 13 (from Publication III) reiterates the applicability of the fully untargeted peptidomics approach, where protein sequence might not be available. Namely, assignment of peptide length and peptide amino acid composition (non-discriminating to leucine and isoleucine) by fully untargeted workflow matched 90.5% and 71.4% of the assignment by the peptide mapping approach (using BSA sequence and non-specific cleavage). Being fully based on the untargeted workflow, Supplementary Study II was most dependent on the statistical evaluation of the samples. Table 3 provides a short recap of the samples studied.

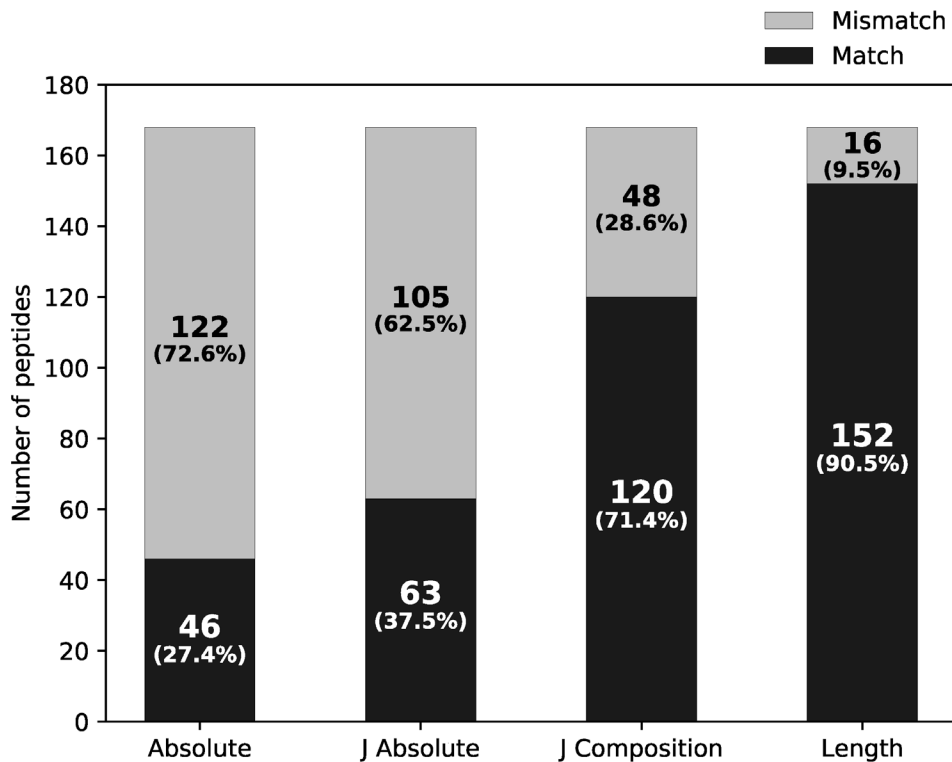


Figure 13. The dependence of the peptide matching specificity criteria on the peptide identification assignment discrepancy (Arju et al., 2022).

Table 3. Malt types, mashing regimes, and respective sample numbers (Figure 10).

		Malt		
		1	2	3
		Château Pilsen 2RS®	Chateau Pale Ale®	Chateau Distilling®
Mashing regime	1	1.1	2.1	3.1
	2	1.2	2.2	3.2
	3	1.3	2.3	3.3

Based on the tentative identification of 183 peptides (ChempSpider_Peptides database) with a fold change over 1.5, the PLS-DA showed a distinct separation of samples prepared with different malt types (Figure 14) and mashing regimes (Figure 15).

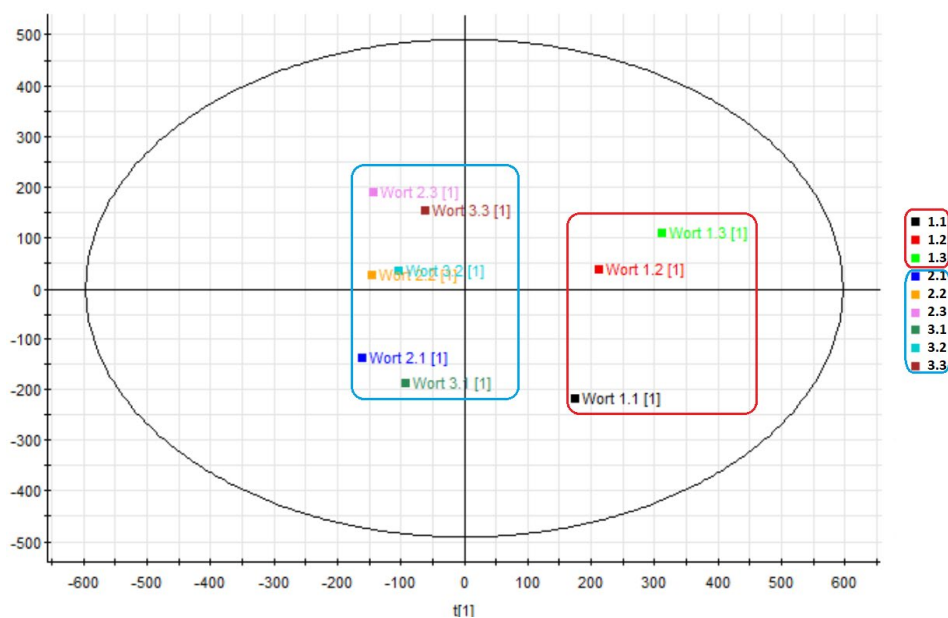


Figure 14. PLS-DA, grouping accordingly to malt type. Blue- Pale Ale and Distilling malts; Red- Pilsner malt (Table 3).

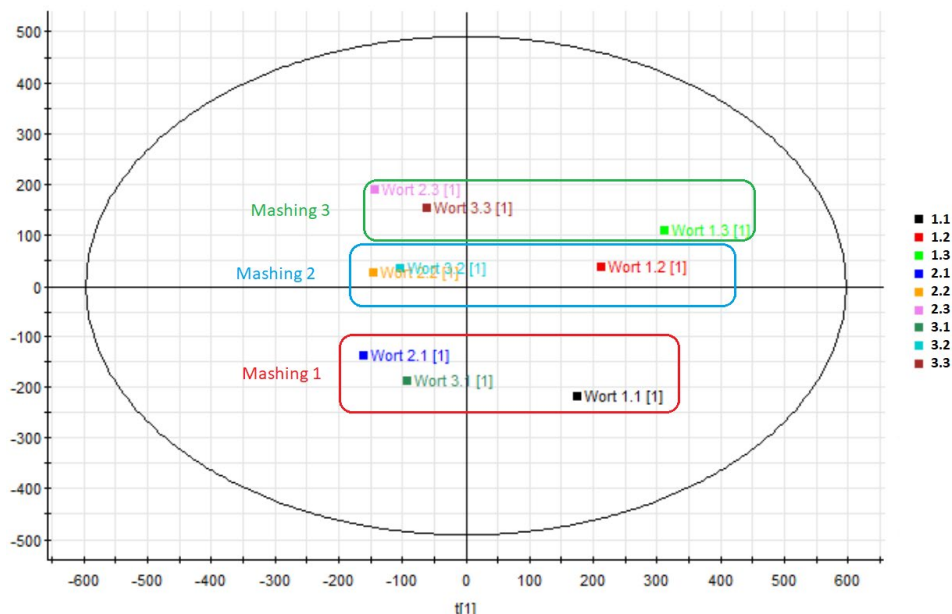


Figure 15. PLS-DA, grouping accordingly to mashing regime. Red- mashing 1; Blue- mashing 2; Green- mashing 3 (Table 3).

Further data review confirmed the trends between relative peptide abundances of Pilsner, Pale Ale, and Distilling malts (Figures 16 and 17) as well as between mashing regimes 1, 2 and 3 (Figures 18 and 19).

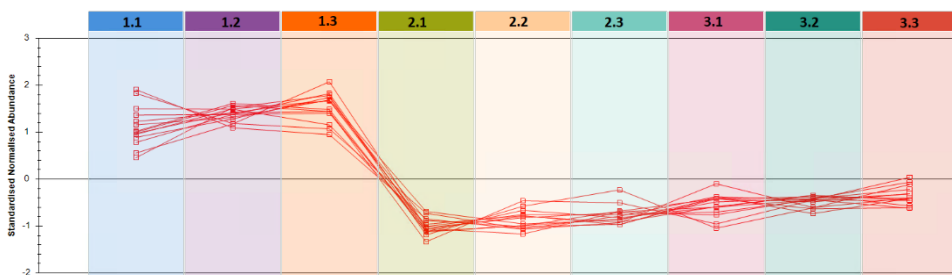


Figure 16. Relative peptide abundance changes due to different malt types (higher in Pilsner malt).

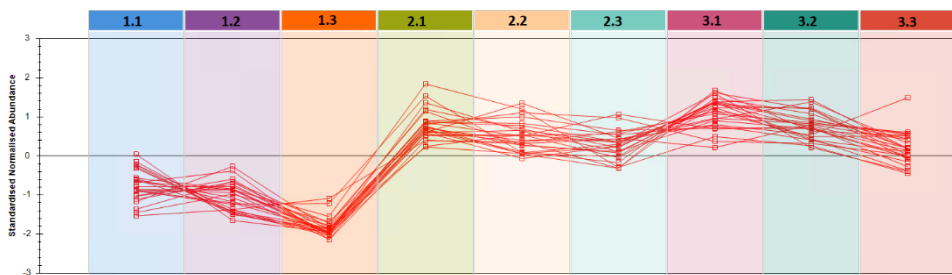


Figure 17. Relative peptide abundance changes due to different malt types (higher in Pale Ale and Distilling malts).

Figures 16 and 17 highlighted two trends that contribute to the differentiation observed in Figure 14. Namely, worts prepared from Pilsner malt (1.1-1.3) differentiated those from Pale Ale (2.1-2.3) and Distilling (3.1-3.3) malts. Each line in the graph represents the relative abundance of the most notable peptides (i.e., those with the highest fold change observed during individual experiments, Table 4).

Table 4. Example of top 5 peptides with highest fold change differentiating the malt types used in Case Study II. Trend 1 represents peptides in Figure 16; Trend 2 represents peptides in Figure 17.

Trend 1 (Total 14 peptides)			Trend 2 (Total 26 peptides)		
Peptide	Max Abundance	Fold Chahnge	Peptide	Max Abundance	Fold Chahnge
Tyr-Leu-Glu-Glu	2800	6	Trp-Thr-Tyr	795	8.7
Lys-His-Asn-Asn	4560	4.4	Arg-Ile-Ile	882	4.5
Lys-Val-His-Asn	865	3.2	Ile-Pro-Val-Gln	584	3.2
Gly-Arg-Gly-Ser	915	3.1	Val-Asp-Lys-Pro	700	3
Glu-Gln-Gln	2225	3	Glu-Pro-Leu-Leu	560	2.7

Figure 15 highlights trends contributing to the differentiation observed in the mashing regimes (Figures 18 and 19), most notable (highest fold change top 5) of which are presented in Table 5.

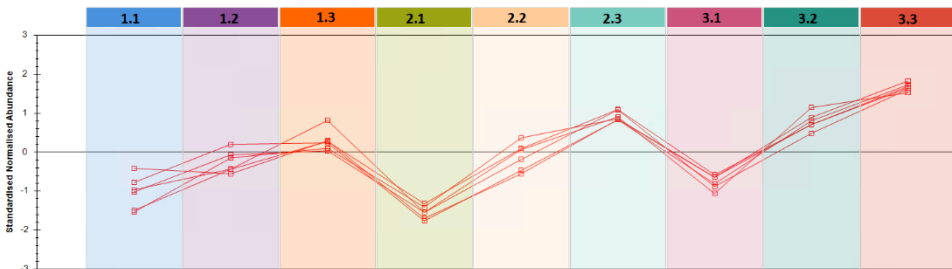


Figure 18. Peptide relative concentration trends due to differences in the mashing regimes (increasing peptide abundance with every additional mashing step).

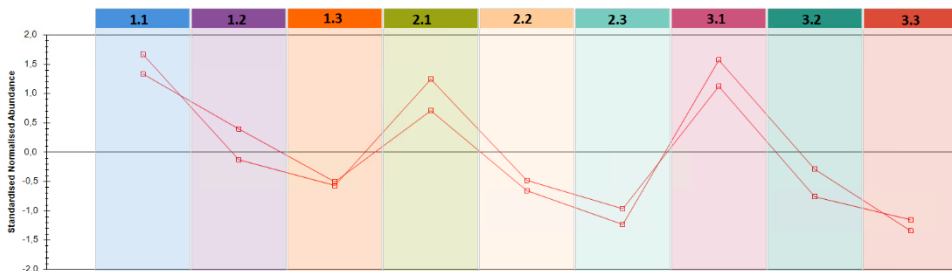


Figure 19. Peptide relative abundance trends due to differences in the mashing regimes (decreasing peptide abundance with every additional mashing step)

Table 5. Peptides with highest fold change differentiating the mashing regimes used in Case Study II. Trend 3 represents peptides in Figure 18; Trend 4 represents peptides in Figure 19.

Trend 3 (Total 6 peptides)			Trend 4 (Total 2 peptides)		
Peptide	Max Abundance	Fold Chahnge	Peptide	Max Abundance	Fold Chahnge
Ile-Glu-Ala	725	2.1	Tyr-Ala-Ala-Phe	802	3
Glu-Leu	690	2	Lys-Asn-Asp	499	1.5
Ile-Val-Asp	634	1.9			
Glu-Val-Val	551	1.8			
Val-Glu-Leu	638	1.7			

Comparison between Tables 4 and 5 revealed that there was not only a higher number of peptides that differentiates the malt types (trend 1 and 2 (14/26) vs trend 3 and 4 (6/2)) but also the maximum fold change was higher (trend 1 and 2 (6/8.7) vs trend 3 and 4 (2.1/3)), compared to what was observed in case of different mashing regimes. These results suggest that malt type might have higher impact on the composition of the peptides than the applied mashing regimes. Thus, peptides composition in the worts prepared from Pale Ale and Distilling malts are much more similar than those from Pilsner malt.

Even though this study was carried out as proof of a concept, it has still managed to provide significant evidence to justify the continuation of the research, due to the high technological value in it. The outcome of this study, demonstrating differences in peptide compositions and relative concentrations in response to malt types and mashing regimes has high importance for the food and fermentation industry and deserves further research with a more optimised experimental design to achieve better differentiation and higher result confidence levels. Additionally, similar studies can be used not only to investigate the effect of different processing parameters on peptide composition but also for authentication of different products or raw materials, based on the peptide profile.

5 Conclusions

The main objective of this thesis was to develop and implement higher throughput methodologies for peptide composition profiling and peptide consumption screening in various biological matrices applicable to food- and biotechnology. The developed methodologies were based on nanoUHPLC-DIA-HRMS and UHPLC-IMS-DIA-HRMS and were tested with three practical matrix-based case studies.

In the first case study, the development and implementation of peptide profiling during cheese ripening were carried out. The nanoUPLC-DIA-HRMS method was developed for peptide profiling of the Gouda-type cheese ripening and a significant reduction in analysis time (70-min) was achieved. However, the drawbacks of the improved throughput were found to be an underutilised elution profile of peptides and a longer equilibration time due to the use of nanoUHPLC. The reduction in the overall number of peptide identifications (c.a. 500 in the most peptide-rich samples) was found to be sufficient for effective sample differentiation and evaluation of the proteolytic activity during cheese ripening. For peptide profiling during Emmental cheese ripening, the UHPLC-IMS-DIA-HRMS method was developed. Not only was a further reduction in analysis time (less than 50-min) achieved, but also an improvement of the peptide elution profile along with simplified sample preparation protocol were realised due to the use of UHPLC. Additionally, the use of IMS has revealed the presence of previously unreported peptide structural isomers, detection of which was made possible due to the combined use of IMS and DIA. However, the use of UHPLC resulted in a lower overall peptide signal response compared to the peptide signal response in the previously described work with Gouda-type cheese. Similar to the methodology used for Gouda-type cheese, the drawback of the improved throughput in this work was a reduction in the overall number of peptide identifications (c.a. 450 in the most peptide-rich samples).

In the second case study, the peptide mapping method was developed for rapid elucidation of proteolytic enzyme cleavage specificity using UHPLC-IMS-DIA-HRMS. The developed method allowed for analysis of peptides composition in protein hydrolysate within less than a 20-min analytical run and consequently has contributed to a significant improvement of the sample throughput. The methodology was then applied for the untargeted screening of short peptide consumption during microbial fermentation. Another important outcome of this study is that the developed methodology can be used to profile the short peptide (2-4 AA) composition of previously unelucidated protein sources with a certain degree of confidence (peptide length and AA composition). Successful completion of this study allowed for a critical large-cohort-data evaluation and proposition of additional changes to further improve the methodology's throughput.

In the third case study, the same UHPLC-IMS-DIA-HRMS methodology was further shown applicable for untargeted evaluation of peptide composition in unelucidated complex natural matrices.

References

- Agyei, D., & Danquah, M. K. (2011). Industrial-scale manufacturing of pharmaceutical-grade bioactive peptides. *Biotechnology Advances*, 29(3), 272–277. <https://doi.org/10.1016/j.biotechadv.2011.01.001>
- Al Feteisi, H., Achour, B., Barber, J., & Rostami-Hodjegan, A. (2015). Choice of LC-MS methods for the absolute quantification of drug-metabolizing enzymes and transporters in human tissue: a comparative cost analysis. *The AAPS Journal*, 17(2), 438–446. <https://doi.org/10.1208/s12248-014-9712-6>
- Badgett, M. J., Boyes, B., & Orlando, R. (2018). Peptide retention prediction using hydrophilic interaction liquid chromatography coupled to mass spectrometry. *Journal of Chromatography. A*, 1537, 58–65. <https://doi.org/10.1016/j.chroma.2017.12.055>
- Bateman, N. W., Goulding, S. P., Shulman, N. J., Gadok, A. K., Szumlinski, K. K., MacCoss, M. J., & Wu, C. C. (2014). Maximizing peptide identification events in proteomic workflows using data-dependent acquisition (DDA). *Molecular & Cellular Proteomics: MCP*, 13(1), 329–338. <https://doi.org/10.1074/mcp.M112.026500>
- Becerra-Rodríguez, C., Taghouti, G., Portier, P., Dequin, S., Casal, M., Paiva, S., & Galeote, V. (2021). Yeast Plasma Membrane Fungal Oligopeptide Transporters Display Distinct Substrate Preferences despite Their High Sequence Identity. *Journal of Fungi (Basel, Switzerland)*, 7(11). <https://doi.org/10.3390/jof7110963>
- Bekker-Jensen, D. B., Martínez-Val, A., Steigerwald, S., Rütther, P., Fort, K. L., Arrey, T. N., Harder, A., Makarov, A., & Olsen, J. V. (2020). A Compact Quadrupole-Orbitrap Mass Spectrometer with FAIMS Interface Improves Proteome Coverage in Short LC Gradients. *Molecular & Cellular Proteomics: MCP*, 19(4), 716–729. <https://doi.org/10.1074/mcp.TIR119.001906>
- Beri, J., Rosenblatt, M. M., Strauss, E., Urh, M., & Bereman, M. S. (2015). Reagent for Evaluating Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Performance in Bottom-Up Proteomic Experiments. *Analytical Chemistry*, 87(23), 11635–11640. <https://doi.org/10.1021/acs.analchem.5b04121>
- Bikaki, M., Shah, R., Müller, A., & Kuhnert, N. (2021). Heat induced hydrolytic cleavage of the peptide bond in dietary peptides and proteins in food processing. *Food Chemistry*, 357, 129621. <https://doi.org/10.1016/j.foodchem.2021.129621>
- Boukil, A., Suwal, S., Chamberland, J., Pouliot, Y., & Doyen, A. (2018). Ultrafiltration performance and recovery of bioactive peptides after fractionation of tryptic hydrolysate generated from pressure-treated β -lactoglobulin. *Journal of Membrane Science*, 556, 42–53. <https://doi.org/10.1016/j.memsci.2018.03.079>
- Chen, M.-X., Zhang, Y., Fernie, A. R., Liu, Y.-G., & Zhu, F.-Y. (2021). SWATH-MS-Based Proteomics: Strategies and Applications in Plants. *Trends in Biotechnology*, 39(5), 433–437. <https://doi.org/10.1016/j.tibtech.2020.09.002>
- DeLano, M., Walter, T. H., Lauber, M. A., Gilar, M., Jung, M. C., Nguyen, J. M., Boissel, C., Patel, A. V., Bates-Harrison, A., & Wyndham, K. D. (2021). Using Hybrid Organic-Inorganic Surface Technology to Mitigate Analyte Interactions with Metal Surfaces in UHPLC. *Analytical Chemistry*, 93(14), 5773–5781. <https://doi.org/10.1021/acs.analchem.0c05203>

- Dicker, L., Lin, X., & Ivanov, A. R. (2010). Increased Power for the Analysis of Label-free LC-MS/MS Proteomics Data by Combining Spectral Counts and Peptide Peak Attributes*. *Molecular & Cellular Proteomics: MCP*, 9(12), 2704–2718. <https://doi.org/10.1074/mcp.M110.002774>
- Fic, E., Kedracka-Krok, S., Jankowska, U., Pirog, A., & Dziedzicka-Wasylewska, M. (2010). Comparison of protein precipitation methods for various rat brain structures prior to proteomic analysis. *Electrophoresis*, 31(21), 3573–3579. <https://doi.org/10.1002/elps.201000197>
- Finoult, I., Pinkse, M., Van Dongen, W., & Verhaert, P. (2011). Sample preparation techniques for the untargeted LC-MS-based discovery of peptides in complex biological matrices. *Journal of Biomedicine & Biotechnology*, 2011, 245291. <https://doi.org/10.1155/2011/245291>
- Frey, B. L., Lador, D. T., Sondalle, S. B., Krusemark, C. J., Jue, A. L., Coon, J. J., & Smith, L. M. (2013). Chemical derivatization of peptide carboxyl groups for highly efficient electron transfer dissociation. *Journal of the American Society for Mass Spectrometry*, 24(11), 1710–1721. <https://doi.org/10.1007/s13361-013-0701-2>
- Gertsman, I., & Barshop, B. A. (2018). Promises and pitfalls of untargeted metabolomics. *Journal of Inherited Metabolic Disease*, 41(3), 355–366. <https://doi.org/10.1007/s10545-017-0130-7>
- Graw, S., Tang, J., Zafar, M. K., Byrd, A. K., Bolden, C., Peterson, E. C., & Byrum, S. D. (2020). proteiNorm - A User-Friendly Tool for Normalization and Analysis of TMT and Label-Free Protein Quantification. *ACS Omega*, 5(40), 25625–25633. <https://doi.org/10.1021/acsomega.0c02564>
- Gritti, F., David, M., Brothy, P., & Lewis, M. R. (2022). Model of retention time and density of gradient peak capacity for improved LC-MS method optimization: Application to metabolomics. *Analytica Chimica Acta*, 1197, 339492. <https://doi.org/10.1016/j.aca.2022.339492>
- Guimaraes, G. J., Sutton, J. M., Gilar, M., Donegan, M., & Bartlett, M. G. (2022). Impact of Nonspecific Adsorption to Metal Surfaces in Ion Pair-RP LC-MS Impurity Analysis of Oligonucleotides. *Journal of Pharmaceutical and Biomedical Analysis*, 208, 114439. <https://doi.org/10.1016/j.jpba.2021.114439>
- Heusel, M., Bludau, I., Rosenberger, G., Hafen, R., Frank, M., Banaei-Esfahani, A., van Drogen, A., Collins, B. C., Gstaiger, M., & Aebersold, R. (2019). Complex-centric proteome profiling by SEC-SWATH-MS. *Molecular Systems Biology*, 15(1), e8438. <https://doi.org/10.15252/msb.20188438>
- Huang, L.-J., Chiang, C.-W., Chen, S.-L., Wei, S.-Y., & Chen, S.-H. (2019). Complete mapping of disulfide linkages for etanercept products by multi-enzyme digestion coupled with LC-MS/MS using multi-fragmentations including CID and ETD. *Journal of Food and Drug Analysis*, 27(2), 531–541. <https://doi.org/10.1016/j.jfda.2018.11.007>
- Isaac, G., Wilson, I. D., & Plumb, R. S. (2022). Application of hybrid surface technology for improving sensitivity and peak shape of phosphorylated lipids such as phosphatidic acid and phosphatidylserine. *Journal of Chromatography. A*, 1669, 462921. <https://doi.org/10.1016/j.chroma.2022.462921>

- Jiang, N., Gao, Y., Xu, J., Luo, F., Zhang, X., & Chen, R. (2022). A data-independent acquisition (DIA)-based quantification workflow for proteome analysis of 5000 cells. *Journal of Pharmaceutical and Biomedical Analysis*, 216, 114795. <https://doi.org/10.1016/j.jpba.2022.114795>
- Kevvai, K., Kütt, M.-L., Nisamedtinov, I., & Paalme, T. (2016). Simultaneous utilization of ammonia, free amino acids and peptides during fermentative growth of *Saccharomyces cerevisiae*. *Journal of the Institute of Brewing. Institute of Brewing*, 122(1), 110–115. <https://doi.org/10.1002/jib.298>
- Kim, S., Mischerikow, N., Bandeira, N., Navarro, J. D., Wich, L., Mohammed, S., Heck, A. J. R., & Pevzner, P. A. (2010). The generating function of CID, ETD, and CID/ETD pairs of tandem mass spectra: applications to database search. *Molecular & Cellular Proteomics: MCP*, 9(12), 2840–2852. <https://doi.org/10.1074/mcp.M110.003731>
- Kosmides, A. K., Kamisoglu, K., Calvano, S. E., Corbett, S. A., & Androulakis, I. P. (2013). Metabolomic fingerprinting: challenges and opportunities. *Critical Reviews in Biomedical Engineering*, 41(3), 205–221. <https://doi.org/10.1615/critrevbiomedeng.2013007736>
- Kultima, K., Nilsson, A., Scholz, B., Rossbach, U. L., Fälth, M., & Andrén, P. E. (2009). Development and Evaluation of Normalization Methods for Label-free Relative Quantification of Endogenous Peptides*. *Molecular & Cellular Proteomics: MCP*, 8(10), 2285–2295. <https://doi.org/10.1074/mcp.M800514-MCP200>
- Li, G., Delafield, D. G., & Li, L. (2020). Improved structural elucidation of peptide isomers and their receptors using advanced ion mobility-mass spectrometry. *Trends in Analytical Chemistry: TRAC*, 124, 115546. <https://doi.org/10.1016/j.trac.2019.05.048>
- Liigand, P., Liigand, J., Cuyckens, F., Vreeken, R. J., & Kruve, A. (2018). Ionisation efficiencies can be predicted in complicated biological matrices: A proof of concept. *Analytica Chimica Acta*, 1032, 68–74. <https://doi.org/10.1016/j.aca.2018.05.072>
- Liu, C., Topchiy, E., Lehmann, T., & Basile, F. (2015). Characterization of the dehydration products due to thermal decomposition of peptides by liquid chromatography-tandem mass spectrometry. *Journal of Mass Spectrometry: JMS*, 50(3), 625–632. <https://doi.org/10.1002/jms.3570>
- Liu, X., Valentine, S. J., Plasencia, M. D., Trimpin, S., Naylor, S., & Clemmer, D. E. (2007). Mapping the human plasma proteome by SCX-LC-IMS-MS. *Journal of the American Society for Mass Spectrometry*, 18(7), 1249–1264. <https://doi.org/10.1016/j.jasms.2007.04.012>
- Liu, X., Wang, Q., & Lauber, M. A. (2022). High sensitivity acidic N-glycan profiling with MS-enhancing derivatization and mixed mode chromatography. *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences*, 1191, 123120. <https://doi.org/10.1016/j.jchromb.2022.123120>
- Martini, S., Solieri, L., & Tagliacucchi, D. (2021). Peptidomics: new trends in food science. *Current Opinion in Food Science*, 39, 51–59. <https://doi.org/10.1016/j.cofs.2020.12.016>
- Meier, F., Brunner, A.-D., Frank, M., Ha, A., Bludau, I., Voytik, E., Kaspar-Schoenefeld, S., Lubeck, M., Raether, O., Bache, N., Aebersold, R., Collins, B. C., Röst, H. L., & Mann, M. (2020). diaPASEF: parallel accumulation-serial fragmentation combined with data-independent acquisition. *Nature Methods*, 17(12), 1229–1236. <https://doi.org/10.1038/s41592-020-00998-0>

Michalski, A., Damoc, E., Lange, O., Denisov, E., Nolting, D., Müller, M., Viner, R., Schwartz, J., Remes, P., Belford, M., Dunyach, J.-J., Cox, J., Horning, S., Mann, M., & Makarov, A. (2012). Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. *Molecular & Cellular Proteomics: MCP*, *11*(3), O111.013698. <https://doi.org/10.1074/mcp.O111.013698>

Morris, L. S., & Marchesi, J. R. (2016). Assessing the impact of long term frozen storage of faecal samples on protein concentration and protease activity. *Journal of Microbiological Methods*, *123*, 31–38. <https://doi.org/10.1016/j.mimet.2016.02.001>

Moughan, P. J., Rutherford, S. M., Montoya, C. A., & Dave, L. A. (2014). Food-derived bioactive peptides--a new paradigm. *Nutrition Research Reviews*, *27*(1), 16–20. <https://doi.org/10.1017/S0954422413000206>

Nakayasu, E. S., Gritsenko, M., Piehowski, P. D., Gao, Y., Orton, D. J., Schepmoes, A. A., Fillmore, T. L., Frohnert, B. I., Rewers, M., Krischer, J. P., Ansong, C., Suchy-Dicey, A. M., Evans-Molina, C., Qian, W.-J., Webb-Robertson, B.-J. M., & Metz, T. O. (2021). Tutorial: best practices and considerations for mass-spectrometry-based protein biomarker discovery and validation. *Nature Protocols*, *16*(8), 3737–3760. <https://doi.org/10.1038/s41596-021-00566-6>

Nanita, S. C., & Kaldon, L. G. (2016). Emerging flow injection mass spectrometry methods for high-throughput quantitative analysis. *Analytical and Bioanalytical Chemistry*, *408*(1), 23–33. <https://doi.org/10.1007/s00216-015-9193-1>

Olsen, J. V., & Mann, M. (2004). Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(37), 13417–13422. <https://doi.org/10.1073/pnas.0405549101>

Pinu, F. R., Goldansaz, S. A., & Jaine, J. (2019). Translational Metabolomics: Current Challenges and Future Opportunities. *Metabolites*, *9*(6). <https://doi.org/10.3390/metabo9060108>

Proust, L., Sourabié, A., Pedersen, M., Besançon, I., Haudebourg, E., Monnet, V., & Juillard, V. (2019). Insights Into the Complexity of Yeast Extract Peptides and Their Utilization by *Streptococcus thermophilus*. *Frontiers in Microbiology*, *10*, 906. <https://doi.org/10.3389/fmicb.2019.00906>

Qiao, X.-Q., Wang, R., Zhang, L.-H., Yang, G.-L., & Zhang, Y.-K. (2012). Recent Advancement of Chemical Derivatization and Its Applications to High Sensitive Analysis of Peptide in Mass Spectrometry. *Chinese Journal of Analytical Chemistry*, *40*(7), 1123–1129. [https://doi.org/10.1016/S1872-2040\(11\)60560-4](https://doi.org/10.1016/S1872-2040(11)60560-4)

Ralston-Hooper, K., Hopf, A., Oh, C., Zhang, X., Adamec, J., & Sepúlveda, M. S. (2008). Development of GCxGC/TOF-MS metabolomics for use in ecotoxicological studies with invertebrates. *Aquatic Toxicology*, *88*(1), 48–52. <https://doi.org/10.1016/j.aquatox.2008.03.002>

Ramachandran, S., & Thomas, T. (2021). FPTMS: Frequency-based approach to identify the peptide from the low-energy collision-induced dissociation tandem mass spectra. *Journal of Proteomics*, *235*, 104116. <https://doi.org/10.1016/j.jprot.2021.104116>

- Romero-Luna, H. E., Hernández-Mendoza, A., González-Córdova, A. F., & Peredo-Lovillo, A. (2022). Bioactive peptides produced by engineered probiotics and other food-grade bacteria: A review. *Food Chemistry: X*, 13, 100196. <https://doi.org/10.1016/j.fochx.2021.100196>
- Sarvin, B., Lagziel, S., Sarvin, N., Mukha, D., Kumar, P., Aizenshtein, E., & Shlomi, T. (2020). Fast and sensitive flow-injection mass spectrometry metabolomics by analyzing sample-specific ion distributions. *Nature Communications*, 11(1), 3186. <https://doi.org/10.1038/s41467-020-17026-6>
- Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D., & McLean, J. A. (2016). Untargeted Metabolomics Strategies-Challenges and Emerging Directions. *Journal of the American Society for Mass Spectrometry*, 27(12), 1897–1905. <https://doi.org/10.1007/s13361-016-1469-y>
- Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W.-M., Fiehn, O., Goodacre, R., Griffin, J. L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A. N., Lindon, J. C., Marriott, P., Nicholls, A. W., ... Viant, M. R. (2007). Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics: Official Journal of the Metabolomic Society*, 3(3), 211–221. <https://doi.org/10.1007/s11306-007-0082-2>
- Taivosalo, A., Kriščiunaite, T., Seiman, A., Part, N., Stulova, I., & Vilu, R. (2018). Comprehensive analysis of proteolysis during 8 months of ripening of high-cooked Old Saare cheese. *Journal of Dairy Science*, 101(2), 944–967. <https://doi.org/10.3168/jds.2017-12944>
- Tejada-Casado, C., Hernández-Mesa, M., Monteau, F., Lara, F. J., Olmo-Iruela, M. D., García-Campaña, A. M., Le Bizec, B., & Dervilly-Pinel, G. (2018). Collision cross section (CCS) as a complementary parameter to characterize human and veterinary drugs. *Analytica Chimica Acta*, 1043, 52–63. <https://doi.org/10.1016/j.aca.2018.09.065>
- Temussi, P. A. (2012). The good taste of peptides. *Journal of Peptide Science: An Official Publication of the European Peptide Society*, 18(2), 73–82. <https://doi.org/10.1002/psc.1428>
- Urbizo-Reyes, U., Liceaga, A. M., Reddivari, L., Kim, K.-H., & Anderson, J. M. (2022). Enzyme kinetics, molecular docking, and in silico characterization of canary seed (*Phalaris canariensis* L.) peptides with ACE and pancreatic lipase inhibitory activity. *Journal of Functional Foods*, 88, 104892. <https://doi.org/10.1016/j.jff.2021.104892>
- Valeja, S. G., Xiu, L., Gregorich, Z. R., Guner, H., Jin, S., & Ge, Y. (2015). Three dimensional liquid chromatography coupling ion exchange chromatography/hydrophobic interaction chromatography/reverse phase chromatography for effective protein separation in top-down proteomics. *Analytical Chemistry*, 87(10), 5363–5371. <https://doi.org/10.1021/acs.analchem.5b00657>
- Valletta, M., Ragucci, S., Landi, N., Di Maro, A., Pedone, P. V., Russo, R., & Chambery, A. (2021). Mass spectrometry-based protein and peptide profiling for food frauds, traceability and authenticity assessment. *Food Chemistry*, 365, 130456. <https://doi.org/10.1016/j.foodchem.2021.130456>

- Van Puyvelde, B., Daled, S., Willems, S., Gabriels, R., Gonzalez de Peredo, A., Chaoui, K., Mouton-Barbosa, E., Bouyssié, D., Boonen, K., Hughes, C. J., Gethings, L. A., Perez-Riverol, Y., Bloomfield, N., Tate, S., Schiltz, O., Martens, L., Deforce, D., & Dhaenens, M. (2022). A comprehensive LFQ benchmark dataset on modern day acquisition strategies in proteomics. *Scientific Data*, *9*(1), 126. <https://doi.org/10.1038/s41597-022-01216-6>
- Wang, B., Ezejias, T., Feng, H., & Blaschek, H. (2008). Sugaring-out: A novel phase separation and extraction system. *Chemical Engineering Science*, *63*(9), 2595–2600. <https://doi.org/10.1016/j.ces.2008.02.004>
- Wang, D., Fang, S., & Wohlueter, R. M. (2009). N-terminal derivatization of peptides with isothiocyanate analogues promoting Edman-type cleavage and enhancing sensitivity in electrospray ionization tandem mass spectrometry analysis. *Analytical Chemistry*, *81*(5), 1893–1900. <https://doi.org/10.1021/ac8021136>
- Windarsih, A., Suratno, Warmiko, H. D., Indrianingsih, A. W., Rohman, A., & Ulumuddin, Y. I. (2022). Untargeted metabolomics and proteomics approach using liquid chromatography-Orbitrap high resolution mass spectrometry to detect pork adulteration in *Pangasius hypophthalmus* meat. *Food Chemistry*, *386*, 132856. <https://doi.org/10.1016/j.foodchem.2022.132856>
- Zhao, Y., & Jensen, O. N. (2009). Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics*, *9*(20), 4632–4641. <https://doi.org/10.1002/pmic.200900398>
- Zhou, Z., Tu, J., & Zhu, Z.-J. (2018). Advancing the large-scale CCS database for metabolomics and lipidomics at the machine-learning era. *Current Opinion in Chemical Biology*, *42*, 34–41. <https://doi.org/10.1016/j.cbpa.2017.10.033>
- Zhu, H., Wu, X., Huo, J., Hou, J., Long, H., Zhang, Z., Wang, B., Tian, M., Chen, K., Guo, D. 'an, Lei, M., & Wu, W. (2021). A five-dimensional data collection strategy for multicomponent discovery and characterization in Traditional Chinese Medicine: *Gastrodia Rhizoma* as a case study. *Journal of Chromatography. A*, *1653*, 462405. <https://doi.org/10.1016/j.chroma.2021.462405>

Acknowledgements

I would like to acknowledge the support of my supervisors, Prof. Raivo Vilu and Dr. Ildar Nisamedtinov.

I would like to thank my colleagues, collaborators, industrial partners, and co-authors for their input in making the thesis happen.

This work has been partially supported by “TUT Institutional Development Program for 2016–2022” Graduate School in Biomedicine and Biotechnology receiving funding from the European Regional Development Fund under program ASTRA 2014–2020.4.01.16-0032 in Estonia and the Estonian Research Council via project RESTA13.

Abstract

Development and implementation of high throughput peptidomics for microbial studies

Peptides are a class of highly diversified biopolymeric molecules that find many practical uses. However, their chemical diversity result in many challenges associated with their analysis , which is especially evident in the case of complex natural matrices containing peptides from undefined proteins and proteases with unelucidated specificities.

This manuscript aims to develop liquid chromatography (LC) mass spectrometry (MS) methodologies for peptide analysis in different food matrices. The primary focus was on the untargeted workflows using ultrahigh pressure liquid chromatography (UHPLC) coupled with the data-independent acquisition (DIA) high-resolution mass spectrometry (HRMS) to improve throughput and flexibility of peptidome profiling and untargeted peptide consumption screening capabilities in various matrices such as cheese, synthetic grape must and beer wort.

Theoretical and practical knowledge created in this work, along with published data and supplementary materials present progression from a rather complex workflow based on nano-UHPLC and DIA-HRMS to a well-streamlined, high-throughput workflow utilising UPLC coupled with ion-mobility separation enabled (IMS) DIA-HRMS with substantially simplified sample preparation applicable for peptidome analysis in food- and biotechnological studies.

Lühikokkuvõte

Suure läbilaskevõimega peptidoomika meetodite arendamine ja juurutamine mikrobioloogilisteks uuringuteks

Peptiidid on paljude erinevate funktsioonidega aminohapetest koosnevad biomolekulid, millede määramine võib omada tihti praktilist tähtsust. Tulenevalt oma keemilisest mitmekesisusest on peptiidide analüüs seotud mitmete väljakutsetega, eelkõige keeruliste bioloogiliste maatriksite puhul, kus peptiidid pärinevad määratlemata aminohappelise järjestusega valkudest ja/või on tekkinud teadmata spetsiifilisusega proteaaside toimetel.

Käesoleva doktoritöö peamiseks eesmärgiks oli töötada välja kombineeritud vedelikkromatograafia ja massispektromeetria meetodikad peptiidide analüüsiks erinevates komplekssetes maatriksites. Töö põhifookus oli suunatud mittesihitud analüüsi meetodite väljatöötamisele, parandades peptidoomiprofiilide määramise jõudlust erinevate maatriksite näitel, nagu juust, viinamarja- ja õllevirre.

Töö teoreetilised ja praktilised väljundid koos avaldatud andmete ja täiendavate materjalidega tutvustavad edasiminekut nano-UHPLC-DIA-HRMS-il (*ingl. k.* nano ultrahigh pressure liquid chromatography data-independent acquisition high-resolution mass spectrometry) põhinevalt suhteliselt komplitseeritud ja töömahukalt meetodikalt suure läbilaskevõimega meetodikale, mis baseerub UPLC-IMS-DIA-HRMS-il (*ingl. k.* ultrahigh pressure liquid chromatography ion mobility separation data-independent acquisition high-resolution mass spectrometry) ja võimaldab seejuures kasutada ka oluliselt lihtsustatud proovide ettevalmistust peptidoomi analüüsiks toidu- ja biotehnoloogilistes protsessides.

Appendix 1

Publication I

Arju, G., Taivosalo, A., Pismennoi, D., Lints, T., Vilu, R., Daneberga, Z., Vorslova, S., Renkonen, R., Joenvaara, S. (2020) Application of the UHPLC-DIA-HRMS Method for Determination of Cheese Peptides. *Foods* 9, 979. doi:10.3390/foods9080979

Article

Application of the UHPLC-DIA-HRMS Method for Determination of Cheese Peptides

Georg Arju ^{1,2,*}, Anastassia Taivosalo ², Dmitri Pismennoi ^{1,2}, Taivo Lints ^{1,2}, Raivo Vilu ², Zanda Daneberga ³ , Svetlana Vorslova ³, Risto Renkonen ^{4,5} and Sakari Joenvaara ^{4,5}

¹ Department of Chemistry and Biotechnology, School of Science, Tallinn University of Technology, Ehitajate tee 5, 12616 Tallinn, Estonia; dmitri@tftak.eu (D.P.); taivo.lints@tftak.eu (T.L.)

² Center of Food and Fermentation Technologies, Akadeemia tee 15A, 12618 Tallinn, Estonia; anastassia@tftak.eu (A.T.); raivo@tftak.eu (R.V.)

³ Institute of Oncology, Riga Stradins University, 13 Pilsonu Str., LV-1002 Riga, Latvia; zanda.daneberga@rsu.lv (Z.D.); svetlana.vorslova@rsu.lv (S.V.)

⁴ Transplantation Laboratory, Haartman Institute, University of Helsinki, FI-00014 Helsinki, Finland; risto.renkonen@helsinki.fi (R.R.); sakari.joenvaara@helsinki.fi (S.J.)

⁵ HUSLAB, Helsinki University Hospital, FI-00029 Helsinki, Finland

* Correspondence: georg@tftak.eu; Tel.: +372-53-401-565

Received: 30 May 2020; Accepted: 22 July 2020; Published: 23 July 2020



Abstract: Until now, cheese peptidomics approaches have been criticised for their lower throughput. Namely, analytical gradients that are most commonly used for mass spectrometric detection are usually over 60 or even 120 min. We developed a cheese peptide mapping method using nano ultra-high-performance chromatography data-independent acquisition high-resolution mass spectrometry (nanoUHPLC-DIA-HRMS) with a chromatographic gradient of 40 min. The 40 min gradient did not show any sign of compromise in milk protein coverage compared to 60 and 120 min methods, providing the next step towards achieving higher-throughput analysis. Top 150 most abundant peptides passing selection criteria across all samples were cross-referenced with work from other publications and a good correlation between the results was found. To achieve even faster sample turnaround enhanced DIA methods should be considered for future peptidomics applications.

Keywords: dairy product analysis; cheese peptidomics; cheesemaking; data-independent acquisition

1. Introduction

During cheese ripening, caseins undergo a progressive breakdown by enzymatic action, releasing peptides and amino acids, which contributes to the development of cheese flavour and texture [1]. The term “peptidomics” has been extensively used in dairy science for comprehensive analysis of peptides released during proteolysis in different cheese varieties [2–4] as well as characterisation of bioactive peptides with potential nutritional and health-promoting effects [5,6]. Several researchers have been focusing on the identification of phosphorylated peptides [7] and the determination of specific bitter peptides and their contribution to cheese flavour [8,9]. Many studies have been carried out to evaluate the effect of different adjunct cultures on the formation of peptides in cheese and thus to adjust the taste and aroma characteristics of a final product [10,11].

The key analytical tool employed in cheese peptidome research (i.e., increasing the knowledge of proteolytic events occurring during ripening) as well as exploring the possibilities of controlling the cheese maturation process, is currently mass spectrometry (MS) coupled with liquid chromatography (LC) [12–14]. The most widely used hyphenation is nano ultra (high) performance liquid chromatography (nanoUHPLC) coupled with high-resolution mass spectrometry (HRMS) based on data-dependent acquisition (DDA) mode [15–17].

The samples complexity combined with the slow acquisition rate of DDA modes as well as the ever-growing demand for higher protein coverage typically results in analytical gradients that exceed 60 or even 120 min, making such methods less appealing for higher-throughput studies [4,14]. Data-independent acquisition (DIA) is an alternative acquisition mode to DDA. DIA, unlike DDA, does not rely on precursor isolation. DIA is based on a principle of a rapid alternation between low and high collision energies to acquire MS¹ and MS² spectra. DIA relies on a chromatographic alignment of MS¹ and MS² for fragment-precursor assignment. Operating at higher acquisition rates and being compatible with faster gradients, DIA has been employed in several food research applications such as food safety, authenticity testing and peptide profiling of various food matrices [18,19]. Using DIA it is possible to simultaneously acquire both qualitative and quantitative data.

The aim of this study was to develop an LC-DIA-MS-based methodology for cheese peptide profiling with a sub-60 min analytical gradient without compromises in chromatographic performance and protein coverage.

2. Materials and Methods

2.1. Materials

Hi3 EColi STD (p/n: 186006012) and [Glu1]-Fibrinopeptide (p/n: 700004729) were purchased from Waters Corporation (Milford, MA, USA). Peptide quantitation was performed using Pierce™ Quantitative Colorimetric Peptide Assay (C/N: 23275, Thermo Fisher Scientific, Waltham, MA, USA). Nanosep® Centrifugal Devices with Omega™ membrane 3 K were obtained from Pall (p/n: OD003C34, Port Washington, NY USA). Ultrapure water (18.2 MΩ.cm) was prepared with MilliQ® Direct-Q® UV (Merck KGaA, Darmstadt, Germany). Acetonitrile (MeCN; LiChrosolv, hypergrade for LC-MS,) and formic acid (FA; LC-MS grade) were acquired from Sigma-Aldrich (Darmstadt, Germany).

2.2. Cheese Manufacture and Sampling

Three Gouda-type cheeses (Cheese 1, Cheese 2 and Cheese 3) were produced using three different DL-starter cultures (DL1 and DL2 by Chr. Hansen Ltd., Hørsholm, Denmark and DL3 by DuPont™ Danisco®, Copenhagen, Denmark) at a dairy plant from 600 L of pasteurised (at 74 °C for 15 s) milk. Animal rennet (25 mL 100 1/L; 230 IMCU 1/g; 20/80) of chymosin and bovine pepsin, (Chr. Hansen Ltd., Hørsholm, Denmark) was added to milk. After coagulation, the curd was cut, whey removed, and cheese grains stirred and heated at 32 °C for 30 min. Cheeses were prepressed under whey, moulded, pressed for 1.5 h, brine salted (pH 5.1) for 36 h, waxed and ripened at 10–15 °C for 90 days. Samples were taken from each cheese at 0 (after salting), 14, 30 and 90 days of ripening, grated and stored at –20 °C for further analysis.

2.3. Sample Preparation

To prepare water-soluble extracts (WSE) of cheeses, 2.5 g of grated cheese was homogenised in 22.5 mL of MilliQ® water (12,500–13,000 rpm) using Polytron PT 2100 dispersing aggregate with a diameter of 20 mm (Kinematica AG, Switzerland). Samples were heated for 10 min at 75 °C and centrifuged for 20 min at 4 °C at 13,302 g. Supernatants were stored in 1.5 mL Eppendorf® Protein LoBind microcentrifuge tubes (Eppendorf AG, Germany) at –20 °C until further purification. Thawed sample supernatants (250 µL) and MilliQ® water (250 µL) were transferred into the Eppendorf® tube and vortexed for 30 s. Obtained mixture (400 µL) was added to Nanosep® with 3 K Omega™ spin filter. Samples were centrifuged at 11,200 g for 15 min. For peptide quantification 30 µL of filtrate was mixed with 970 µL of working reagent. After 30 min incubation at room temperature, absorbance was measured at 480 nm (Helios Gamma, Thermo Electron Corporation, Waltham, MA, USA) and concentrations calculated using a blank-adjusted calibration curve. For HRMS analysis, 30 µL of filtrate was mixed with 970 µL of MilliQ® water (1% MeCN and 0.1% FA). Using results from the peptide concentration measurement samples were further diluted to result in 100 ng/µL equalising column

load across all the samples. The sample (195 μL) was transferred into a vial and spiked with 5 μL of 1 pmol/ μL of Hi3 EColi STD.

2.4. Liquid Chromatography Mass Spectrometry

Samples were analysed using Waters nanoAcquity UPLC[®] system (Waters Corporation, Milford, MA, USA) coupled with a Waters MALDI SYNAPT G2-Si Mass Spectrometer equipped with NanoLockSpray Exact Mass Ionisation Source and controlled by Waters MassLynx 4.1 (V4.1 SCN916, Waters Corporation, Milford, MA). Mobile phases were as follows: (A) MilliQ[®] + 0.1% formic acid and (B) MeCN + 0.1% formic acid. Injection volume was 2 μL . Samples were loaded onto Acquity UPLC[®] Symmetry C18 Nanoacquity 10 k 2 g V/M Trap column (100A, 5 μm , 180 μm \times 20 mm, Waters Corporation, Milford, MA, USA). Loading was carried out for 1 min at 5 $\mu\text{L}/\text{min}$ using 1% B. Loaded sample was further analysed using Acquity UPLC[®] M-Class HSS T3 Column (1.8 μm , 75 μm \times 150 mm, Waters Corporation, Milford, MA, USA) kept at 40 °C. The gradient was as follows: 0–1 min hold at 1% B, 1–10 min linear gradient 1–15% B, 10–40 min linear gradient 15–35% B, 40–45 min linear gradient 35–95% B, 45–53 min hold at 95% B, 53–55 min linear gradient 95–1% B, 55–70 min hold at 1% B. Analytical flow rate was 0.3 $\mu\text{L}/\text{min}$.

The instrument was operated in positive polarity and resolution mode (35000 FWHM at 785.8426 m/z). Data were acquired in MS^E mode with a scan time of 0.5 s between 1 and 50 min. Recorded mass range was from 50 to 2000 m/z for both low and high energy spectra. The collision energy was ramped from 15 to 45 V in the trap cell of the instrument. Cone voltage was set to 40 V and capillary voltage was set to 2.4 kV. Source offset was 60, source temperature was 80 °C. Cone gas was 50 L/h, nano flow gas was 0.3 bar and purge gas was 100 L/h. [Glu1]-Fibrinopeptide was used as LockMass for mass axis correction and was acquired every 30 s.

2.5. Raw Data Processing

The raw data files were imported to the Progenesis QI for proteomics software (Nonlinear Dynamics, Newcastle, UK). During the import masses were lock mass corrected with 785.8426 m/z , corresponding to doubly charged [Glu1]-Fibrinopeptide B. Default parameters for peak picking and alignment algorithm were used.

The peptides were searched against beta-casein (β -CN; P02666), alpha_{s1}-casein (α _{s1}-CN; P02662) and alpha_{s2}-casein (α _{s2}-CN; P02663) sequences from bovine species obtained with the UniProt database [20].

The protein identifications were done against sequences added with a spike in Hi3 standard ClpB protein sequence, CLPB_ECOLI (P63285). Nonspecific cleavage was chosen and zero missed cleavages were allowed. Fragment and peptide error tolerances were set to auto and false discovery rate to <1%. One or more fragment ions per peptide were required for ion matching. The following variable post-translational modifications were used in the analysis: oxidation (M), acetyl-(protein N-terminal), deamidation (NQ) and phosphorylation (STY). The analysis of each post-translational modification was done separately, and the results were combined. Absolute mass error for a peptide was set to 5 ppm and we included peptides with one to three charges in the analysis. In the sample grouping, the within-subject design was used, fold changes and repeated measures of ANOVA were used for statistics. Filtered data were exported and then subjected to the normalisation of peptide abundances based on the coefficients of each sample dilution.

2.6. Data Analysis

An additional batch of samples was analysed using the methodology described by Taivosalo et al. [14] to highlight differences in results between two approaches. DDA experiment raw data files were imported to MaxQuant proteomics software (<https://www.maxquant.org/>) for data analysis as described in the publication and subsequently exported for intramethod correlation analysis. Filtered data were exported and then subjected to the normalisation of peptides abundances based on

the coefficients of each sample dilution. Normalised abundances were used to construct a data matrix to identify differences between sample peptide compositions.

The comparison of the DIA and DDA methods was done with the help of in-house data analysis and visualisation scripts written in the Python™ programming language (Python Software Foundation). For both methods for each measured sample, the top 150 peptides with the highest intensities were found. The locations of those peptides were then found on the protein sequences the peptides originated from, and peptide coverage profiles were created for each casein in every sample for both methods, showing the peptides coloured by the intensity and laid out on their corresponding protein sequences.

3. Results

Overall, 558 peptides were identified among the analysed samples across 90 days of ripening using our method. The variation in peptide profiles and abundances was evaluated.

Ten per cent of samples (Day 0 of each cheese) were injected as triplicates. Relative standard deviation for all replicates equated to 10.88%.

It was found that at Day 90, Cheese 2 had the lowest number of peptides (Figure 1).

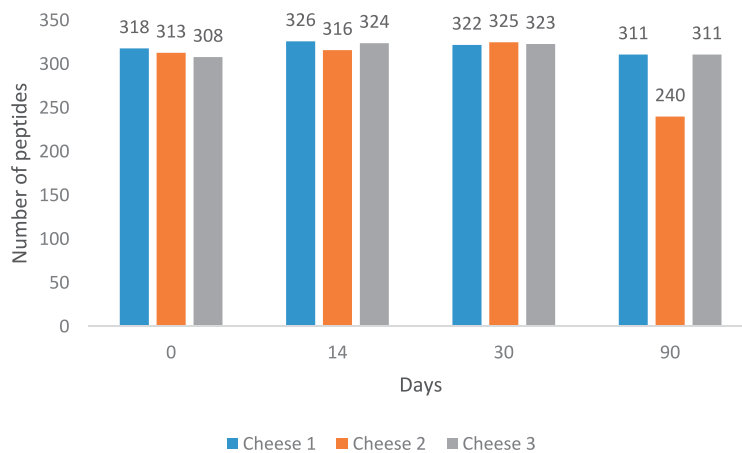


Figure 1. Number of identified peptides with unique amino acid sequences across 90 days of ripening.

Cheese 1 and Cheese 3 had 6.7- and 6-times higher summed peptide intensities compared to Cheese 2 (Figure S1).

At the same time, the average length of peptides in Cheese 2 was found to be longer than in other cheeses (Figure S2). All cheeses were subjected to comparative analysis to identify unique peptides at each measured point during cheese ripening. Results of the comparison are illustrated in Figure 2 that displays four Venn diagrams [21] for different days of ripening.

It was found that identified peptides during the first month of ripening were highly similar and accounted for approximately 93% identified peptides in all samples. During the ripening process, similarities in peptide composition between cheeses started to decrease. At the 90th day of ripening, Cheese 1 and Cheese 3 were more similar in peptide composition compared to Cheese 2, including over 60 identified peptides that were not present in Cheese 2.

Figure 3 indicates the difference in peptide accumulation and degradation pattern between days 0 and 90 for all three cheeses. Cheese 2, unlike the other two, displays the prevalence of peptide degradation compared to accumulation. This pattern is also consistent with the summed peptide intensities of each sample.

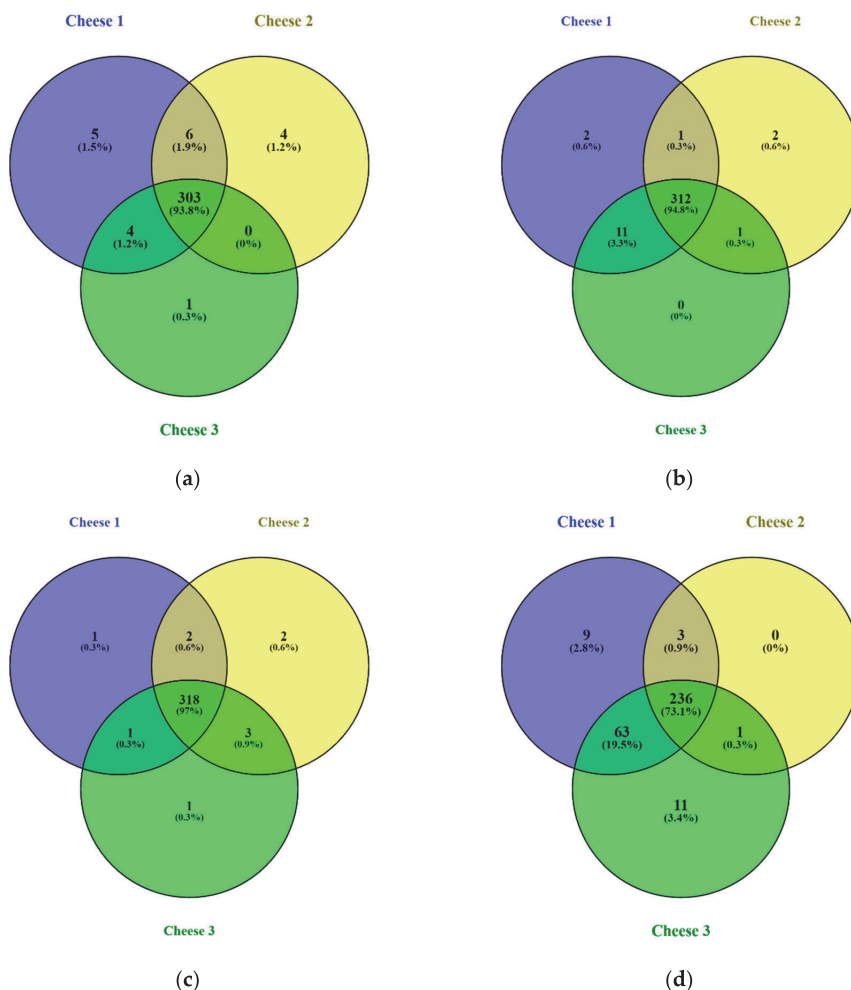


Figure 2. Venn diagrams of peptide distribution for Cheese 1, 2 and 3: (a) 0 days of ripening, (b) 14 days of ripening, (c) 30 days of ripening, (d) 90 days of ripening. Percentages in brackets denote proportion of all identified peptides across all cheeses.

For comparison between DDA and DIA-based approaches, the top 150 most abundant peptides per method were selected by their normalised intensities at the 90th day of ripening. It was found that 70 unique peptide sequences with a median length of nine amino acids were similar between DIA and DDA approaches, based on the identified peptides from Cheese 1. DIA-based approach results showed 80 unique peptide sequences with a median length of a peptide of seven amino acids. On the other hand, the DDA-based approach results showed 80 unique peptide sequences with a median length of 10 amino acids (Figure 4).

As for peptides identified in Cheese 2 sample, 58 unique peptide sequences (median length: 9 AA) were common between two methodologies, 92 unique peptide sequences (median length: 7 AA) belonged to DIA-based approach results and 92 unique peptide sequences (median length: 10 AA) were found only in the DDA-based approach results (Figure 5).

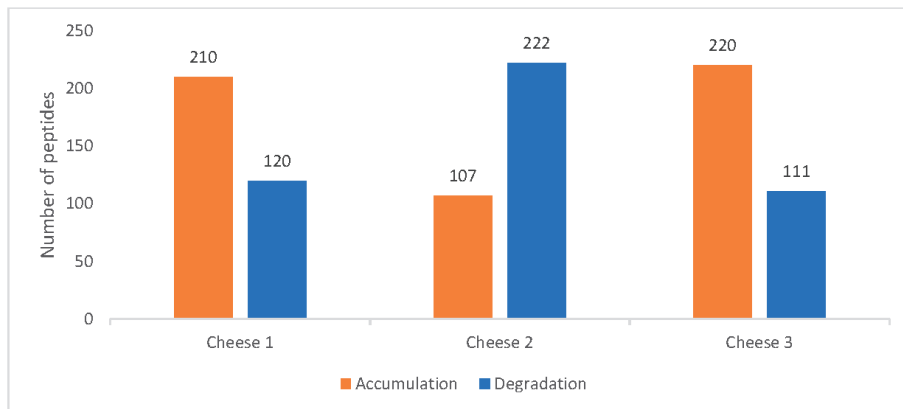


Figure 3. Cheese peptide profile trends between Days 0 and 90.

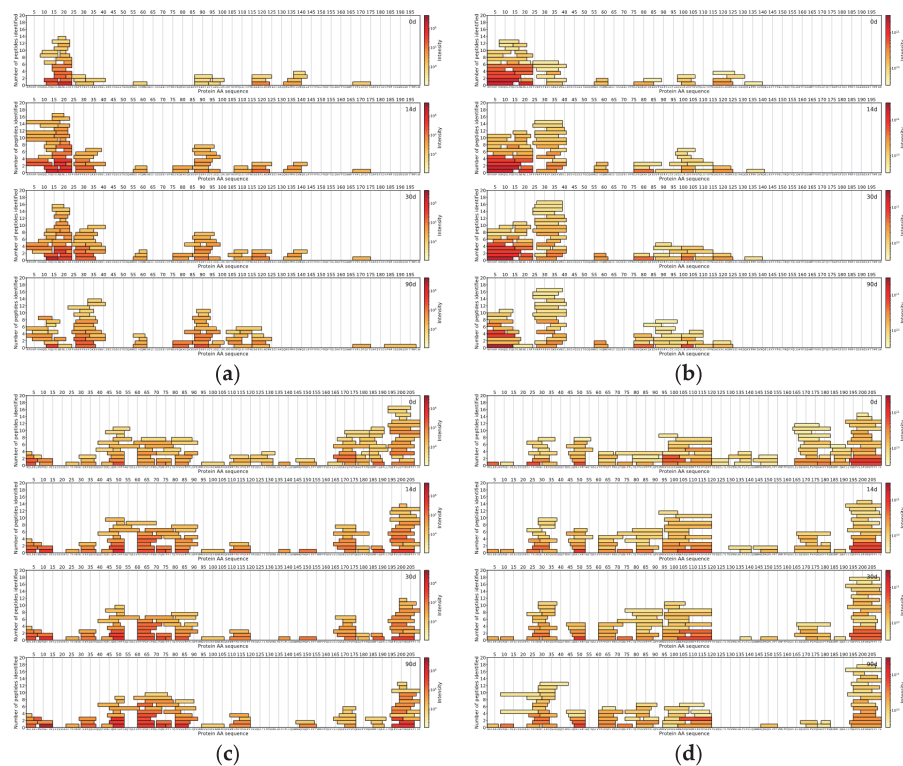


Figure 4. Cheese 1 α_{s1} -CN: (a) data-independent acquisition (DIA)-MS and (b) data-dependent acquisition (DDA)-MS. Cheese 1 β -CN: (c) DIA-MS and (d) DDA-MS. The X-axis represents the casein amino acid sequence, the Y-axis represents a number of peptides and colour represents the logarithmically scaled intensity of peptides.

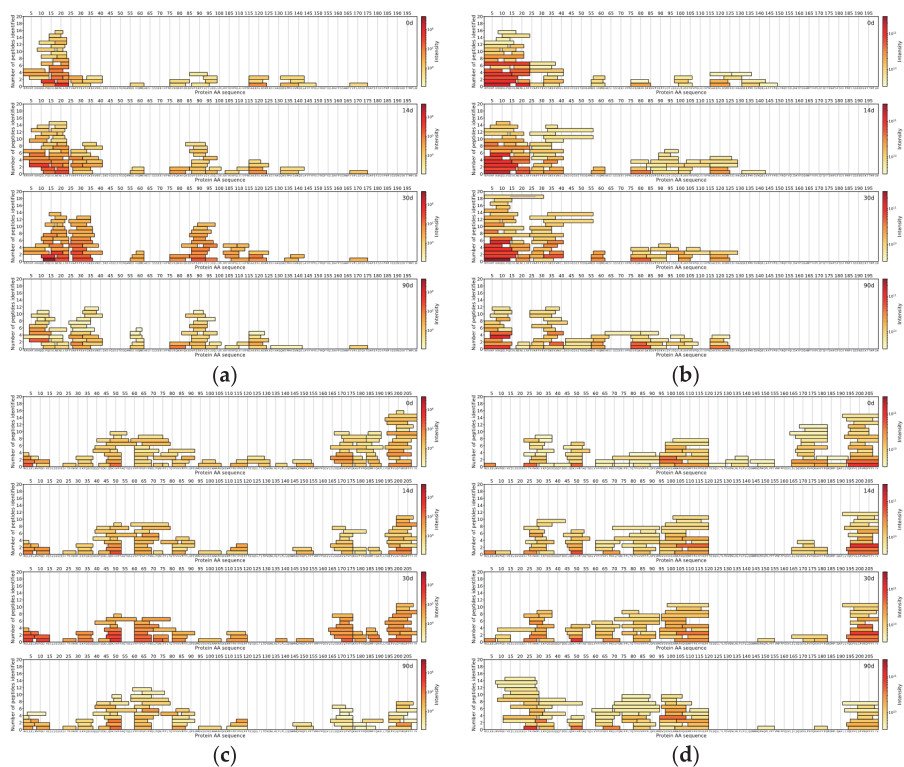


Figure 5. Cheese 2 α_{s1} -CN: (a) DIA-MS and (b) DDA-MS. Cheese 2 β -CN: (c) DIA-MS and (d) DDA-MS. The X-axis represents the casein amino acid sequence, the Y-axis represents a number of peptides and colour represents the logarithmically scaled intensity of peptides.

Peptides identified in Cheese 3 showed a similar trend as 74 unique peptide sequences (median length: 8 AA) were in common between two methodologies, 76 unique peptide sequences (median length: 7 AA) were found in DIA-based approach results and 76 unique peptide sequences (median length: 11 AA) belonged to the DDA-based approach (Figure 6).

Across all samples analysed with either DIA- or DDA-based approaches, peptides from α_{s1} -CN and β -CN comprised the majority of all detected peptides in the top 150 most abundant peptides.

In this study, we have also found several peptides, that have been previously reported to show bioactivity [6,22]: VPITPT (α_{s2} -CN f117-122), MPFPKYPVEPF (β -CN f109-119), EPVLGPVRGPF (β -CN f195-206), DKIHFP (β -CN f47-52), YPFPGIPN (β -CN f60-68), TPVVVPPFLQPE (β -CN f80-91), VPGEIVE (β -CN f8-14), VPSERYL (α_{s1} -CN f86-92), VLGPPVRGPF (β -CN f197-206) and YPFPGPI (β -CN f60-66).

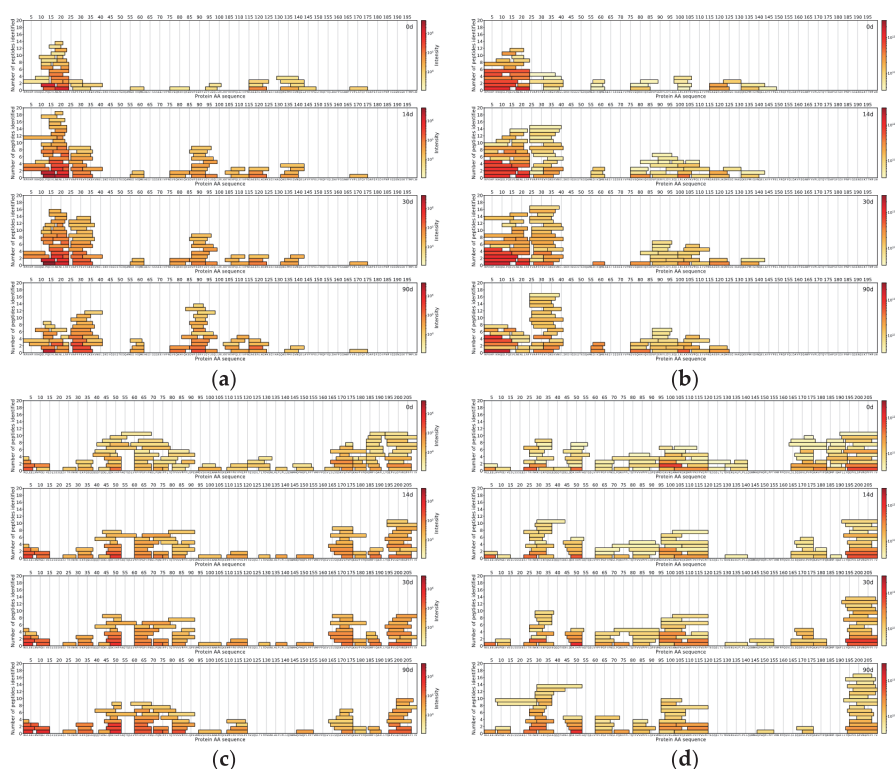


Figure 6. Cheese 3 α_1 -CN: (a) DIA-MS and (b) DDA-MS. Cheese 3 β -CN: (c) DIA-MS and (d) DDA-MS. The X-axis represents the casein amino acid sequence, the Y-axis represents a number of peptides and colour represents the logarithmically scaled intensity of peptides.

4. Discussion

During chromatographic method development, three peptide elution profiles were evaluated. The separation was performed with 40, 60 and 120 min analytical gradients to compare chromatographic performances and MS method compatibility. Figure 7 displays a base peak intensity chromatogram for a 40 min analytical gradient method. The narrowest extracted peak chromatogram was at 10 s at the base of the peak, providing a sufficient number of data points per peak (Figure S3). Therefore, as 60 and 120 min methods did not result in a higher number of identified peptides, a 40 min analytical gradient was selected as the one facilitating the best throughput.

Although the conventional DDA approach provides cleaner MS² spectra due to active isolation of the precursor, it suffers from a phenomenon known as data completeness problem [19]. While most abundant species get their corresponding MS² spectra recorded, less abundant species can potentially be missed. Due to sample-to-sample variation in analyte concentration in combination with precursor selection criteria even a peptide eluting at the same time might be missed out.

DIA is not only not subjected to the aforementioned drawback of DDA, but it also operates at a significantly higher acquisition speed due to minimising the time between MS¹ and MS² scan acquisition. Therefore, faster acquisition rate does not only allow to record data qualitatively, but also quantitatively. However, DIA-based methods with a quadrupole set for static transition exceedingly rely on chromatographic separation to minimise peptide coelution and hence, acquire cleaner MS² spectra.

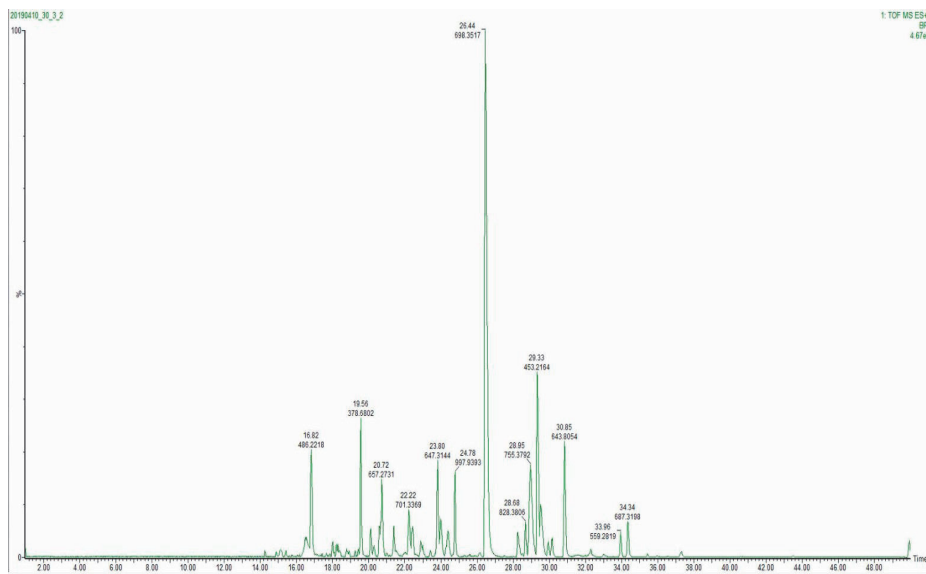


Figure 7. Base peak intensity (BPI) chromatogram of 40 min analytical gradient.

Furthermore, during synthesis in the mammary gland, caseins undergo post-translational changes in their primary structure [23]. One of the most important post-translational modifications in caseins is phosphorylation (at Ser, Thr and Tyr residues) and thus, analysis of phosphorylated peptides requires additional enrichment and purification step to decrease ion ionisation competition between non- and phosphorylated peptides [24]. With the current method, it is not possible to robustly analyse the peptides with every possible modification.

In recent years methods of further enhancement of a conventional DIA approach are gaining significant popularity. Active scanning (SONAR[®], Waters and Scanning SWATH[®]) or stepped (SWATH[®], Sciex) quadrupole and ion mobility separation (HDMS^E[®], Waters and PASEF[®], Bruker)-based DIA methods further expand the capabilities of DIA [25]. Active scanning or stepping quadrupole-based DIA methods significantly improve spectral clarity of MS² spectra by allowing fragmentation of only the ions confining within a quadrupole transmission profile. However, it loses a portion of the beam not confining to a quadrupole transmission window and hence, results in decreased overall sensitivity. Ion mobility separation based DIA, on the other hand, operates under a principle of preion mobility separation ion accumulation and subsequent release and hence, does not suffer from the ion loss of the quadrupole-based methods. As fragments can only exist when a precursor is present and fragments are inheriting the same drift time as the precursor due to the fact that mobility separation takes place before the fragmentation, it has been reported that ion mobility separation achieves a similar type of MS² clarity using the alignment of drift times and chromatographic profile of a precursor and fragments (HDMS^E[®]/PASEF[®]) [26,27]. Implementation of enhanced DIA methods would allow for even faster gradients and is worth further investigation.

A cut-off filter (3 kDa) was selected for sample preparation due to the unclear interaction of shorter cheese peptides with reversed-phase solid-phase extraction. In our work, we observed bias towards shorter peptides compared to Taivosalo et al. [14]. This bias could have been caused by either a natural bias of the given system towards shorter peptides, a decreased loss of shorter peptides due to not using of reverse-phase solid-phase extraction, or an increased loss of longer peptides due to the use of 3 kDa cut-off filter. As the overall number of peptides identified was lower than anticipated the use of

3 kDa cut-off filter should be further reviewed for its performance against conventional reversed-phase solid-phase extraction methods.

5. Conclusions

A rapid method was developed and successfully applied to the cheese peptidomics studies. The study allowed to indicate differences in cheese ripening caused due to the use of different starter cultures. Further methodology development is possible via the deployment of enhanced DIA approaches. Enhanced transmission of shorter peptides presents additional interest for future studies due to recorded bioactivity and sensory effects.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2304-8158/9/8/979/s1>, Figure S1: Summed peptide intensities of Cheese 1, 2 and 3 at the day 90, Figure S2: Peptide length distribution across 3 cheeses at the day 90, Figure S3: Overlay of Base Peak Intensity and Extracted Ion Chromatogram for narrowest peak corresponding to an identified peptide.

Author Contributions: Conceptualisation, G.A., A.T. and S.J.; methodology, G.A.; software, G.A. and S.V.; instrumental setup Z.D.; data curation, S.J. and R.R.; writing—original draft preparation, G.A.; writing—review and editing, A.T., D.P., S.J. and R.V.; visualisation, D.P. and T.L.; project administration, G.A.; funding acquisition, R.V. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by “TUT Institutional Development Program for 2016–2022” Graduate School in Biomedicine and Biotechnology receiving funding from the European Regional Development Fund under program ASTRA 2014–2020.4.01.16-0032 in Estonia.

Acknowledgments: Authors would like to acknowledge Tiina Kriščiunaite for input in data interpretation.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Fox, P.F.; McSweeney, P.L.H. Proteolysis in cheese during ripening. *Food Rev. Int.* **1996**, *12*, 457–509. [[CrossRef](#)]
2. Fontenele, M.A.; Bastos, M.S.R.; dos Santos, K.M.O.; Bemquerer, M.P.; do Egito, A.S. Peptide profile of Coalho cheese: A contribution for Protected Designation of Origin (PDO). *Food Chem.* **2017**, *219*, 382–390. [[CrossRef](#)]
3. Galli, B.D.; Baptista, D.P.; Cavalheiro, F.G.; Negrão, F.; Eberlin, M.N.; Gigante, M.L. Peptide profile of Camembert-type cheese: Effect of heat treatment and adjunct culture *Lactobacillus rhamnosus* GG. *Food Res. Int.* **2019**, *123*, 393–402. [[CrossRef](#)]
4. Sforza, S.; Cavatorta, V.; Lambertini, F.; Galaverna, G.; Dossena, A.; Marchelli, R. Cheese peptidomics: A detailed study on the evolution of the oligopeptide fraction in Parmigiano-Reggiano cheese from curd to 24 months of aging. *J. Dairy Sci.* **2012**, *95*, 3514–3526. [[CrossRef](#)]
5. Sánchez-Rivera, L.; Martínez-Maqueda, D.; Cruz-Huerta, E.; Miralles, B.; Recio, I. Peptidomics for discovery, bioavailability and monitoring of dairy bioactive peptides. *Food Res. Int.* **2014**, *63*, 170–181. [[CrossRef](#)]
6. Nielsen, S.D.; Beverly, R.L.; Qu, Y.; Dallas, D.C. Milk bioactive peptide database: A comprehensive database of milk protein-derived bioactive peptides and novel visualization. *Food Chem.* **2017**, *232*, 673–682. [[CrossRef](#)]
7. Ardö, Y.; Lilbæk, H.; Kristiansen, K.R.; Zakora, M.; Otte, J. Identification of large phosphopeptides from β -casein that characteristically accumulate during ripening of the semi-hard cheese Herrgård. *Int. Dairy J.* **2007**, *17*, 513–524. [[CrossRef](#)]
8. Karametsi, K.; Kokkinidou, S.; Ronningen, I.; Peterson, D.G. Identification of Bitter Peptides in Aged Cheddar Cheese. *J. Agric. Food Chem.* **2014**, *62*, 8034–8041. [[CrossRef](#)] [[PubMed](#)]
9. Fallico, V.; McSweeney, P.L.H.; Horne, J.; Pediliggieri, C.; Hannon, J.A.; Carpino, S.; Licitra, G. Evaluation of Bitterness in Ragusano Cheese. *J. Dairy Sci.* **2005**, *88*, 1288–1300. [[CrossRef](#)]
10. Baptista, D.P.; Galli, B.D.; Cavalheiro, F.G.; Negrão, F.; Eberlin, M.N.; Gigante, M.L. *Lactobacillus helveticus* LH-B02 favours the release of bioactive peptide during Prato cheese ripening. *Int. Dairy J.* **2018**, *87*, 75–83. [[CrossRef](#)]
11. Reale, A.; Ianniello, R.G.; Ciocia, F.; Di Renzo, T.; Boscaino, F.; Ricciardi, A.; Coppola, R.; Parente, E.; Zotta, T.; McSweeney, P.L.H. Effect of respirative and catalase-positive *Lactobacillus casei* adjuncts on the production and quality of Cheddar-type cheese. *Int. Dairy J.* **2016**, *63*, 78–87. [[CrossRef](#)]

12. Gagnaire, V.; Mollé, D.; Herrouin, M.; Léonil, J. Peptides Identified during Emmental Cheese Ripening: Origin and Proteolytic Systems Involved. *J. Agric. Food Chem.* **2001**, *49*, 4402–4413. [[CrossRef](#)] [[PubMed](#)]
13. Sforza, S.; Ferroni, L.; Galaverna, G.; Dossena, A.; Marchelli, R. Extraction, Semi-Quantification, and Fast On-line Identification of Oligopeptides in Grana Padano Cheese by HPLC–MS. *J. Agric. Food Chem.* **2003**, *51*, 2130–2135. [[CrossRef](#)] [[PubMed](#)]
14. Taivosalo, A.; Kriščiunaite, T.; Seiman, A.; Part, N.; Stulova, I.; Vilu, R. Comprehensive analysis of proteolysis during 8 months of ripening of high-cooked Old Saare cheese. *J. Dairy Sci.* **2018**, *101*, 944–967. [[CrossRef](#)]
15. Baptista, D.P.; Araújo, F.D.S.; Eberlin, M.N.; Gigante, M.L. A Survey of the Peptide Profile in Prato Cheese as Measured by MALDI-MS and Capillary Electrophoresis: Peptide profile of Prato cheese. *J. Food Sci.* **2017**, *82*, 386–393. [[CrossRef](#)]
16. Rehn, U.; Petersen, M.A.; Saedén, K.H.; Ardö, Y. Ripening of extra-hard cheese made with mesophilic DL-starter. *Int. Dairy J.* **2010**, *20*, 844–851. [[CrossRef](#)]
17. Toelstede, S.; Hofmann, T. Sensomics Mapping and Identification of the Key Bitter Metabolites in Gouda Cheese. *J. Agric. Food Chem.* **2008**, *56*, 2795–2804. [[CrossRef](#)]
18. Yılmaz, M.T.; Kesmen, Z.; Baykal, B.; Sagdic, O.; Kulen, O.; Kacar, O.; Yetim, H.; Baykal, A.T. A novel method to differentiate bovine and porcine gelatins in food products: NanoUPLC-ESI-Q-TOF-MSE based data independent acquisition technique to detect marker peptides in gelatin. *Food Chem.* **2013**, *141*, 2450–2458. [[CrossRef](#)]
19. Hernández-Mesa, M.; Escourrou, A.; Monteau, F.; Le Bizec, B.; Dervilly-Pinel, G. Current applications and perspectives of ion mobility spectrometry to answer chemical food safety issues. *Trends Anal. Chem.* **2017**, *94*, 39–53. [[CrossRef](#)]
20. The UniProt Consortium. The Universal Protein Resource (UniProt). *Nucleic Acids Res.* **2008**, *36*, D190–D195. [[CrossRef](#)]
21. Oliveros, J.C.; Venny. An Interactive Tool for Comparing Lists with Venn’s Diagrams. 2007–2015. Available online: <https://bioinfogp.cnb.csic.es/tools/venny/index.html> (accessed on 15 December 2019).
22. Sebald, K.; Dunkel, A.; Hofmann, T. Mapping Taste-Relevant Food Peptidomes by Means of Sequential Window Acquisition of All Theoretical Fragment Ion–Mass Spectrometry. *J. Agric. Food Chem.* **2019**. [[CrossRef](#)] [[PubMed](#)]
23. Farrell, H.M.; Jimenez-Flores, R.; Bleck, G.T.; Brown, E.M.; Butler, J.E.; Creamer, L.K.; Hickss, C.L.; Hollar, C.M.; NG-Kwai-Hang, K.F.; Swaisgood, H.E. Nomenclature of the proteins of cows’ milk—sixth revision. *J. Dairy Sci.* **2004**, *87*, 1641–1674. [[CrossRef](#)]
24. Larsen, M.; Thingholm, T.; Jensen, O.; Roepstorff, P.; Jorgensen, T. Highly Selective Enrichment of Phosphorylated Peptides from Peptide Mixtures Using Titanium Dioxide Microcolumns. *Mol. Cell Proteomics* **2005**, *4*, 873–886. [[CrossRef](#)]
25. Ludwig, C.; Gillet, L.; Rosenberger, G.; Amon, S.; Collins, B.C.; Aebersold, R. Data-independent acquisition-based SWATH-MS for quantitative proteomics: A tutorial. *Mol. Syst. Biol.* **2018**, *14*, e8126. [[CrossRef](#)] [[PubMed](#)]
26. Alves, T.O.; D’Almeida, C.T.S.; Victorio, V.C.M.; Souza, G.H.M.F.; Cameron, L.C.; Ferreira, M.S.L. Immunogenic and allergenic profile of wheat flours from different technological qualities revealed by ion mobility mass spectrometry. *J. Food Compos. Anal.* **2018**, *73*, 67–75. [[CrossRef](#)]
27. Jeanne Dit Fouque, K.; Fernandez-Lima, F. Recent advances in biological separations using trapped ion mobility spectrometry–mass spectrometry. *Trends Anal. Chem.* **2019**, *116*, 308–315. [[CrossRef](#)]



Appendix 2

Publication II

Xia, X., Arju, G., Taivosalo, A., Lints, T., Kriščiunaite, T., Vilu, R., Corrigan, B.M., Gai, Nan., Fenelon, M.A., Toin, J.T., Kilcawley, K., Kelly, A.L., McSweeney, P.L.H., Sheehan, J.J. (2023) Effect of β -casein reduction and high heat treatment of micellar casein concentrate on proteolysis, texture and the volatile profile of Emmental cheese during ripening. *International Dairy Journal* 138, 105540. doi.org/10.1016/j.idairyj.2022.105540



ELSEVIER

Contents lists available at ScienceDirect

International Dairy Journal

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

Effect of β -casein reduction and high heat treatment of micellar casein concentrate on proteolysis, texture and the volatile profile of resultant Emmental cheese during ripening

Xiaofeng Xia^{a, b}, Georg Arju^{c, d}, Anastassia Taivosalo^c, Taivo Lints^c, Tiina Kriščiunaite^c, Raivo Vilu^{c, d}, Bernard M. Corrigan^a, Nan Gai^b, Mark A. Fenelon^a, John T. Tobin^a, Kieran Kilcawley^a, Alan L. Kelly^b, Paul L.H. McSweeney^b, Jeremiah J. Sheehan^{a, *}

^a Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, P61 C996, Ireland

^b School of Food and Nutritional Sciences, University College Cork, Cork, T12 YN60, Ireland

^c Centre of Food and Fermentation Technologies (TFTAK), Akadeemia Tee 15a, 12618, Tallinn, Estonia

^d Department of Chemistry and Biotechnology, School of Science, Tallinn University of Technology, Ehitajate Tee 5, 12616, Tallinn, Estonia

ARTICLE INFO

Article history:

Received 29 July 2022

Received in revised form

27 October 2022

Accepted 29 October 2022

Available online 8 November 2022

ABSTRACT

The effect of β -casein reduction (LB) and high heat treatment (HHT, 120 °C \times 15 s) of whey protein-reduced milk on the ripening properties of Emmental cheese was investigated. Cheeses were manufactured from low-heat skim milk powder (LH), micellar casein concentrate powder (MCC) and β -casein reduced MCC powder (LB- or HHT LB) respectively, then ripened for 120 days. There was no significant difference between the ripening properties of LB and MCC cheeses. Compared with the other treatments, HHT LB cheese had significantly lower plasmin activity (day 120 of ripening) and flowability (days 70 and 120 of ripening), higher level of a-value (redness), and similar numbers of peptides and levels of casein hydrolysis (days 1 and 120 of ripening). Overall, reduction of β -casein by 4.25% in MCC did not influence the maturation properties of Emmental cheese; however, high heat treatment of whey protein-reduced milk impaired cheese functionality and appearance.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

Microfiltration (MF) can selectively separate native whey protein from milk and produce micellar casein concentrate (MCC) as retentate and a virgin whey protein stream as permeate (Ardisson-Korat & Rizvi, 2004). The virgin whey protein is free from fat, chymosin, starter culture, cheese fines, cheese colorant, and caseinomacropptide, and has been called 'ideal whey' due to its superior technological and functional properties compared with cheese whey (Bacher & Königsfeldt, 2000). MF is usually carried out at 45–55 °C (warm MF) to achieve high permeate flux (Holland, Corredig, & Alexander, 2011). However, microfiltration of refrigerated milk at low temperature (4–15 °C, cold MF) is attracting research interest, as it can minimise bacterial growth and heat damage to protein (France, Kelly, Crowley, & O'Mahony, 2021; Raghunath & Hibbard, 1997).

Since the association of β -casein with casein micelles is temperature-dependent, holding milk at 1–4 °C for 12–48 h can

weaken the hydrophobic interactions between β -casein and casein micelles and liberate β -casein from the casein micelles into the milk serum phase (O'Mahony, Smith, & Lucey, 2014). As a result, in addition to native whey protein, cold MF can also remove a portion of β -casein from milk into the permeate (O'Mahony et al., 2014). β -Casein is (i) an effective stabiliser for emulsion and foam systems (Li, Auty, O'Mahony, Kelly, & Brodtkorb, 2016), (ii) a precursor for bioactive compounds (Kamiński, Cieślińska, & Kostyra, 2007), and (iii) an important component in infant milk formulae (Li et al., 2019). Thus, isolation of β -casein from milk is desirable. As the by-product of both β -casein isolate and whey protein isolate, β -casein-reduced MCC, when used for the production of model Cheddar cheese, showed increased meltability on heating (O'Mahony et al., 2014). However, the effect of using β -casein-reduced MCC as a basis for cheese-making on the proteolysis, volatile profile and other maturation indexes of such cheeses have not been reported previously.

Plasmin, the primary indigenous proteinase in milk, mainly acts on β -casein and is the most important milk-derived proteinase for cheese ripening, especially so for cheese types made with high cooking temperature (≥ 50 °C) like Emmental (Ardo, McSweeney,

* Corresponding author.

E-mail address: Diarmuid.Sheehan@teagasc.ie (J.J. Sheehan).

Magboul, Upadhyay, & Fox, 2017), due to inactivation of chymosin during such cooking. After heat treatment, the free thiol-disulphide groups exposed on the denaturation of β -lactoglobulin (β -LG) can form covalent bonds with the thiol-rich region of plasmin and plasminogen, and the activity of plasmin and plasminogen is reduced subsequently (Aaltonen & Ollikainen, 2011). Heating milk at 90 °C for 15 s significantly decreased the plasmin activity in the resultant cheese (Benfeldt, Sørensen, Ellegård, & Petersen, 1997). However, the thermal inactivation of plasmin is decreased by partly removing whey protein from milk using MF before heating (95 °C \times 15 s) (Aaltonen & Ollikainen, 2011). One focus of the current study was to determine whether the plasmin activity of whey protein-reduced milk produced by MF was influenced by application of a heat treatment of 120 °C \times 15 s, as opposed to 95 °C \times 15 s in previous studies.

Xia et al. (2022) reported that partially removing β -casein from milk did not influence the rennet coagulability of cheesemilk or the composition and yield of Emmental cheeses made therefrom; applying high heat treatment (120 °C \times 15 s) to whey protein-reduced milk significantly impaired its rennet coagulation properties without affecting the composition (except fat content) and yield of cheese made therefrom. The objective of the current study was to determine the effect of β -casein reduction and high heat treatment of micellar casein concentrate on the patterns and levels of proteolysis, texture and volatile profile of such Emmental cheeses during ripening.

Proteolysis is widely considered one of the most important activities during cheese maturation, and tests such as measurement of levels of pH 4.6 soluble nitrogen, urea-PAGE and plasmin activity test are typically conducted to evaluate cheese proteolysis. However, these measurements are only able to quantify the intact casein level in cheese and, overall levels of breakdown of α_{S1} -, α_{S2} - and β -casein without giving a more comprehensive and accurate picture of cheese proteolysis. LC-MS offers the capability to separate and identify peptides generated from cheese proteolysis, giving researchers much more information on casein breakdown pathways (Taivosalo et al., 2018). LC-MS has been used to characterise proteolysis in Cheddar cheese (Upadhyay et al., 2004), Camembert cheese (Mane & McSweeney, 2020) and Old Saare cheese (Taivosalo et al., 2018); one objective of this research was to use LC-MS to evaluate proteolysis in Emmental cheese for the first time.

2. Materials and methods

2.1. Chemicals

Ultrapure water (18.2 M Ω cm) was prepared with MilliQ® IQ 7000 equipped with LC-Pak (Merck KGaA, Darmstadt, Germany). Acetonitrile (MeCN; LiChrosolv, hypergrade for LC-MS) and formic acid (FA; LC-MS grade) were acquired from Sigma-Aldrich (Darmstadt, Germany). Hi3 *E.coli* STD (p/n: 186006012) and Leucine-Enkephalin (p/n: WT186006013) were purchased from Waters Corporation (Milford, MA, USA).

2.2. Preparation of Emmental cheese

Miniature (500 g) Emmental cheeses were manufactured from low heat skim milk powder (LHSMP), micellar casein concentrate powder (MCC powder) and β -casein reduced micellar casein concentrate powder (LB MCC powder) at pilot scale at Moorepark Technology Limited, Fermoy, Co. Cork, Ireland as described by Xia et al. (2022).

2.2.1. Powder production

Pasteurised skim milk (Dairygold, Mitchelstown, Co Cork, Ireland, 1000 kg) was chilled and divided into three portions; one portion (400 kg) of milk was evaporated and spray-dried, the powder generated therefrom was called LHSMP. The second portion (300 kg) was microfiltered at 47 °C, with 56.38% of whey protein being removed, and the resultant whey protein reduced-milk was called MCC. The third portion (300 kg) was microfiltered at 8.5 °C, depleting 53.60% of whey protein and 4.25% of β -casein from the milk to generate LB MCC. Both MCC and LB MCC were subsequently evaporated and spray-dried producing MCC or LB MCC powder respectively. All powders were packed, sealed and stored at 4 °C after production.

2.2.2. Cheese manufacture

On day 1, four fat-free milk types, i.e., LH-, MCC-, LB- and HHT LB milk, were prepared from LHSMP, MCC powder, LB MCC powder, milk permeate powders (Profile™ P8 permeate powder W463, Kerry Group, Listowel, Ireland), CaCl₂ (calcium chloride anhydrous powder Reag. Ph Eur, Merck Chemicals Ltd, Nottingham, UK) and RO water as described in Xia et al. (2022), the milk was stored at 4 °C. The contents of casein, lactose and ash in milks were standardised to 2.89%, 5.07% and 0.77%, respectively. On day 2, the LH-, MCC- and LB MCC milk samples were pasteurised at 72 °C for 15 s and HHT LB milk was high heat treated at 120 °C for 15 s. Milk samples were chilled to 4 °C immediately after heat treatment and were refrigerated at this temperature afterwards. On day 3, pasteurised cream was added to each milk to standardise the cheesemilks (CMs) to a casein-to-fat ratio of 0.79:1, and Emmental cheeses were then manufactured in 10 L cheese vats as described in Xia et al. (2022). Subsequent to pressing and brining, the cheeses were vacuum-packed and ripened at 8 °C from day 0 to day 14 (pre-ripening), at 22 °C from day 15–70 (warm room ripening) and at 4 °C from day 71–120 (cold room ripening). Quadruplicate trials (trials 1, 2, 3 and 4) were carried out.

2.3. Texture, pH, colour and flowability

Emmental cheeses were sampled at day 1, 14, 70 and 120 of ripening, where the texture profile (fracture stress, fracture strain and firmness, tested at day 1 and 14 of maturation), pH, colour (levels of L*, a* and b*) and flowability (tested at day 70 and 120 of maturation) of cheese samples were measured as described by Hou, Hannon, McSweeney, Beresford, and Guinee (2012), Fenelon, O'connor, and Guinee (2000), Sheehan, Patel, Drake, and McSweeney (2009) and McCarthy, Wilkinson, Kelly, and Guinee (2016), respectively.

2.4. Plasmin activity

The plasmin activity in LB MCC milks (unheated, pasteurised or high heat treated), cheesemilks and experimental cheeses at day 120 of maturation was monitored by the coumarin peptide assay (Richardson & Pearce, 1981). Results were expressed as nmol AMC mL⁻¹ min⁻¹, which equates to one unit of plasmin activity.

2.5. Urea-PAGE and pH 4.6 SN

Emmental cheeses were sampled at day 1, 14, 70 and 120 of ripening, and primary proteolysis levels were determined in two ways: (i) by urea-polyacrylamide gel electrophoresis (PAGE) on a cheese weight basis as described by Lamichhane, Sharma, Kennedy, Kelly, and Sheehan (2019) and Xia et al. (2021); (ii) by measuring levels of % pH 4.6 soluble nitrogen in a fixed weight of cheese (pH 4.6) as described by Fenelon et al. (2000).

2.6. Liquid chromatography–mass spectrometry (LC–MS)

The peptide profile in the pH 4.6 soluble extracts from day 1 and 120 of ripening in trial 1 and trial 2 were measured by LCMS in triplicate.

2.6.1. Sample preparation

Prior to peptide analysis, pH 4.6- water-soluble extracts (100 μ L) were mixed with acetonitrile (100 μ L) in Eppendorf® tubes (Eppendorf AG, Germany) to precipitate proteins, vortexed for 30 s and centrifuged at 11,200 \times g for 15 min. Then, the supernatant (100 μ L) and MilliQ® water (900 μ L) were transferred into a new Eppendorf® tube and vortexed for 30 s. The obtained mixture (130 μ L) was transferred into a low-volume insert and spiked with 20 μ L of 1 pmol μ L⁻¹ Hi3 *E.coli* STD as housekeeping ions.

2.6.2. LC–MS

Peptides were identified using liquid chromatography–mass spectrometry. The system comprised a Waters I-Class Plus (BSM, SM-FL and CH-A) ultra-performance liquid chromatography (UPLC®) system hyphenated to a Waters Vion ion mobility separation-quadrupole time-of-flight (IMS-QTOF) mass spectrometer (3.1.0) equipped with LockSpray exact mass ionisation source and controlled by Waters UNIFI (1.9.4). Nitrogen was used as collision gas. Samples (1 μ L) were injected onto Waters Acquity UPLC® HSS T3 column (1.8 μ m, 1 \times 150 mm, p/n:186003537) in triplicate. The column temperature was maintained at 40 °C while samples were kept pre-injection at 8 °C in a temperature-controlled autosampler. Mobile phase A consisted of 0.1% formic acid in MilliQ® water, mobile phase B consisted of 0.1% formic acid in acetonitrile, weak needle wash consisted of 90:10 MilliQ® water: acetonitrile, strong needle wash consisted of 10:90 MilliQ® water: acetonitrile, and seal wash consisted of 50:50 MilliQ® water: acetonitrile. The initial analytical flow rate was 0.1 mL min⁻¹. Samples were eluted with the following gradient: 0–0.5 min hold at 1% B, 0.5–40.5 min linear gradient 1–30% B, 40.5–42.5 min linear gradient 50–99% B accompanied by linear flow rate increase 0.1–0.2 mL min⁻¹, 42.5–44.5 min hold at 99% B, 44.5–45 min linear gradient 99–1% B accompanied by linear flow rate decrease 0.2–0.1 mL min⁻¹, 45–48.5 min hold at 1% B.

The mass spectrometer was operated in positive polarity and sensitivity mode (35,000 FWHM at 556.2771 *m/z*). Data were acquired in HDMSE mode with a scan time of 0.3 s. The recorded mass range was from 50 to 2000 *m/z* for both low and high energy spectra. The collision energy was ramped from 10 to 63 V. Capillary voltage was set to 0.5 kV, cone voltage to 30 V, and source offset to 50 V. Source temperature was set to 120 °C, desolvation temperature to 500 °C, desolvation gas flow to 700 L h⁻¹ and cone gas flow to 50 L h⁻¹. The system was calibrated using Waters Major Mix (p/n:186008113) and a solution of leucine enkephalin (Leu–Enk, 50 pg μ L⁻¹) was infused at 15 μ L min⁻¹ for automatic lock mass correction every 5 min, as well as before and after each run.

2.6.3. Raw data processing and analysis

Peptide identification and quantification was carried out using UNIFI Large Molecule Package with Peptide Map (IMS) workflow. Peptide mapping was carried out using UNIFI Large Molecule Package (LMP) by searching for peptides against α _{S1}-casein (α _{S1}-CN; P02662), α _{S2}-casein (α _{S2}-CN; P02663), β -casein (β -CN; P02666) and κ -casein (κ -CN, P02668) sequences from bovine species obtained with the UniProt database. Peptide search was set to non-specific digest with no missed cleavages allowed, and minimum sequence length was set to 4. The following filters were applied: mass error between –5 and 5 ppm, matched first gen primary ions greater than or equal to 1, fragment label not containing in-source

fragments, loss of H₂O and loss of NH₃. The following variable post-translational modifications were used in the analysis: oxidation (M), acetyl-(protein N-terminal), deamidation (NQ) and phosphorylation (STY).

Processed data from UNIFI were exported into Microsoft Excel for grouping, alignment, and filtration based on the retention time (0.1 min), observed molecular mass (0.01 Da per charge) and collision cross section (10% RSD). The aligned data were then further used to construct a data matrix for revision and removal of outliers and false-positive identifications.

The identified and quantified peptides, as well as their normalised and averaged intensities, were further analysed and visualised with the help of in-house data analysis and visualisation scripts written in the Python™ programming language (Python Software Foundation). Averaged values of peptide intensities of technical triplicates were further averaged to produce average peptide intensity values of biological replicates. Both sets of data were plotted as peptide maps and cleavage sites, as described by Taivosalo et al. (2018).

2.7. Volatile compounds

The volatile profiles of each of the cheeses were measured at day 120 of ripening by solid phase micro-extraction (SPME)–gas chromatography–mass spectrometry (GC–MS) as described by Lamichhane et al. (2018b). All tests were performed in triplicate.

2.8. Statistical analysis

Emmental cheeses were manufactured in quadruplicate pilot plant trials over a period of 5 weeks. One-way ANOVA analysis was carried out to compare the effect of treatments on the plasmin activities in milk samples, cheesemilk and cheeses as well as on the volatile compounds in experimental cheeses, using IBM SPSS statistics 24.0 (IBM Corp., 2016, Chicago, IL, USA). The effect of treatments, ripening time and their interactions on the levels of pH, pH 4.6-SN, texture profile, flowability and colour in cheeses were determined by a split-plot design, split-plot design analysis was carried out using the PROC MIXED procedure of SAS software version 9.4 (SAS Institute, 2012). The level of significance ($P < 0.05$) of differences between means was determined by Tukey's test.

3. Results and discussion

3.1. Composition of cheese

The gross composition and pH of experimental cheeses at day 14 of maturation were reported in detail by Xia et al. (2022) and are summarised in Supplementary material Table S1. In brief, among the four cheeses types, LH cheese had the highest contents of moisture and MNFS, as well as the lowest contents of calcium and calcium per gram of protein, MCC cheese had the highest pH, and HHT LB cheese had the lowest FDM content. The levels of protein, salt and S/M were similar between the experimental cheeses.

3.2. pH

Curd washing with water is sometimes applied to decrease the residual lactose content in Emmental cheeses (Fröhlich-Wyder et al., 2017; Heino, 2008) and, as a result of this and of metabolism of lactate, the pH of Emmental cheeses often increases from ~5.2 in day 1–5.6–5.7 by day 52–180 (Lamichhane, Kelly, & Sheehan, 2018a; Lawrence, Creamer, & Gilles, 1987; Lopez, Camier, & Gassi, 2007). However, the pH of experimental Emmental cheeses in the current study decreased significantly

Table 1

Statistical significances (adjusted *P*-values) for mean changes of maturation indexes during ripening in Emmental cheeses as affected by treatment or time as well as the interactive effect of treatment and time.^a

Parameter	Treatment	Time	Interactive effect (treatment × time)
pH	NS	***	NS
% pH 4.6 SN 100 g ⁻¹ cheese	NS	***	NS
Firmness (N)	*	NS	*
Fracture stress (N)	NS	NS	NS
Fracture strain	NS	NS	NS
Flowability (%)	***	NS	NS
L* (whiteness)	NS	***	NS
a* (redness)	***	***	NS
b* (yellowness)	NS	***	NS

^a The levels of pH, % pH 4.6 SN 100 g⁻¹ cheese, L*, a* and b* in cheese were measured and compared at day 1, 14, 70 and 120 of ripening separately; the firmness, fracture stress and fracture strain in cheese were monitored at day 1 and day 14 of maturation respectively; the cheese flowability was determined at day 70 and 120. Significant levels: NS, *P* > 0.05; **P* < 0.05; ***P* < 0.01; ****P* < 0.001.

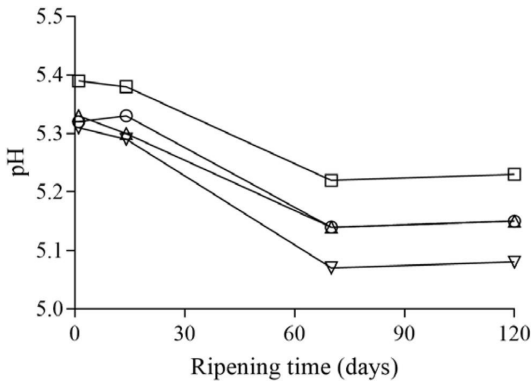


Fig. 1. pH during ripening of Emmental cheeses manufactured from low heat skim milk powder (○), micellar casein concentrate powder (□) or β-casein reduced micellar casein concentrate powder that were pasteurised (△) or high heat treated (▽).

during ripening, from 5.30–5.40 on day 1 to 5.10–5.25 on day 120 (*P* < 0.0001, Table 1, Fig. 1). The lactose content of the four experimental milks was standardised to 5.07%, which was close to the average lactose content in bovine milk (4.9%) (Fox, Uniacke-Lowe, McSweeney, & O'Mahony, 2015). The absence of curd washing to control final cheese pH in this study might have led to the presence of residual lactose in the experimental cheeses, resulting in a decrease in pH in cheese over maturation due to bacterial fermentation of residual lactose (O'Sullivan, McSweeney, Cotter, Giblin, & Sheehan, 2016).

The pH of cheese can influence cheese texture by affecting the level of casein hydration and the levels of primary proteolysis in cheese through modulating the activity of plasmin or residual chymosin (Hickey et al., 2017; Li et al., 2020). Similarly, levels of primary proteolysis can also affect the intensity of flavour in cheese (McSweeney & Sousa, 2000) through the production of substrates for secondary proteolysis. By obtaining the desired pH value for a given cheese type, cheesemakers can produce cheeses with desired texture, taste and flavour characteristics. During the formulation of MCC-, LB- and HHT LB milks, milk permeate powder was added to fortify the lactose content in milk. By adding a lower quantity of milk permeate powder, the lactose content in these milks could be decreased from 5.07% to a level where water addition could be eliminated to achieve the desired cheese pH (Heino, 2008).

Although the MCC cheese had a significantly higher pH at d 14 of ripening, when mean pH levels across ripening were considered, there was no significant effect of reduction of β-casein levels in the cheesemilk nor application of high heat treatment (120 °C × 15 s) to the milk on cheese pH during maturation, which may be due to their similar contents of protein, moisture and ash (Supplementary material Table S1).

3.3. Texture

The firmness of experimental cheeses sampled on day 1 of ripening was not significantly different from that of cheese sampled on day 14 (Table 1), which was not surprising, since the main decrease of firmness in brine-salted semi-hard cheeses is usually observed during warm room ripening (Lamichhane et al., 2019). Due to eye formation after 14 days of maturation, the texture of Emmental cheeses was not measured at later time points.

On both day 1 and day 14 of ripening, the firmness and fracture stress in MCC cheese were higher than the other cheeses (Fig. 2), which might be related to the higher calcium content (Lawrence, Heap, & Gilles, 1984) and significantly higher pH level in this cheese (Supplementary material Table S1). A higher cheese pH can increase the amount of calcium bound to para-casein, resulting in greater levels of para-casein aggregation and a firmer cheese gel structure, and thus a higher cheese firmness (Hou, Hannon, McSweeney, Beresford, & Guinee, 2014). O'Mahony et al. (2014) reported that removing 9–10% of the β-casein from milk significantly reduced the firmness of a model Cheddar cheese. However, and possibly due to the much lower level of β-casein reduction in the present study (4.25%), no significant differences in levels of fracture stress, fracture strain and firmness between LH cheese and cheeses LB- and HHT LB were observed (*P* > 0.05).

3.4. Plasmin activity

The plasmin activity in LB MCC decreased significantly when increasing the heat treatment from no heat treatment to pasteurisation (72 °C for 15 s) to HHT (120 °C for 15 s) (Fig. 3A). Since only 53.60% of whey proteins were removed from LB MCC by microfiltration (data not shown), heat-induced denaturation of β-LG and binding with plasmin and plasminogen might explain the very low plasmin activity in HHT LB MCC. As a result, the cheese milk and Emmental cheese made from HHT LB MCC also had a much lower plasmin activity compared with those prepared from PS LB MCC (Fig. 3B and C).

No significant difference in plasmin activity was observed between MCC- and LB CM and their associated cheeses, as expected (Fig. 3B and C). Whey proteins, especially β-LG, can decrease the activity of plasmin and plasminogen activators in milk and, as both plasmin and plasminogen are associated with casein micelles (France et al., 2021), they can be retained in whey protein-reduced MCC after microfiltration, resulting in MCC with higher plasmin activity than its original feed milk, as shown by Aaltonen and Ollikainen (2011). Thus, it was not surprising to observe a significantly lower plasmin activity in LH CM compared with those in MCC- and LB CM (Fig. 3B), with similar results also reported by Li et al. (2020). However, this difference was not carried forward to the respective cheeses, in contrast to the findings of Li et al. (2020). Further research is required to determine the reason for this.

3.5. Urea-PAGE

The patterns of proteolysis of α_{S1}- and β-caseins in Emmental cheeses were evaluated using urea-PAGE, where cheese samples of similar protein levels were loaded on a weight basis (Fig. 4). For

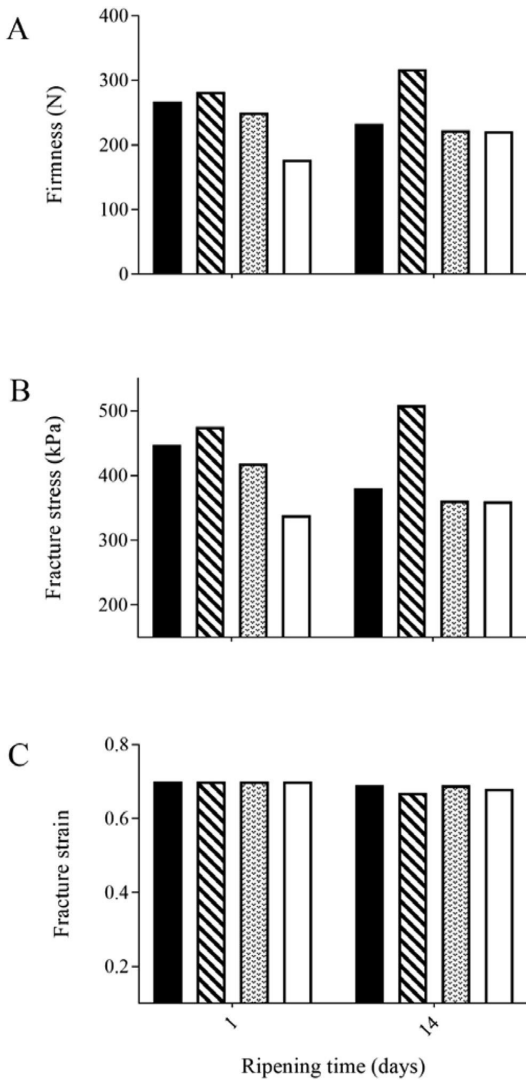


Fig. 2. Fracture stress (A), fracture strain (B) and firmness (C) of Emmental cheeses manufactured from low heat skim milk powder (○), micellar casein concentrate powder (□) or β-casein reduced micellar casein concentrate powder that were pasteurised (Δ) or high heat treated (▽) at day 1 and day 14 of ripening.

each treatment, significant hydrolysis of α_{S1} - and β -caseins was observed during warm-room ripening, with little further breakdown during cold-room ripening. The high temperature (22 °C) of the warm ripening can accelerate the hydrolysis of both α_{S1} - and β -caseins, as reported by Sheehan, Wilkinson, and McSweeney (2008) and Lamichhane et al. (2019).

From day 14 of maturation onwards, proteolysis of α_{S1} - and β caseins in MCC-, LB- and HHT LB cheeses was much more extensive than those in the LH cheese. Plasmin has been widely considered the main proteinase that hydrolyses β -casein during the course of cheese maturation (Ardo et al., 2017). The level of intact β casein in LH cheese was much higher than those in cheeses MCC- and LB, suggesting a lower plasmin activity in LH cheese as a result of lower

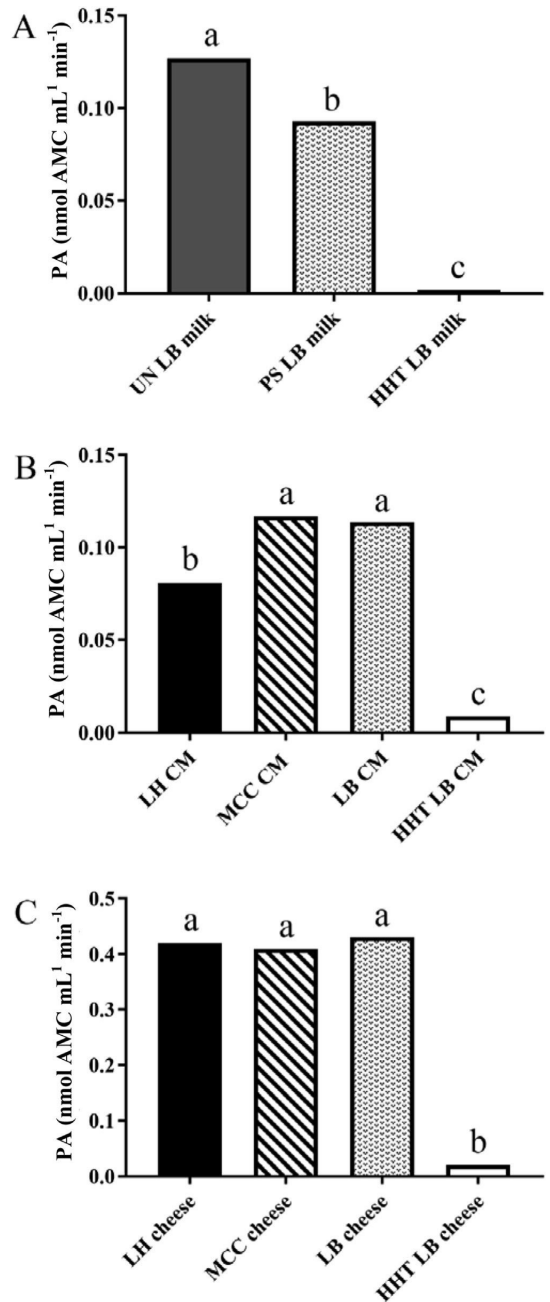


Fig. 3. Plasmin activity in micellar casein concentrates (A), cheese milks (B) and Emmental cheeses at day 120 of ripening (C). Abbreviations are: plasmin activity (PA), milk prepared from β-casein reduced micellar casein concentrate were not heated (UN LB milk); pasteurised (PS LB milk) or high heat treated (HHT LB milk), cheese milks were prepared from low heat skim milk powder (LH CM), micellar casein concentrate powder (MCC CM) or β-casein reduced micellar casein concentrate powder that were pasteurised (LB CM) or high heat treated (HHT LB CM); the Emmental cheeses made therefrom were called LH-, MCC-, LB- or HHT LB cheese respectively. Results are means of quadruplicate trials, values within a figure not sharing the same lowercase letter differ significantly ($P < 0.05$).

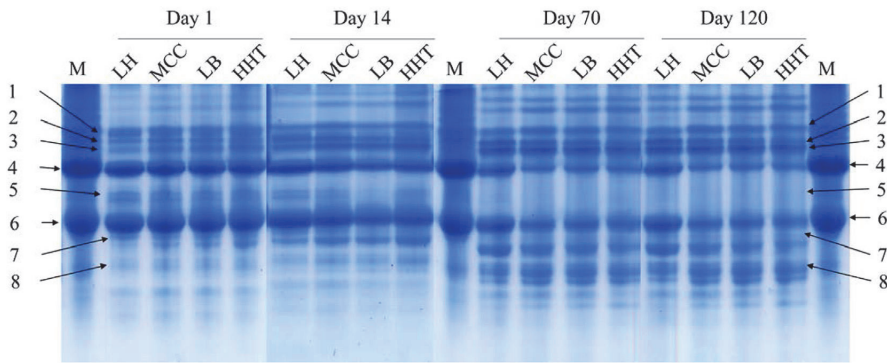


Fig. 4. Urea-polyacrylamide gel electrophoretograms of Emmental cheeses at day 1, 14, 70 and 120 of ripening. Abbreviations are: M, marker (sodium caseinate); Emmental cheeses manufactured from low heat skim milk powder (LH), micellar casein concentrate powder (MCC) or β -casein reduced micellar casein concentrate powder that were pasteurised (LB) or high heat treated (HHT). Numbers are: 1, β -casein (f106–209); 2, β -casein (f29–209); 3, β -casein (f108–209); 4, β -casein; 5, β -casein (f1–192); 6, α_{S1} -casein; 7, α_{S1} -casein (f102–199); 8, α_{S1} -casein (f24–199).

plasmin activity in its cheese milk. A much higher level of β -casein breakdown was expected in the MCC- and LB cheeses than that in HHT LB cheese due to their higher plasmin activities (Fig. 3C); however, similar levels of proteolysis of β -casein in these three cheeses suggest that analytical approaches more sensitive than urea-PAGE, such as mass spectrometry, may be required to discriminate the effect of treatments on cheese proteolysis.

Hydrolysis of α_{S1} -casein was also observed in the four types of experimental cheese. Both *Cryphonectria parasitica* proteinase (also called *Endothia parasitica* proteinase, used as coagulant in this research) (Gagnaire, Mollé, Herrouin, & Léonil, 2001) and cathepsin D (Hayes et al., 2001) can hydrolyse α_{S1} -casein during cheese ripening. It has been reported that the activity of *C. parasitica* proteinase in cheese may be totally inactivated after cooking at 50–55 °C for 30–60 min (Garnot & Mollé, 1987; Winwood, 1989). However, in this study, the curds were cooked to 50 °C and drained immediately after reaching this temperature; and the size of the derived Emmental cheeses (0.5 kg per cheese) were also much smaller than those of industrial Emmental (75 and 120 kg) (Fröhlich-Wyder et al., 2017). As a result, the temperature of the experimental cheeses is likely to have decreased from 50 °C towards room temperature shortly after cooking and, for the reason of having an elevated temperature for a relatively short time, the *C. parasitica* proteinase in experimental cheeses might only have been partly inactivated.

As the indigenous milk proteinase cathepsin D can partly survive pasteurisation of cheesemilk (72 °C \times 15 s, 8% survival), cooking in cheesemaking (55 °C \times 30 min, 45% survival) (Hayes et al., 2001) or even heating at 99 °C for 60 s (Larsen et al., 2000), it was possible that cathepsin D also played a role in the hydrolysis of α_{S1} -casein in experimental cheeses during maturation. The levels of α_{S1} -casein hydrolysis in MCC-, LB- and HHT LB cheeses were much higher than that in LH cheese. However, more studies are required to (i) elucidate which enzyme(s) is responsible for the hydrolysis of α_{S1} -casein in this research and (ii) explain why LH cheese had higher levels of intact α_{S1} -casein than MCC-, LB and HHT LB cheeses.

3.6. Levels of pH 4.6 SN 100 g⁻¹ cheese

The level of primary proteolysis in cheese was expressed as a percentage of pH4.6 soluble nitrogen (SN) 100 g⁻¹ cheese, where a higher content of pH 4.6 SN 100 g⁻¹ cheese indicates a higher level of primary proteolysis in experimental cheese (McCarthy,

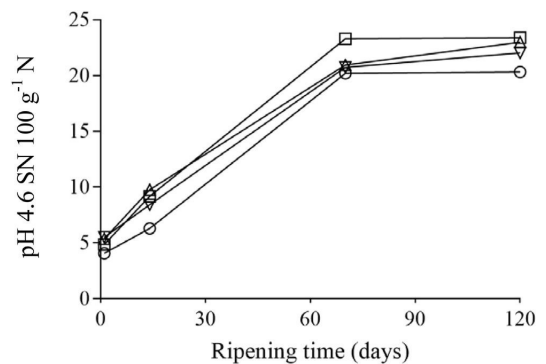


Fig. 5. pH 4.6 SN 100 g⁻¹ cheese of Emmental cheeses manufactured from low heat skim milk powder (\circ), micellar casein concentrate powder (\square) or β -casein reduced micellar casein concentrate powder that were pasteurised (Δ) or high heat treated (∇).

Wilkinson, & Guinee, 2017). The levels of pH4.6-SN 100 g⁻¹ cheese in all experimental cheeses increased significantly during the warm room period (day 14 to day 70, 22 °C, $P < 0.0001$) (Table 1, Fig. 5), in line with the previous reports for Swiss-type (O'Sullivan et al., 2016; Sheehan et al., 2008) and Maasdam (Lamichhane et al., 2018a) cheeses. However, the change was not significant from day 1 to day 14 (8 °C, $P > 0.05$) or from day 70 to day 120 (4 °C, $P > 0.05$) of maturation (results not shown), presumably due to the lower ripening temperature.

Variations in the contents of whey protein, β -casein and the thermal history of the cheese milk did not affect the levels of pH 4.6 SN 100 g⁻¹ cheese ($P > 0.05$), suggesting that these treatments did not affect levels of primary proteolysis. This contrasted with urea-PAGE results, where it was seen that LH cheese had much higher levels of intact casein than those in MCC, LB and HHT LB cheeses at day 14, 70 and 120 of ripening. As a result, LC–MS was carried out to provide a more detailed understanding of cheese proteolysis.

3.7. Peptide analysis by LC–MS

A total number of 177 peptides with unique sequences were identified in pH 4.6-soluble extracts from LH cheese on day 1 of ripening, and this number increased to 488 on day 120, as a result

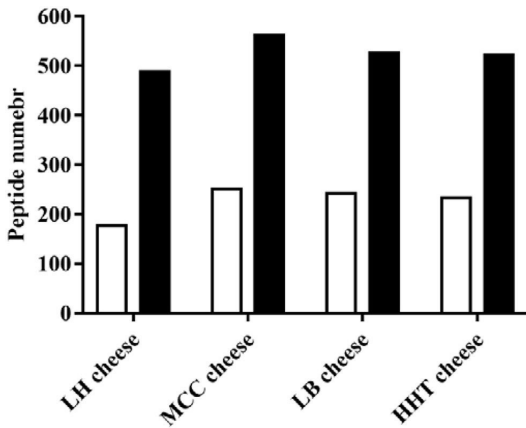


Fig. 6. Number of unique peptides identified in pH4.6 soluble extracts of Emmental cheeses manufactured from low heat skim milk powder (LH cheese), micellar casein concentrate powder (MCC cheese) or β -casein reduced micellar casein concentrate powder that were pasteurised (LB cheese) or high heat treated (HHT LB cheese) at day 1 (white) or day 120 (black) of ripening.

of ongoing proteolysis (Fig. 6). A higher number of peptides was identified from MCC, LB, and HHT LB cheeses compared with the LH cheese on both ripening days (Fig. 6). The total intensity of peptides originating from α_{S1} -, α_{S2} -, β - or κ -casein also increased from day 1 to day 120 of maturation (Fig. 7) for all cheeses.

At both ripening times, the overall numbers of κ -casein-derived peptides were much lower than those of α_{S1} -, α_{S2} - or β -casein-derived peptides (Fig. 7) as was also found in Old Saare cheese (Taivosalo et al., 2018) and Irish farmhouse Camembert cheese (Mane & McSweeney, 2020). On both day 1 and day 120 of ripening, the summed intensities of peptides originating from α_{S1} -, α_{S2} -, β - or κ -casein in MCC, LB and HHT LB cheeses were roughly 1.5 to 3 times higher than those in LH cheeses (Fig. 7). This was in agreement with the urea-PAGE results and indicates higher levels of intact caseins and a lower level of primary proteolysis in LH cheese (Fig. 4).

On day 1, the summed intensity of β -casein derived peptides in LH cheese was higher than that of α_{S1} -casein derived peptides (Fig. 7A). Taivosalo et al. (2018) reported the same phenomenon in Old Saare cheese and suggested this could be because the hydrolysis of β -casein by plasmin started in milk, whereas the proteolysis of α_{S1} -casein only took place during cheese ripening. Since the plasmin activity in both MCC and LB CMs was higher than that in LH CM (Fig. 3B), this could also explain the higher level of peptide intensity derived from β -casein in the 1 day old MCC and LB cheeses than that in LH cheese (Fig. 7A). As the plasmin activity in HHT LB CM was much lower than that in the other CMs, it was hard to explain the higher intensity of β -casein derived peptides in 1 day old HHT LB cheese than that in LH cheese. The reason for the higher level of α_{S1} -casein proteolysis than that of β -casein in MCC, LB as well as HHT LB cheeses on day 1 of ripening (Fig. 7A) requires further exploration as well.

In all experimental cheeses, the intensities of α_{S1} -casein-derived peptides were higher than those from β -casein at day 120 of ripening (Fig. 7B). In addition, the proportion of α_{S1} - and β -casein derived peptides increased and decreased (Supplementary material Fig. S1), respectively, from day 1 to day 120 of ripening, which might be related to the lower pH in day 120 cheeses (Fig. 1), being more favourable for chymosin action (Li et al., 2020).

In Emmental cheese, plasmin is the primary enzyme driving the hydrolysis, preferentially cleaving β - and α_{S2} -caseins after the Lys

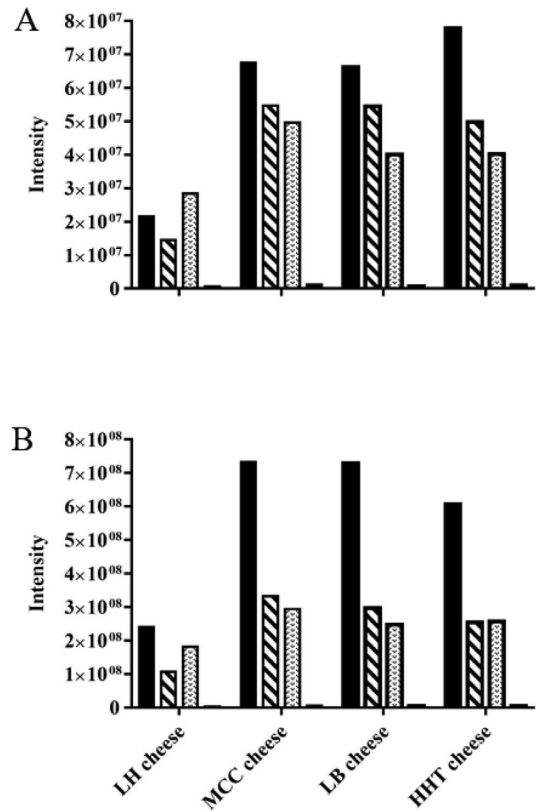


Fig. 7. Summed intensity of peptides derived from α_{S1} - (black), α_{S2} - (stripe), β - (dot), or κ -casein (white) identified in pH4.6 soluble extracts of Emmental cheeses manufactured from low heat skim milk powder (LH cheese), micellar casein concentrate powder (MCC cheese) or β -casein reduced micellar casein concentrate powder that were pasteurised (LB cheese) or high heat treated (HHT LB cheese) at day 1 (A) and day 120 (B) of ripening.

and Arg residues (Upadhyay et al., 2004). Fig. 8 shows the hydrolysis patterns of β -, α_{S2} -, and α_{S1} -caseins for all cheeses at two ripening points. For that, the intensities of all amino acids located at the same positions from all identified peptides were totalled and located on the corresponding casein sequences. Thus, the most intensively hydrolysed regions and regions with a low abundance of peptides can be seen and compared between the cheeses. Most of the β -casein derived peptides were produced from the N-terminal region f1–119 of β -casein, and the region f164–209 was less hydrolysed in all cheeses at both ripening time points (Fig. 8A). Proteolytic patterns of MCC, LB, and HHT LB cheeses were very similar, whereas LH differed considerably from other cheeses at both ripening points (Fig. 8A).

To further evaluate the enzyme cleavage sites, we summed the intensities of all identified peptides starting or ending with the same respective N- or C-terminal amino acids corresponding to the abundance of terminal amino acids adjacent to the cleavage sites, and plotted them on the β -casein sequence (Fig. 9). Hereafter, these are referred to as the intensities of terminal amino acids around the cleavage sites. A higher intensity of terminal amino acids around the cleavage sites was evident in the day 120 cheeses compared with the beginning of ripening (data not shown). The enzyme cleavage sites analysis of β -casein showed no considerable

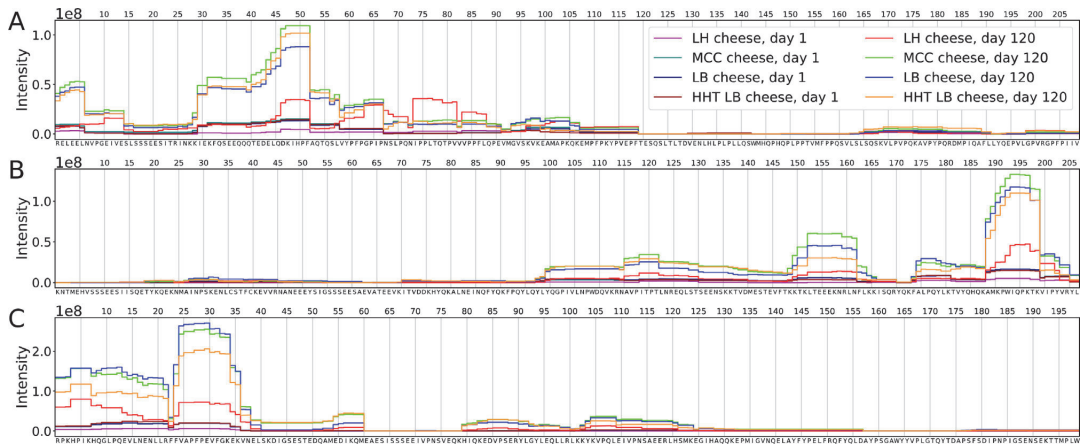


Fig. 8. The regions of β - (A), α_{S2} - (B) and α_{S1} -casein (C) with the identified peptides from Emmental cheeses manufactured from low heat skim milk powder (LH), micellar casein concentrate powder (MCC) or β -casein reduced micellar casein concentrate powder that were pasteurised (LB) or high heat treated (HHT LB) at day 1 and day 120 of ripening (x-axis, the sequences of the caseins; y-axis, summed intensities of the amino acids in all peptides identified in pH4.6 soluble extracts of cheeses by LC–MS).

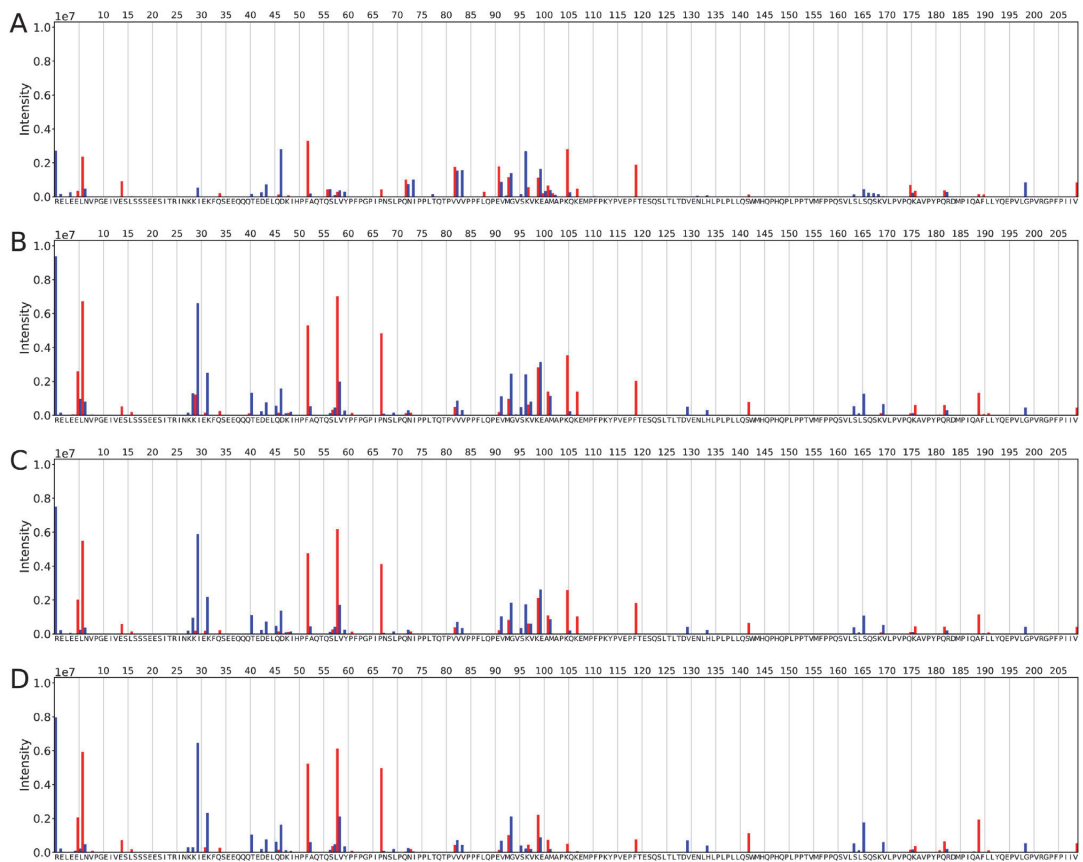


Fig. 9. Enzyme cleavage sites of β -casein identified in LH (A), MCC (B), LB (C), and HHT LB (D) cheeses at 1 day [x-axis, the sequence of β -casein; y-axis, summed intensities of all identified peptides starting or ending with the same respective N- (blue bars) or C-terminal (red bars) amino acids, corresponding to the abundance of terminal amino acids adjacent to the cleavage sites].

difference between MCC and LB cheeses; in general, the intensity of terminal amino acids around the cleavage sites on β -casein was lower in HHT LB cheese. In LH cheese, β -casein was hydrolysed at the same bonds but the intensity of terminal amino acids around the cleavage sites was lower (Fig. 9). The abundant cleavage sites were Leu₆–Asn₇ [possibly hydrolysed by cell envelope proteinases (CEP) of *Lactococcus lactis* (*L. lactis*) (Kunji, Mierau, Hagting, Poolman, & Konings, 1996)], Glu₁₄–Ser₁₅ [(by enzymes of thermophilic starter) (Miclo et al., 2012)], Lys₂₈–Lys₂₉ [(by plasmin) (Fox & McSweeney, 1996)], Lys₂₉–Ile₃₀, Glu₃₁–Lys₃₂, Gln₄₀–Thr₄₁ [(by CEP of lactococci) (Fox & McSweeney, 1996)], Glu₄₄–Lue₄₅, Gln₄₆–Asp₄₇ [(by CEP of lactococci) (Sadat-Mekmene et al., 2011)], Phe₅₂–Ala₅₃ [by CEP of *L. lactis* (Kunji et al., 1996), enzymes of *Streptococcus thermophilus* (*St. thermophilus*) (Miclo et al., 2012), cathepsin D (Larsen, Benfeldt, Rasmussen, & Petersen, 1996)], Gln₅₆–Ser₅₇ [by CEP of lactococci (Miclo et al., 2012; Sadat-Mekmene et al., 2011), enzymes of thermophilic bacteria] Leu₅₈–Val₅₉, Pro₆₂–Asn₆₃, Val₈₂–Val₈₃ [(CEP of lactococci (Miclo et al., 2012; Sadat-Mekmene et al., 2011), enzymes of thermophilic bacteria)] Met₉₃–Gly₉₄ [CEP of lactococci (Kunji et al., 1996; Miclo et al., 2012; Sadat-Mekmene et al., 2011), enzymes of thermophilic bacteria], Lys₉₉–Glu₁₀₀ (Fig. 8).

Looking specifically at plasmin activity based on the peptides identified in pH4.6 soluble extracts, the abundance of plasmin primary cleavage sites Lys₂₈–Lys₂₉, Lys₁₀₅–His₁₀₆/Gln₁₀₆, and Lys₁₀₇–Glu₁₀₈ (Fox & McSweeney, 1996) was considerably lower in LH and HHT LB cheeses compared with MCC and LB cheeses (Fig. 9). Nevertheless, the proteolysis driven by other enzymes towards plasmin-derived peptides seemed to be scarcely affected, which could be seen from the numerous cleavage sites and production of corresponding peptides.

In the case of α ₂-casein, the majority of peptides arose from the C-terminal part, f98–207, of the sequence, while the N-terminal part was less hydrolysed (Fig. 8B). The difference in the proteolytic pattern of LH cheese from other ones is greater than between other cheeses at both 1 and 120 days of ripening (Fig. 8B, red lines). Considering plasmin activity toward α ₂-casein, the intensity of terminal amino acids around the plasmin-preferred cleavage sites Arg₁₁₄–Asn₁₁₅, Lys₁₄₉–Lys₁₅₀, Lys₁₅₀–Thr₁₅₁, Lys₁₈₈–Asp₁₈₉, Lys₁₉₇–Thr₁₉₈ (Le Bars & Gripon, 1989) was lower in HHT LB cheese compared with MCC and LB cheeses, similarly to the case with β -casein (Miclo et al., 2012; Monnet, Ley, & González, 1992; Sadat-Mekmene et al., 2011) (Supplementary material Fig. S2). However, bacterial proteases could also largely contribute to the cleavage of these sites. The plasmin-specific cleavage site Lys₂₁–Gln₂₂, usually not hydrolysed by other proteases, was not seen in the pattern of α ₂-casein cleavages.

Proteolytic patterns of α ₁-casein indicated the same differences in the profiles of cleavage sites between cheeses like in the case of β - and α ₂-casein; more intense hydrolysis, represented by a higher intensity of terminal amino acids around the cleavage sites, was seen in MCC and LB cheeses compared with LH and HHT (Supplementary material Fig. S3). α ₁-casein is known to be hydrolysed by plasmin more slowly than β - and α ₂-casein, while the activity towards κ -casein (Sousa, Ardö, & McSweeney, 2001) is very low. Three major plasmin-specific cleavage sites were found on the sequence of α ₁-casein, Lys₃–His₄, Lys₇₉–His₈₀, and Arg₉₀–Tyr₉₁ (McSweeney, Olson, Fox, Healy, & Højrup, 1993); the intensity of terminal amino acids around the cleavage sites was lower in LH cheese compared with other ones, but no evident differences were seen between MCC, LH, and HHT cheeses. Other plasmin-specific cleavage sites, Lys₄₂–Asp₄₃, Lys₅₈–Gln₅₉, and Arg₁₁₉–Leu₁₂₀ (McSweeney et al., 1993), were not discernible in the pattern of α ₁-casein cleavages (Supplementary material Fig. S3).

3.8. Flowability

The flowability of the cheeses was measured at day 70 and day 120 of ripening, with no significant change observed in all treatments during this time (Table 1, Fig. 10), possibly due to the absence of significant changes in levels of primary proteolysis as demonstrated by levels of pH4.6-SN 100 g⁻¹ cheese. Increased levels of casein breakdown are conducive to increased levels of casein hydration, encouraging the heat-induced detachment of adjacent layers of the *para*-casein network upon heating (Sheehan & Guinee, 2004).

On increasing the heat treatment temperature from 72 to 120 °C, the flowability of HHT LB cheese was significantly lower at both day 70 and day 120 of ripening in comparison with LH, MCC and LB cheeses, possibly due to the higher levels of denatured β -LG in HHT LB cheese (Xia et al., 2022). Rynne, Beresford, Kelly, and Guinee (2004) also reported a negative impact of increased heating temperature applied to cheese milk, from 72 to 87 °C, on the flowability of half-fat Cheddar cheese. It was suggested that the continuity of the protein network in cheese made from HHT milk was enhanced due to the inclusion of denatured β -LG in the cheese matrix through disulphide bonds formed between β -LG and *para*-casein, as well as increased hydrophobic interactions between hydrophobic sites on denatured β -LG in heated cheese (Rynne et al., 2004). Reduction in the level of (intact) β -casein in the cheese through partly depleting β -casein from the cheesemilk (by 9–10%) can enhance its meltability upon heating (O'Mahony et al., 2014); however, the lower amount of β -casein reduction in LB MCC (by 4.25%) achieved in the present study did not result in cheeses with a significantly higher flowability than the LH- and MCC ($P > 0.05$, Fig. 10) cheeses.

3.9. Colour

The colour of Emmental cheeses was monitored over 120 days of maturation, and was expressed as L* (lightness, 0–100, from black to white), a* (–100 to 100, from green to red) and b* values (–100 to 100, from blue to yellow) (Table 1, Fig. 11). For cheeses subjected to different treatments, the levels of L*, a* and b* were significantly affected by ripening time ($P < 0.0001$) (Table 1). The whiteness (L*) ($P < 0.0001$) and redness (a*) ($P < 0.0001$) of experimental cheeses

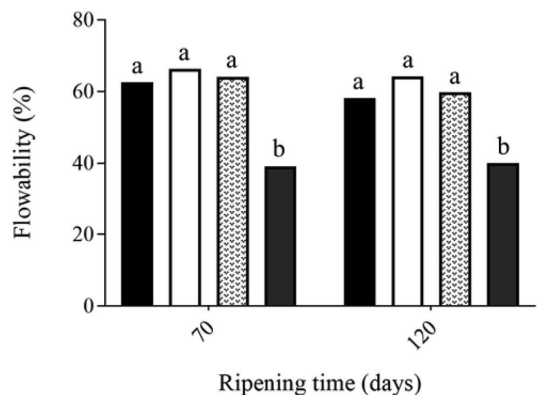


Fig. 10. Flowability of Emmental cheeses manufactured from low heat skim milk powder (black), micellar casein concentrate powder (white) or β -casein reduced micellar casein concentrate powder that were pasteurised (dot) or high heat treated (striped) at day 70 and 120 of maturation.

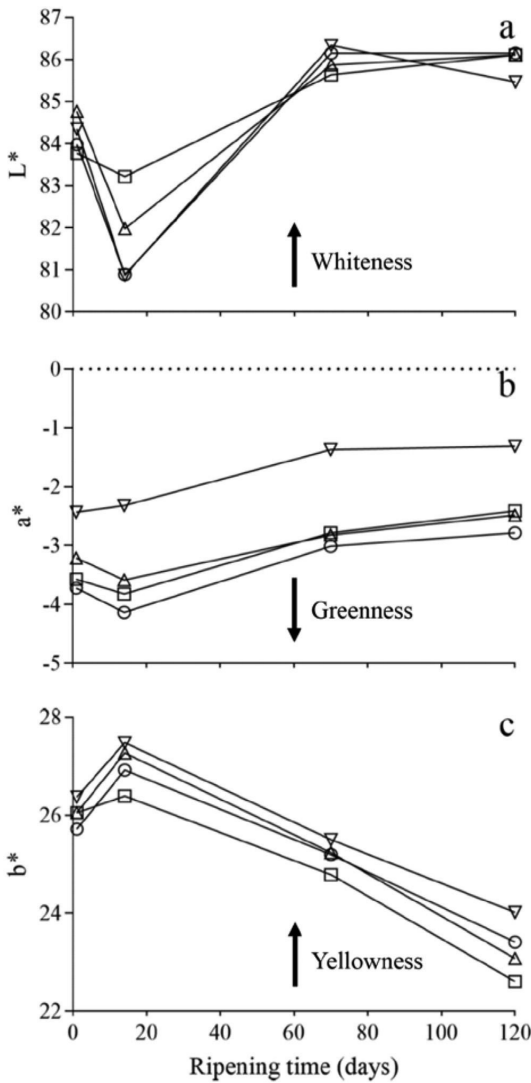


Fig. 11. Colour expressed as L* (a), a* (b) and b* (c) values of Emmental cheeses manufactured from low heat skim milk powder (○), micellar casein concentrate powder (□) or β-casein reduced micellar casein concentrate powder that was pasteurised (△) or high heat treated (▽) at different ripening stage.

increased significantly during warm room period; however, the yellowness (b*) decreased significantly during both the warm room (P < 0.05) and the subsequent cold room ripening (P < 0.01) periods (data not shown).

The hue for each cheeses, on visual observation, was yellow on day 14 and turned more red at day 70 of ripening, in contrast to the results of Rohm and Jaros (1996), who found that the yellowness (b*) of Emmental cheese increased after 10 weeks of maturation, giving the cheese a slightly orange appearance. Sulejmani and Hayaloglu (2016) suggested that the increasing concentration of β-carotene in cheese during storage caused by moisture loss might explain increased yellowness over cheese maturation. Dairy products, including heat-treated milk, powder and cheese manufactured

from heated milk, darken upon intensive heat treatment (temperature and holding time) and storage due to the Maillard Reaction (MR), where lactose and lysine-rich proteins react to generate protein bound-brown compounds (Al-Saadi, Easa, & Deeth, 2013; Morales & Van Boekel, 1998). Due to the absence of water addition in our trials, the residual lactose content in the resultant cheeses might be much higher compared with that in typical Emmental cheese, as was also postulated previously in this study. Reducing sugars like lactose in cheese can inhibit cheese discolouration through the MR during storage (Igoshi, Sato, Kameyama, & Murata, 2017), while high storage temperature can also enhance this reaction and the formation of brown pigments (Rufián-Henares, Guerra-Hernandez, & García-Villanova, 2006).

From day 1 to day 120 of ripening, the redness (a*) of LH-, MCC- and LB cheeses were not significantly different from each other; however, they were all significantly lower than that in HHT LB cheese (Table 1, Fig. 11b). Similar results were also reported by Aydemir and Dervisoglu (2010), who reported that Kulek cheese made from heat-treated milk (75 °C × 5 min) is redder than that

Table 2
Mean volatile compounds peak areas in Emmental cheeses manufactured from whey protein reduced cheese milk of different thermal histories and β-casein contents at day 180.^a

Volatile compounds	LH cheese	MCC cheese	LB cheese	HHT LB cheese
Acids				
Acetic acid	6044958 ^a	7120178 ^a	6576524 ^a	6325249 ^a
Propanoic acid	3012263 ^a	2931604 ^a	1485808 ^a	1499614 ^a
Butanoic acid	4531009 ^a	4325022 ^a	4870161 ^a	4762408 ^a
3-Methylbutanoic acid	307866 ^a	703223 ^a	753178 ^a	428731 ^a
2-Methylbutanoic acid	336575 ^a	974505 ^a	779338 ^a	608154 ^a
Pentanoic acid	57626 ^a	0 ^b	32619 ^a	50209 ^a
Hexanoic acid	3073759 ^a	2152036 ^a	3214334 ^a	2896043 ^a
Octanoic acid	826304 ^a	258498 ^a	655585 ^a	636370 ^a
Decanoic acid	79378 ^a	0 ^b	33720 ^a	57047 ^a
Alcohols				
Ethanol	237154 ^a	372296 ^a	660507 ^a	1193719 ^a
Isopropyl alcohol	58527 ^a	44355 ^a	20227 ^a	29643 ^a
1-Propanol	4418 ^a	13491 ^a	6435 ^a	5655 ^a
2-Butanol	49922 ^a	105283 ^a	7500 ^a	129936 ^a
Phenylethyl alcohol	6512 ^a	4912 ^a	6753 ^a	6861 ^a
Aldehydes				
Benzaldehyde	99779 ^a	11873 ^a	16877 ^a	21383 ^a
Benzeneacetaldehyde	7065 ^a	4149 ^a	4472 ^a	6203 ^a
Esters				
Ethyl propanoate	10611 ^a	64753 ^a	110199 ^a	207135 ^a
Ethyl butanoate	0 ^b	0 ^b	133562 ^a	198964 ^a
Ethyl acetate	13052 ^a	25702 ^a	28688 ^a	0 ^b
Ethyl hexanoate	11414 ^a	16896 ^a	73062 ^a	106291 ^a
Ethyl octanoate	3450 ^a	5821 ^a	28791 ^a	56891 ^a
Ethyl decanoate	0 ^b	0 ^b	4571 ^a	8294 ^a
Ketones				
Acetone	139185 ^a	108964 ^a	111559 ^a	137318 ^a
2,3-Butanedione	1271319 ^a	1265061 ^a	866712 ^a	1239631 ^a
2-Butanone	143795 ^a	100391 ^a	73052 ^a	153294 ^a
2-Pentanone	994334 ^a	352255 ^{a,b}	319853 ^b	654446 ^{a,b}
Acetoin	3550901 ^a	2638519 ^a	2931821 ^a	1521708 ^a
2-Heptanone	1708946 ^a	496740 ^a	766794 ^a	1311057 ^a
2-Nonanone	310660 ^a	75267 ^a	129052 ^a	222331 ^a
2-Undecanone	17716 ^a	10474 ^a	13362 ^a	15054 ^a
Lactones				
γ-Decalactone	3928 ^a	5101 ^a	3760 ^a	3688 ^a
Sulphur compounds				
Dimethyl sulphide	1636 ^a	1132 ^a	1194 ^a	2212 ^a
Dimethyl sulphone	7158 ^a	0 ^b	0 ^b	0 ^b
Terpene				
D-Limonene	1938 ^c	34787 ^a	14424 ^b	14484 ^b

^a Emmental cheeses were manufactured from low heat skim milk powder (LH cheese), micellar casein concentrate powder (MCC cheese) or β-casein reduced micellar casein concentrate powder that were pasteurised (LB cheese) or high heat treated (HHT LB cheese). Results are means of triplicate trials, values within a row not sharing the same superscript letter differ significantly (P < 0.05).

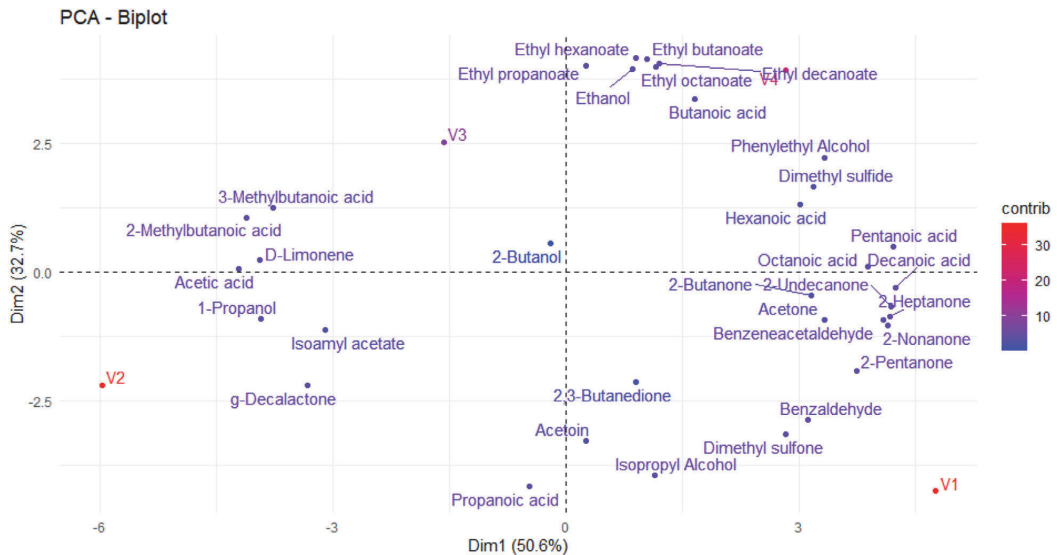


Fig. 12. Principal component analysis (PCA) bi-plot of separation based on volatile profile of Emmental cheese at day 120. Emmental cheeses manufactured from low heat skim milk powder (LH cheese, V1), micellar casein concentrate powder (MCC cheese, V2) or β -casein reduced micellar casein concentrate powder that were pasteurised (LB cheese, V3) or high heat treated (HHT LB cheese, V4). Results are means of triplicate trials.

made from raw milk. The increase in a^* value indicated the formation of brown pigment in HHT LB cheese (Le, Bhandari, Holland, & Deeth, 2011), which may be correlated with the visually increased brownness observed in HHT LB cheese in comparison with those in the other cheeses. No significant difference was observed in colour between unheated, pasteurised and high heat-treated LB milk (data not shown); however, HHT LB cheese was redder than the other cheeses from day 1 of ripening (Fig. 11b). This suggests that low levels of MR might have occurred in HHT LB milk, but the level of browning reaction was so low that the measured milk colour was not changed after being heated at 120 °C for 15 s. Since the brown compounds produced from MR are bound to casein (Morales & Van Boekel, 1998), the colour formed in heat-treated milk might be retained and concentrated in the resultant cheese, giving the fresh HHT LB cheese a brown look.

The Maillard reaction is undesirable for cheese makers due to discolouration and impaired nutritional value (measured by lysine loss) in cheese (Aydemir & Dervisoglu, 2010; Patel, Modi, Patel, & Aparnathi, 2013; Rufián-Henares et al., 2006). It has been shown that processed cheese made from natural cheese with a darker appearance tends to show higher levels of discolouration (Arai et al., 2020). Since lactose is a limiting factor in MR, normalising the lactose content in cheese milk by adding less milk permeate powder (which provided lactose to the cheese milk in this study), or applying HHT to milk before adding milk permeate, might avoid undesirable MR.

Regardless of the treatment applied, the whiteness and yellowness of experimental cheeses were not significantly different (Table 1).

3.10. Volatile compounds

Thirty-four volatile compounds were detected in experimental cheeses with HS-SPME GC-MS, which included 9 acids, 5 alcohols, 2 aldehydes, 6 esters, 8 ketones, 1 lactone, 2 sulphur compounds and 1 terpene (Table 2). LH cheese had higher levels of aldehydes (2

out of 2), ketones (8 out of 8) and sulphur (1 out of 2) compounds than MCC-, LB- and HHT LB cheeses (Table 2), indicating higher levels of amino acid metabolites, lipid oxidation and lactose metabolism in LH cheese.

Principal component analysis was carried out to evaluate the effect of treatments on volatile profile in Emmental cheeses; the total variance was 83.3% with axis PC-1 accounting for 50.6% of the difference and PC-2 axis 32.7% of the difference (Fig. 12). Cheeses made from β -casein reduced MCC, i.e., LB- and HHT LB cheeses, were located on the positive side of PC-2, and the other two cheeses on the negative side (Fig. 12). LB- and HHT LB cheeses were characterised by higher levels of esters (5 out of 6) and ethanol as well as lower levels of isopropyl alcohol and propionic acid compared with LH- and MCC cheeses (Table 2). The high levels of ethanol in cheeses LB- and HHT LB, which suggest high level of lactose degradation, could explain the high level of ethyl ester in these cheeses, since ethanol is often considered as the rate-limiting step for ethyl ester formation (Thierry, Maillard, Richoux, & Lortal, 2006).

It was proposed earlier that the Maillard reaction might have taken place in HHT LB cheese; however, MR-related volatile compounds, such as furfural and furfuryl alcohol (Newton, Fairbanks, Golding, Andrewes, & Gerrard, 2012), were not detected in such cheese.

4. Conclusion

Reduction of β -casein level in MCC by 4.25% did not influence the level of pH, plasmin activity, primary proteolysis (Urea-PAGE), production of water-soluble peptides (LC-MS), texture profile, flowability or colour in Emmental cheeses made therefrom. However, by partly removing whey proteins, which are inhibitors of plasmin and plasminogen activators, from pasteurised skim milk using microfiltration, the plasmin activity in MCC- and LB cheesemilks was significantly higher than in LH cheesemilk. This resulted in lower levels of α _{s1}- and β -casein hydrolysis in the resultant LH

cheese compared with MCC- and LB cheeses (as measured by urea-PAGE and LC-MS) even though there was no significant difference in plasmin activity between these cheeses, further research is required to understand this phenomenon.

Although 53.60% of whey protein was removed from feed milk prior to high heat treatment, the plasmin activity in HHT LB milk and its cheese was largely inhibited, probably due to the formation of disulphide bonds between remaining denatured β -LG and plasmin. As a result, the intensity of plasmin primary cleavage sites on β -casein was significantly lower in HHT LB cheese in comparison with LB cheese (as demonstrated using LC-MS). HHT LB cheese also had a significantly lower level of flowability as well as a higher redness than other treatments, possibly due to the enhanced hydrophobic interaction between denatured whey protein or early-stage Maillard reaction, respectively.

Both β -casein and native whey protein are valuable dairy ingredients; based on our results, these two components can be partly removed from milk before cheesemaking without changing Emmental cheese quality during ripening. Due to its impaired flowability on heating and colour, the Emmental cheese made from high heat-treated whey protein-reduced milk may not be appealing to consumers. Future research should focus on the potential to mitigate this by increasing the casein content on a total protein basis, as well as decreasing the lactose content in whey protein-reduced-milk before high heat treatment; and developing cheeses with higher levels of plasmin activity, primary proteolysis and flowability, together with a less red colour. Overall, as a side stream in the production of both native whey protein and β -casein, β -casein reduced MCC shows good heat stability and can be used to manufacture Emmental cheese of high quality.

Declaration of competing interest

None.

Acknowledgements

Xiaofeng Xia was in receipt of a Teagasc Walsh Fellowship and the study was funded by the Department of Agricultural, Food and Marine under the Food Institutional Research Measure (15/F/683).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.idairyj.2022.105540>.

References

- Aaltonen, T., & Ollikainen, P. (2011). Effect of microfiltration of milk on plasmin activity. *International Dairy Journal*, *21*, 193–197.
- Al-Saadi, J. M., Easa, A. M., & Deeth, H. C. (2013). Effect of lactose on cross-linking of milk proteins during heat treatments. *International Journal of Dairy Technology*, *66*, 1–6.
- Arai, A., Igoshi, A., Inoue, A., Noda, K., Tsutsuura, S., & Murata, M. (2020). Relationship between lactose utilization of lactic acid bacteria and browning of cheese during storage. *Bioscience, Biotechnology, and Biochemistry*, *84*, 1886–1893.
- Ardisson-Korat, A., & Rizvi, S. (2004). Vatless manufacturing of low-moisture part-skim Mozzarella cheese from highly concentrated skim milk microfiltration retentates. *Journal of Dairy Science*, *87*, 3601–3613.
- Ardo, Y., McSweeney, P. L. H., Magboul, A. A. A., Upadhyay, V. K., & Fox, P. F. (2017). Biochemistry of cheese ripening: Proteolysis. In P. L. H. McSweeney, P. F. Fox, P. D. Cotter, & D. W. Everett (Eds.), *Cheese: Chemistry, physics and microbiology* (4th ed., pp. 445–482). San Diego, CA, USA: Academic Press.
- Aydemir, O., & Dervisoglu, M. (2010). The effect of heat treatment and starter culture on colour intensity and sensory properties of Kulek cheese. *International Journal of Dairy Technology*, *63*, 569–574.
- Bacher, T., & Königsfeldt, P. (2000). WPI by microfiltration of skim milk. *Scandinavian Dairy Information*, *3*, 18–20.
- Benfeldt, C., Sørensen, J., Ellegård, K. H., & Petersen, T. E. (1997). Heat treatment of cheese milk: Effect on plasmin activity and proteolysis during cheese ripening. *International Dairy Journal*, *7*, 723–731.
- Fenelon, M. A., O'connor, P., & Guinee, T. P. (2000). The effect of fat content on the microbiology and proteolysis in Cheddar cheese during ripening. *Journal of Dairy Science*, *83*, 2173–2183.
- Fox, P. F., & McSweeney, P. L. H. (1996). Proteolysis in cheese during ripening. *Food Reviews International*, *12*, 457–509.
- Fox, P. F., Uniacke-Lowe, T., McSweeney, P., & O'Mahony, J. (2015). *Dairy chemistry and biochemistry*. New York, NY, USA: Springer.
- France, T. C., Kelly, A. L., Crowley, S. V., & O'Mahony, J. A. (2021). The effects of temperature and transmembrane pressure on protein, calcium and plasmin partitioning during microfiltration of skim milk. *International Dairy Journal*, *114*, Article 104930.
- Fröhlich-Wyder, M.-T., Bisig, W., Guggisberg, D., Jakob, E., Turgay, M., & Wechsler, D. (2017). Cheeses with propionic acid fermentation. In P. L. H. McSweeney, P. F. Fox, P. D. Cotter, & D. W. Everett (Eds.), *Cheese: Chemistry, physics and microbiology* (4th ed., pp. 889–910). San Diego, CA, USA: Academic Press.
- Gagnaire, V., Mollé, D., Herrouin, M., & Léonil, J. (2001). Peptides identified during Emmental cheese ripening: Origin and proteolytic systems involved. *Journal of Agricultural and Food Chemistry*, *49*, 4402–4413.
- Garnot, P., & Mollé, D. (1987). Heat-stability of milk-clotting enzymes in conditions encountered in Swiss cheese making. *Journal of Food Science*, *52*, 75–77.
- Hayes, M. G., Hurley, M. J., Larsen, L. B., Heegaard, C. W., Magboul, A. A., Oliveira, J. C., et al. (2001). Thermal inactivation kinetics of bovine cathepsin D. *Journal of Dairy Research*, *68*, 267–276.
- Heino, A. (2008). Microfiltration of milk I: Cheese milk modification by micro- and ultrafiltration and the effect on Emmental cheese quality. *Milchwissenschaft*, *63*, 279–283.
- Hickey, C. D., Diehl, B. W. K., Nuzzo, M., Millqvist-Feurby, M. A., Wilkinson, M. G., & Sheehan, J. J. (2017). Influence of buttermilk powder or buttermilk addition on phospholipid content, chemical and bio-chemical composition and bacterial viability in Cheddar style-cheese. *Food Research International*, *102*, 748–758.
- Holland, B., Corredig, M., & Alexander, M. (2011). Gelation of casein micelles in β -casein reduced milk prepared using membrane filtration. *Food Research International*, *44*, 667–671.
- Hou, J., Hannon, J. A., McSweeney, P. L. H., Beresford, T. P., & Guinee, T. P. (2012). Effect of curd washing on composition, lactose metabolism, pH, and the growth of non-starter lactic acid bacteria in full-fat Cheddar cheese. *International Dairy Journal*, *25*, 21–28.
- Hou, J., Hannon, J. A., McSweeney, P. L. H., Beresford, T. P., & Guinee, T. P. (2014). Effect of curd washing on cheese proteolysis, texture, volatile compounds, and sensory grading in full fat Cheddar cheese. *International Dairy Journal*, *34*, 190–198.
- Igoshi, A., Sato, Y., Kameyama, K., & Murata, M. (2017). Galactose is the limiting factor for the browning or discoloration of cheese during storage. *Journal of Nutritional Science and Vitaminology*, *63*, 412–418.
- Kamiński, S., Cieślińska, A., & Kostyra, E. (2007). Polymorphism of bovine beta-casein and its potential effect on human health. *Journal of Applied Genetics*, *48*, 189–198.
- Kunji, E. R., Mierau, I., Hagting, A., Poolman, B., & Konings, W. N. (1996). The proteolytic systems of lactic acid bacteria. *Antonie Van Leeuwenhoek*, *70*, 187–221.
- Lamichhane, P., Kelly, A. L., & Sheehan, J. J. (2018a). Effect of milk centrifugation and incorporation of high-heat-treated centrifugate on the composition, texture, and ripening characteristics of Maasdam cheese. *Journal of Dairy Science*, *101*, 5724–5737.
- Lamichhane, P., Pietrzyk, A., Feehily, A. C., Cotter, P. D., Mannion, D. T., Kilcawley, K. N., et al. (2018b). Effect of milk centrifugation and incorporation of high heat-treated centrifugate on the microbial composition and levels of volatile organic compounds of Maasdam cheese. *Journal of Dairy Science*, *101*, 5738–5750.
- Lamichhane, P., Sharma, P., Kennedy, D., Kelly, A. L., & Sheehan, J. J. (2019). Microstructure and fracture properties of semi-hard cheese: Differentiating the effects of primary proteolysis and calcium solubilization. *Food Research International*, *125*, Article 108525.
- Larsen, L. B., Benfeldt, C., Rasmussen, L. K., & Petersen, T. E. (1996). Bovine milk procathepsin D and cathepsin D: Coagulation and milk protein degradation. *Journal of Dairy Research*, *63*, 119–130.
- Larsen, L. B., Wium, H., Benfeldt, C., Heegaard, C. W., Ardö, Y., Qvist, K. B., et al. (2000). Bovine milk procathepsin D: Presence and activity in heated milk and in extracts of rennet-free UF-Feta cheese. *International Dairy Journal*, *10*, 67–73.
- Lawrence, R. C., Creamer, L. K., & Gilles, J. (1987). Texture development during cheese ripening. *Journal of Dairy Science*, *70*, 1748–1760.
- Lawrence, R., Heap, H., & Gilles, J. (1984). A controlled approach to cheese technology. *Journal of Dairy Science*, *67*, 1632–1645.
- Le Bars, D., & Gripon, J. C. (1989). Specificity of plasmin towards bovine α ₂-casein. *Journal of Dairy Research*, *56*, 817–821.
- Le, T. T., Bhandari, B., Holland, J. W., & Deeth, H. C. (2011). Maillard reaction and protein cross-linking in relation to the solubility of milk powders. *Journal of Agricultural and Food Chemistry*, *59*, 12473–12479.
- Li, M., Auty, M. A. E., Crowley, S. V., Kelly, A. L., O'Mahony, J. A., & Brodtkorb, A. (2019). Self-association of bovine β -casein as influenced by calcium chloride, buffer type and temperature. *Food Hydrocolloids*, *88*, 190–198.
- Li, M., Auty, M. A. E., O'Mahony, J. A., Kelly, A. L., & Brodtkorb, A. (2016). Covalent labelling of β -casein and its effect on the microstructure and physico-chemical

- properties of emulsions stabilized by β -casein and whey protein isolate. *Food Hydrocolloids*, 61, 504–513.
- Li, B., Waldron, D. S., Tobin, J. T., Subhir, S., Kelly, A. L., & McSweeney, P. L. H. (2020). Evaluation of production of Cheddar cheese from micellar casein concentrate. *International Dairy Journal*, 107, Article 104711.
- Lopez, C., Camier, B., & Gassi, J.-Y. (2007). Development of the milk fat microstructure during the manufacture and ripening of Emmental cheese observed by confocal laser scanning microscopy. *International Dairy Journal*, 17, 235–247.
- Mane, A., & McSweeney, P. L. (2020). Proteolysis in Irish farmhouse Camembert cheese during ripening. *Journal of Food Biochemistry*, 44, Article 13101.
- McCarthy, C. M., Wilkinson, M. G., & Guinee, T. P. (2017). Effect of coagulant type and level on the properties of half-salt, half-fat Cheddar cheese made with or without adjunct starter: Improving texture and functionality. *International Dairy Journal*, 75, 30–40.
- McCarthy, C. M., Wilkinson, M. G., Kelly, P. M., & Guinee, T. P. (2016). Effect of salt and fat reduction on proteolysis, rheology and cooking properties of Cheddar cheese. *International Dairy Journal*, 56, 74–86.
- McSweeney, P. L. H., Olson, N. F., Fox, P. F., Healy, A., & Højrup, P. (1993). Proteolytic specificity of plasmin on bovine α ₅₁-casein. *Food Biotechnology*, 7, 143–158.
- McSweeney, P. L. H., & Sousa, M. J. (2000). Biochemical pathways for the production of flavour compounds in cheeses during ripening: A review. *Le Lait*, 80, 293–324.
- Miclo, L., Roux, E., Genay, M., Brusseau, E., Poirson, C., Jameh, N., et al. (2012). Variability of hydrolysis of β -, α ₅₁-, and α ₅₂-caseins by 10 strains of *Streptococcus thermophilus* and resulting bioactive peptides. *Journal of Agricultural and Food Chemistry*, 60, 554–565.
- Monnet, V., Ley, J. P., & González, S. (1992). Substrate specificity of the cell envelope-located proteinase of *Lactococcus lactis* subsp. *lactis* NCDO 763. *International Journal of Biochemistry*, 24, 707–718.
- Morales, F., & Van Boekel, M. (1998). A study on advanced Maillard reaction in heated casein/sugar solutions: Colour formation. *International Dairy Journal*, 8, 907–915.
- Newton, A. E., Fairbanks, A. J., Golding, M., Andrewes, P., & Gerrard, J. A. (2012). The role of the Maillard reaction in the formation of flavour compounds in dairy products—not only a deleterious reaction but also a rich source of flavour compounds. *Food & Function*, 3, 1231–1241.
- O'Mahony, J. A., Smith, K. E., & Lucey, J. A. (2014). Purification of beta casein from milk. Patent 11/272, 331 (US 2014/8889208 B2).
- O'Sullivan, D. J., McSweeney, P. L. H., Cotter, P. D., Giblin, L., & Sheehan, J. J. (2016). Compromised *Lactobacillus helveticus* starter activity in the presence of facultative heterofermentative *Lactobacillus casei* DPC6987 results in atypical eye formation in Swiss-type cheese. *Journal of Dairy Science*, 99, 2625–2640.
- Patel, K. N., Modi, R., Patel, H., & Aparnathi, K. (2013). Browning, its chemistry and implications in dairy products: A review. *Indo-American Journal of Agricultural and Veterinary Sciences*, 1, 1–12.
- Raghunath, B., & Hibbard, D. C. (1997). Ultrafiltration of cooled milk. US Patent No. 5, 654, 025.
- Richardson, B. C., & Pearce, K. N. (1981). Determination of plasmin in dairy products. *New Zealand Journal of Dairy Science & Technology*, 16, 209–220.
- Rohm, H., & Jaros, D. (1996). Colour of hard cheese. *Zeitschrift für Lebensmittel-Untersuchung und Forschung*, 203, 241–244.
- Rufián-Henares, J.Á., Guerra-Hernandez, E., & García-Villanova, B. (2006). Colour measurement as indicator for controlling the manufacture and storage of enteral formulas. *Food Control*, 17, 489–493.
- Rynne, N. M., Beresford, T. P., Kelly, A. L., & Guinee, T. P. (2004). Effect of milk pasteurization temperature and in situ whey protein denaturation on the composition, texture and heat-induced functionality of half-fat Cheddar cheese. *International Dairy Journal*, 14, 989–1001.
- Sadat-Mekmene, L., Jardin, J., Corre, C., Mollé, D., Richoux, R., Delage, M. M., et al. (2011). Simultaneous presence of PrtH and PrtH2 proteinases in *Lactobacillus helveticus* strains improves breakdown of the pure α ₅₁-casein. *Applied and Environmental Microbiology*, 77, 179–186.
- Sheehan, J. J., & Guinee, T. P. (2004). Effect of pH and calcium level on the biochemical, textural and functional properties of reduced-fat Mozzarella cheese. *International Dairy Journal*, 14, 161–172.
- Sheehan, J. J., Patel, A., Drake, M., & McSweeney, P. L. H. (2009). Effect of partial or total substitution of bovine for caprine milk on the compositional, volatile, non-volatile and sensory characteristics of semi-hard cheeses. *International Dairy Journal*, 19, 498–509.
- Sheehan, J. J., Wilkinson, M. G., & McSweeney, P. L. H. (2008). Influence of processing and ripening parameters on starter, non-starter and propionic acid bacteria and on the ripening characteristics of semi-hard cheeses. *International Dairy Journal*, 18, 905–917.
- Sousa, M. J., Ardó, Y., & McSweeney, P. L. H. (2001). Advances in the study of proteolysis during cheese ripening. *International Dairy Journal*, 11, 327–345.
- Sulejmani, E., & Hayaloglu, A. A. (2016). Influence of curd heating on proteolysis and volatiles of Kashkaval cheese. *Food Chemistry*, 211, 160–170.
- Taivosalo, A., Kriščiunaitė, T., Seiman, A., Part, N., Stulova, I., & Vilu, R. (2018). Comprehensive analysis of proteolysis during 8 months of ripening of high-cooked Old Saare cheese. *Journal of Dairy Science*, 101, 944–967.
- Thierry, A., Maillard, M.-B., Richoux, R., & Lortal, S. (2006). Ethyl ester formation is enhanced by ethanol addition in mini Swiss cheese with and without added propionibacteria. *Journal of Agricultural and Food Chemistry*, 54, 6819–6824.
- Upadhyay, V. K., Sousa, M. J., Ravn, P., Israelsen, H., Kelly, A. L., & McSweeney, P. L. H. (2004). Use of exogenous streptokinase to accelerate proteolysis in Cheddar cheese during ripening. *Le Lait*, 84, 527–538.
- Winwood, J. (1989). Rennet and rennet substitutes. *International Journal of Dairy Technology*, 42, 1–2.
- Xia, X., Kelly, A. L., Tobin, J. T., Meng, F., Fenelon, M. A., Li, B., et al. (2021). Effect of heat treatment on whey protein-reduced micellar casein concentrate: A study of texture, proteolysis levels and volatile profiles of Cheddar cheeses produced therefrom. *International Dairy Journal*, 129, Article 105280.
- Xia, X., Tobin, J. T., Subhir, S., Fenelon, M. A., Corrigan, B. M., McSweeney, P. L., et al. (2022). Effect of β -casein reduction and high heat treatment of micellar casein concentrate on the rennet coagulation properties, composition and yield of Emmental cheese made therefrom. *International Dairy Journal*, 126, Article 105240.

Appendix 3

Publication III

Arju, G., Berg, H.Y., Lints, T., Nisamedtinov, I. (2022) Methodology for Analysis of Peptide Consumption by Yeast during Fermentation of Enzymatic Protein Hydrolysate Supplemented Synthetic Medium Using UPLC-IMS-HRMS. *Fermentation* 8, 145. doi.org/10.3390/fermentation8040145



Article

Methodology for Analysis of Peptide Consumption by Yeast during Fermentation of Enzymatic Protein Hydrolysate Supplemented Synthetic Medium Using UPLC-IMS-HRMS

Georg Arju ^{1,2,*}, Hidde Yaël Berg ^{1,2} , Taivo Lints ^{1,2} and Ildar Nisamedtinov ^{2,3}

¹ Center of Food and Fermentation Technologies, 12618 Tallinn, Estonia; hidde.berg@tftak.eu (H.Y.B.); taivo.lints@tftak.eu (T.L.)

² Department of Chemistry and Biotechnology, School of Science, Tallinn University of Technology, 12168 Tallinn, Estonia; inisamedtinov@lallemand.com

³ Lallemand Inc., Montreal, QC H1W 2N8, Canada

* Correspondence: georg@tftak.eu; Tel.: +372-53-401-565

Abstract: Several studies have shown the ability of yeast to consume peptides as a nitrogen source in single-peptide containing media. However, a suitable and cost-effective methodology to study the utilization of peptides by yeast and other microorganisms in a complex peptide mixture has yet to be put forward. This article addresses this issue by presenting a screening methodology for tracking the consumption of peptides by yeast during alcoholic fermentation. As a peptide source, the methodology makes use of an in-house prepared peptide-mapped bovine serum albumin (BSA) proteolytic digest, which was applied to a synthetic grape must. The peptide uptake was analyzed using high-throughput ultra-high-pressure liquid chromatography coupled to data-independent acquisition-based ion mobility separation-enabled high-resolution mass spectrometry (UPLC-DIA-IMS-HRMS) analysis. The relative changes of abundance of 123 di- to hexapeptides were monitored and reported during fermentations with three commercial wine strains, demonstrating different uptake kinetics for individual peptides. Using the same peptide-mapped BSA hydrolysate, the applicability of an untargeted workflow was additionally assessed for peptide profiling in unelucidated matrixes. The comparison of the results from peptide mapping and untargeted analysis experiments highlighted the ability of untargeted analysis to consistently identify small molecular weight peptides on the length and amino acid composition. The proposed method, in combination with other analytical techniques, such as gene or protein expression analysis, can be a useful tool for different metabolic studies related to the consumption of complex nitrogen sources by yeast or other microorganisms.

Keywords: liquid chromatography mass spectrometry; ion mobility separation; yeast assimilable peptides; alcoholic fermentation; protein hydrolysate



Citation: Arju, G.; Berg, H.Y.; Lints, T.; Nisamedtinov, I. Methodology for Analysis of Peptide Consumption by Yeast during Fermentation of Enzymatic Protein Hydrolysate Supplemented Synthetic Medium Using UPLC-IMS-HRMS.

Fermentation **2022**, *8*, 145.
<https://doi.org/10.3390/fermentation8040145>

Academic Editors:
Spiros Paramithiotis,
Yorgos Kotseridis and
Maria Dimopoulou

Received: 3 February 2022
Accepted: 24 March 2022
Published: 26 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sufficient assimilable nitrogen content is a prerequisite for successful alcoholic fermentation. *Saccharomyces cerevisiae*, the main microorganism used in the production of alcoholic beverages and bioethanol, is known to utilize ammonia, free amino acids and short peptides as nitrogen sources [1]. While the consumption preferences of *S. cerevisiae* are well studied for ammonia and free amino acids, the role of peptides as a yeast nitrogen source requires further elucidation.

Many feedstocks used in alcoholic fermentation contain a significant proportion of peptides that could be considered as an important nitrogen source for yeast. Peptides thus have a significant technological value. For example, in grape must, about 17% of the total nitrogen can be attributed to oligopeptides, while for beer wort, this number is estimated at around 40% [2,3]. The nitrogen levels in grape must are dependent on the grape variety

and grape cultivation conditions used. Due to these factors, the nitrogen levels in some cases can be deficient [4]. Similarly, high gravity fermentation with substantial content of adjuncts in beer and whisky production is a process that may suffer from nitrogen limitation [2,5].

In any fermentation, nitrogen deficiency can result in premature cessation or even complete arrest of the fermentation process, as well as in poor flavor properties of the product [6–8]. Fermentation nutrients, such as yeast autolysates, rich in free amino acids and peptides, are often used in industry to overcome this problem. Regardless of whether the peptides are present in the fermentation feedstock naturally or are supplemented with nutrients, a suitable screening methodology is required to characterize yeast strains for their ability to take up different peptides. The information obtained from such a screening method would be useful to better anticipate nitrogen deficiency and to avoid related fermentation problems through either specific nitrogen supplementation or the selection of yeast strains with broader peptide consumption capabilities.

S. cerevisiae can internalize peptides of different sizes through multiple membrane peptide transporters [9]. Di- and tripeptide transport is facilitated by Ptr2p, Dal5p and fungal oligopeptide transporters (Fot1-3p) found in oenological yeast [10–14]. However, it is not excluded that FOT transporters can also transport longer peptides [9]. Oligopeptide transporters (Opt1p and Opt2p) have been shown to be involved in tetra- and pentapeptide transport [15–17].

The current knowledge on functionality of different yeast peptide transporters has been mainly obtained using experiments with single synthetic peptides as the sole nitrogen source [10,14–18]. This type of experiment provides accurate information about the ability of yeast strains to utilize the selected peptide as a nitrogen source. However, they do not provide information on the kinetics of peptide utilization throughout different time points during fermentation on complex peptide-rich media, such as wort, grain mash and grape must. Moreover, the characterization of peptide transporters in more complex matrices and peptide utilization capability by different yeasts is challenging due to the lack of suitable cost-efficient screening media with defined and diverse peptide composition. Another challenge is the lack of accurate semi-quantitative analyses for monitoring the uptake of these peptides during fermentation.

Liquid chromatography coupled to mass spectrometry (LC-MS) is currently one of the most promising techniques for short peptide analysis [19–22]. The two biggest analytical challenges are insufficient short peptide identification accuracy for unambiguous identity assignment and throughput limitation caused by data-dependent acquisition (DDA). Unlike longer peptides (pentapeptides and longer) that, due to a higher number of peptide bonds, tend to ionize with a higher charge state ($\geq 2+$), shorter peptides (di- to tetrapeptides) most commonly ionize as singly charged ions. The singly charged peptides produce fewer selective fragments, which leads to reduced identification accuracy [23].

Even though DDA is known to produce higher quality fragmentation spectra due to active quadrupole-based selection of precursor ions prior to fragmentation, it comes at the cost of a reduced acquisition rate that often results in missing data and prolonged analytical gradients [24]. In contrast, data-independent acquisition (DIA) minimizes the likelihood of missing data by rapidly alternating between low and high collision energies and assigning fragments based on the chromatographic elution profile [25]. For fermentation matrices with highly diverse peptide composition, the main drawback of such data acquisition is the generation of complex fragmentation spectra that are often caused by simultaneous fragmentation of closely eluting precursor ions.

A rapid gas-phase separation of ions, also commonly referred to as ion mobility (IM) separation, is often used as an orthogonal method of separation that complements chromatography and mass spectrometry [26]. Hyphenation of travelling wave ion mobility separation (TWIMS) with DIA results in similar drift times for both fragment and precursor ions, allowing alignment by both the chromatographic and ion mobility elution profiles and thus improving the spectral clarity. Moreover, the “trap–release” mode of operation

of the TWIMS results in not only increased peak capacity but also in increased duty cycle of the mass spectrometer. This allows for shorter analytical gradients and the use of the collisional cross section (CCS) as an additional identification qualifier.

This study reports a screening methodology for tracking the consumption of a wide range of peptides by yeast during alcoholic fermentation. The methodology makes use of an in-house prepared peptide-mapped (against a known protein sequence) proteolytic digest of bovine serum albumin (BSA) as a peptide source that was applied to a synthetic grape must. The peptide composition in the hydrolysate as well as the consumption of peptides by yeast in the synthetic grape must were analyzed using high-throughput ultra-high-pressure liquid chromatography coupled with data-independent acquisition-based ion mobility separation-enabled high-resolution mass spectrometry (UPLC-DIA-IMS-HRMS).

To simulate the analysis of an unelucidated matrix, untargeted data analysis (that does not rely on a protein sequence) of the BSA hydrolysate was additionally performed and compared to the results of peptide mapping to assess the identification accuracy of the untargeted analysis.

2. Materials and Methods

An LC-MS methodology was developed for analyzing peptide composition in BSA proteolytic digest (peptide mapping), as well as for monitoring the relative concentration changes of these peptides during fermentation with three commercial wine yeast strains (screening). A modified synthetic grape must with a reduced concentration of free amino acids and ammonia was used as a base fermentation medium and was supplemented with additional nitrogen in the form of peptides from proteolytic digest of BSA. The relative changes of peptide abundances during fermentation were monitored and reported.

Peptide mapping requires a protein sequence to identify peptides. However, when analysing fermentation feedstocks with unknown protein/peptide content, such information is not always available. Therefore, the peptide identification accuracy of an untargeted workflow should be investigated. For this purpose, the peptide identification results obtained by untargeted data analysis against a short peptide database were compared to the results obtained by peptide mapping.

2.1. Chemicals

Hi3 *E.coli* STD (p/n: 186006012) and Leucine-Enkephalin (p/n: WT186006013) were purchased from Waters Corporation (Milford, MA, USA). Ultrapure water (18.2 M Ω × cm) was prepared with MilliQ[®] IQ 7000 equipped with LC-Pak (Merck KGaA, Darmstadt, Germany). Acetonitrile (MeCN; LiChrosolv, hypergrade for LC-MS), formic acid (FA; LC-MS grade), bovine serum albumin (lyophilized powder, ≥96%), fructose, cobalt chloride, boric acid, ammonium molybdate tetrahydrate, ergosterol, oleic acid, all vitamins and cysteine were acquired from Sigma-Aldrich (Darmstadt, Germany).

Histidine, methionine, glutamic acid and arginine were acquired from SERVA Electrophoresis GmbH (Heidelberg, Germany). Alanine, aspartic acid, cysteine, glutamine, glycine, isoleucine, leucine, lysine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine and valine were acquired from Thermo Fisher Scientific (Waltham, MA, USA). Hydrogen chloride (HCl, 36%) was acquired from Labbox Labware (Barcelona, Spain). Liquefied phenol was acquired from Avantor[®] (Wayne, PA, USA).

2.2. Fermentation Media

2.2.1. BSA Hydrolysate

Hydrolysis of BSA (40 g/L dH₂O) was conducted in a 1 L benchtop fermenter (Applikon, Delft, The Netherlands) using the industrial protease COROLASE[®]7089 (AB Enzymes, Darmstadt, Germany), which was selected based on its endoproteolytic activity. This facilitated the greater production of low molecular weight peptides, which were potentially assimilable by yeast. The protease dose rate was 0.5% *w/w* of BSA, and hydrolysis was performed at pH 7 (maintained by titration using 2 M NaOH) for 20 h at 50 °C. Af-

ter completion of the BSA hydrolysis, the hydrolysate was filtered with a Vivaflow[®]200 10 kDa cut-off Hydrosart crossflow cassette (Sartorius, Göttingen, Germany) to remove the protease as well as larger peptides. It was then freeze-dried and stored at $-20\text{ }^{\circ}\text{C}$ until further use.

As the scope of this research was to develop a methodology for following small molecular weight peptide consumption by yeast during alcoholic fermentation, the nitrogen content of peptides smaller than 1 kDa in the fraction was determined from the bound amino acid content and used to calculate the amount of total hydrolysate to be added into the nitrogen deficient synthetic grape must. The sub-1-kDa fraction for amino acid analysis was prepared by rehydrating the BSA hydrolysate (1 g/L dH₂O) and filtering through a Pall[®] Microsep Advance 1K Omega[™] centrifugal filter device (Pall[®] Corporation, Port Washington, NY, USA).

The amounts of free and total amino acids in this fraction were analyzed on a Waters ACQUITY UPLC[®] system (Waters Corporation, Milford, MA, USA) that was coupled to a TUV detector after derivatization using Waters AccQ-Tag chemistry as described by Fiechter and Mayer [27]. For the total amino acids, the samples were first hydrolyzed with 6M HCl + 1% (v/v) phenol for 22 h in a vacuum using an Eldex H/D workstation (Eldex Laboratories Inc., Napa, CA, USA). The amount of bound amino acids was calculated by subtracting the free amino acids from the total amino acid concentration.

2.2.2. Synthetic Grape Must

A synthetic grape must (MS300) as described by Salmon and Barre [28] was prepared with a reduced initial nitrogen content. The free amino acid and NH₃ concentrations were both reduced by 75% to provide 116 mg/L of nitrogen, which was considered insufficient to ferment 210 g/L of fermentable sugars in the synthetic grape must [29,30]. The amino acid and NH₃ concentrations can be found in Table A2. An additional 136 mg/L of nitrogen was then supplemented in the form of peptides from the BSA hydrolysate, resulting in 252 mg/L nitrogen in the synthetic grape must.

2.3. Fermentation

Three commercial *S. cerevisiae* wine yeast strains (Lalvin[™] ICV Opale 2.0, Lalvin[™] Persy and Lalvin[™] QA23), provided by Lallemand Inc. (Montreal, QC, Canada), were used in this study. All strains were obtained as active dry yeast cultures. Inoculum was prepared by rehydrating 1 g of dry culture in 10 mL of sterile 0.9 % NaCl for 15 min at room temperature. The synthetic grape must was then pitched to give an initial yeast concentration of about $5 \cdot 10^6$ cells/mL.

All fermentations were performed at $24\text{ }^{\circ}\text{C}$ in 250 mL Pyrex[™] bottles equipped with a GL45 open top PBT screw cap and PYREX[™] Media Bottle Septum (Corning Inc., Corning, NY, USA). A sampling port was assembled with size 13 Masterflex[™] Norprene[™] Food L/S[™] Precision Pump Tubing (Masterflex SE, Gelsenkirchen, Germany), connected with a metal tubing piercing through the septum. A gas outlet was installed to prevent overpressure by piercing the septum with a Sterican[®] Ø 0.8 × 40 mm single-use hypodermic needle (B. Braun, Melsungen, Germany) attached to a Millex-FG 0.2 µm hydrophobic PTFE filter (Merck KGaA, Darmstadt, Germany).

The initial volume of synthetic grape must was 200 mL. Samples (3 mL) were collected at 0, 24, 48, 72 and 168 h. For each strain, a control fermentation was performed in parallel without the addition of BSA hydrolysate.

2.4. Sample Preparation

Prior to peptide analysis, samples (200 µL) were mixed with MeCN (200 µL) in an Eppendorf[®] tube (Eppendorf AG, Hamburg, Germany) to precipitate proteins, vortexed for 30 s and centrifuged at $11,200 \times g$ for 15 min. The supernatant (100 µL) and MilliQ[®] water (900 µL) were then transferred into a new Eppendorf[®] tube and vortexed for 30 s.

The obtained mixture (100 μL) was transferred into a low-volume insert and spiked with 20 μL of 1 pmol/ μL of Hi3 *E.coli* STD.

Before and after injection of the batch of samples, a blank (MilliQ[®] water) and a blank (100 μL MilliQ[®] water) spiked with 20 μL of 1 pmol/ μL of Hi3 *E.coli* STD were injected to determine the baseline and system variability. A pooled sample (100 μL) spiked with 20 μL of 1 pmol/ μL of Hi3 *E.coli* STD was also injected three times at the beginning of the experiment and one time after each strain to determine the intra-batch variability.

2.5. Liquid Chromatography Mass Spectrometry

Samples were analyzed using Waters I-Class Plus (SM-FL) UPLC[®] system (Waters Corporation, Milford, MA, USA) coupled with a Waters Vion IMS-QToF Mass Spectrometer equipped with LockSpray II Exact Mass source enclosure and ESI Tool-free Mk3 UPC² probe (50 $\mu\text{m} \times 750 \text{ mm}$, p/n: 700011376) directly connected to the column outlet. Nitrogen was used as collision gas. The instrument was controlled by Waters UNIFI 1.9.4 (3.1.0, Waters Corporation, Milford, MA, USA).

Mobile phases (MP) were as follows: (A) MilliQ[®] + 0.1% formic acid and (B) MeCN + 0.1% formic acid. The weak needle wash was 90:10 (MilliQ[®]:MeCN), the strong needle wash was 10:90 (MilliQ[®]:MeCN), and the seal wash was 50:50 (MilliQ[®]:MeCN). Injection volume was 2 μL . Samples were analyzed using Acquity UPLC[®] HSS T3 Column (1.8 μm , 1 \times 150 mm, Waters Corporation, Milford, MA, USA) kept at 40 °C. The initial flow rate was 0.1 mL/min. The gradient was as follows: a 0–0.5 min hold at 1% B, 0.5–10.5 min linear gradient 1–50% B, 10.5–12.5 min linear gradient, 50–99% B accompanied by a linear flow rate increase of 0.1–0.2 mL/min, 12.5–14.5 min hold at 99% B, 14.5–15 min linear gradient 99–1% B accompanied by a linear flow rate decrease of 0.2–0.1 mL/min, and 15–18.5 min hold at 1% B.

The instrument was operated in positive polarity, sensitivity mode (33,000 FWHM at 556.2766 m/z) and labile ion mobility tune. The analysis type was set as Peptide Map (IMS) and the experiment type was set to MSe. Data was acquired in HDMSe mode with a scan time of 0.2 s. The following manual quadrupole profile was used: mass 150/300/600 (m/z), dwell time 50/30 (%scan time), ramp time 10/10 (%scan time). The recorded mass range was from 50 to 1000 m/z for both low and high energy spectra. The collision energy was ramped from 10 to 63 V.

The cone voltage was set to 30 V, capillary voltage was set to 0.5 kV and source offset was set to 50 V. Source temperature was set to 120 °C and desolvation temperature set to 500 °C. Cone gas flow rate was set to 50 L/h and desolvation gas flow rate was set to 700 L/h. Leucine-enkephalin (50 pg/ μL at 15 $\mu\text{L}/\text{min}$) was used as LockMass for mass axis correction and was acquired before and after each acquisition as well as every 5 min during the acquisition.

2.6. Data Processing

For peptide identification in the BSA hydrolysate, peptide mapping was conducted using UNIFI Large Molecule Package (LMP) by searching the peptides against the sequence of bovine serum albumin (ALBU_BOVIN; P02769). Low energy intensity threshold of 2000 counts and high energy intensity threshold of 200 counts were used. The retention time range was set between 1 and 12 min. Peptide search was set to non-specific digest with no missed cleavages allowed, and the minimum sequence length was set to 2.

The following filters were applied: mass error between -3 and 3 ppm, matched first gen primary ions greater than or equal to 1, fragment label not containing in-source fragments, loss of H₂O and loss of NH₃. For relative quantification (screening of peptide uptake by yeast) and untargeted analysis of short peptides the data was exported using UNIFI Export Package (UEP) to Progenesis Q1 (PQ1) software (Nonlinear Dynamics, Newcastle, UK). During the import, masses were lock mass corrected with 556.2766 m/z , corresponding to singly charged Leucine-enkephalin. Data was subjected to normalization of compound

abundances based on a set of reference housekeeping peptides (spiked Hi3 *E.coli* STD). Automatic sensitivity at level 2 was used.

Automatic retention time alignment was conducted, and the retention time limits were set between 1 and 12 min. Fragment sensitivity of 1% of the base peak was used. The following adducts were used for peptide mapping: M+H, M+2H and M+3H. ChemSpider data source: peptides (PQI_CS_Peptides), containing in-silico database of short peptides (2-4 AA), were used for de novo identification. The precursor ion tolerance was set to 3 ppm and fragment ion tolerance was set to 5 ppm. One or more fragment ions per peptide were required for ion matching.

2.7. Data Analysis

Peptide peaks identified by peptide mapping and untargeted analysis of short peptides were aligned using observed mass to charge ratio, charge, observed retention time (min) as well as collision cross section (\AA^2). Peptides without a distinct consumption trend were filtered out using PQI software. Peptides with up to six amino acids in length and an initial abundance higher than 0.1% of the summed abundance of all peptides were considered and reported for each fermentation experiment.

The aligned data was further used to construct a data matrix for the aforementioned data processing methods. The comparison between two identification methods was performed with the help of in-house data analysis and visualization scripts written in the PythonTM programming language (Python Software Foundation, Wilmington, DE, USA). The identification accuracy and reproducibility assessment between peptide mapping and untargeted analysis was performed based on four categories: absolute amino acid sequence match (Absolute), amino acid sequence match while leucine (L) and isoleucine (I) were not differentiated and annotated as "J" (J Absolute), and amino acid composition match (J Composition) and peptide length match (Length).

3. Results

3.1. BSA Hydrolysate Composition

The total nitrogen in the BSA hydrolysate was determined in three fractions: free amino acids, peptides with molecular weight (MW) smaller than 1 kDa and peptides with MW higher than 1 kDa. Most of the nitrogen (69%; 81.4 mg/g) came from peptides with MW below 1 kDa. Endoproteolytic activity of COROLASE[®] 7089 was apparent as only 7% of total nitrogen could be assigned to free amino acids (8.1 mg/g). The remaining 24% of nitrogen (29 mg/g) was present in the higher than 1 kDa fraction. The degree of hydrolysis was calculated, as described by Adler-Nissen [31] and was 40%.

3.2. LC-MS Method Reproducibility

We observed that the default settings of the ion optics and quadrupole of Vion IMS-Qtof resulted in a lower response of the lower molecular weight peptides. Combination of the labile ion mobility tune and a manual short peptide specific quadrupole transmission profile resulted in an increased response of di- and tripeptides. The system performance was found to be highly reproducible as illustrated by Figure 1 (overlay of three technical replicates). The standard deviation of the peptide retention times was below 0.1 min, and the PQI alignment score was over 95% for the six QC samples and over 92% for all 36 samples.

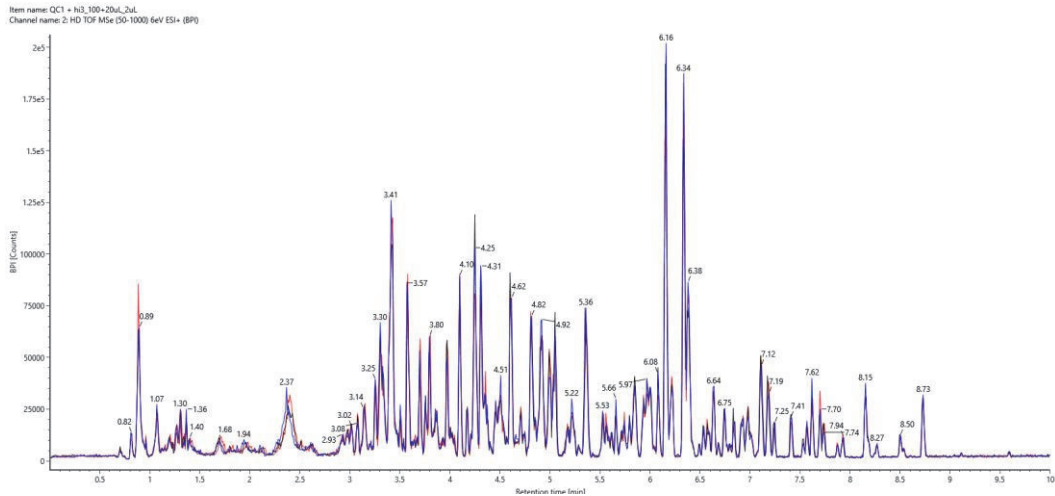


Figure 1. Overlay of triplicate injection of Hi3 *E.coli* STD spiked pooled-sample base peak intensity (BPI) chromatograms.

Moreover, the starting points of all three fermentations display an average relative standard deviation of 5.27% for the intensities of the peptides of interest. The average peak width at 10% of peak height was 8 s, thus, resulting in not only enough data points for reliable qualitative and quantitative integration but also in a sufficient number of data points for accurate tracking of the peak's lift-off, apex and touch-down points for improved high-to-low energy spectra association.

3.3. Fermentation and Peptide Uptake Kinetics

The maximum CO₂ production rate was reached at about 48 h in case of all strains. However, the fermentations supplemented with the BSA hydrolysate showed increased sugar consumption, higher CO₂ production and faster biomass accumulation compared to the corresponding controls without the hydrolysate (Figure 2). At the final experimental point (168 h), up to 95% of sugars were consumed in fermentations with added BSA hydrolysate, whereas up to 83% of sugars were consumed without hydrolysate addition (Table A1).

In total, 123 peptide candidates complying with data analysis filtration criteria (19 dipeptides, 41 tripeptides, 31 tetrapeptides, 19 pentapeptides and 13 hexapeptides) were monitored throughout the fermentation experiments. The relative changes of individual peptide candidate intensities during the fermentation of the synthetic grape must are shown in Figures 3–7 and Tables A4–A8.

All three strains demonstrated similar uptake trends of di-, tri- and tetrapeptides. For all strains, di- and tripeptides were consumed during the first 48 h of fermentation, simultaneously with the free amino acids (Figure 8). The uptake rate of di- and tripeptides was higher by Lalvin Persy™ compared to Lalvin ICV Opale 2.0™ and Lalvin QA23™, as several peptides were consumed by this strain already at 24 h.

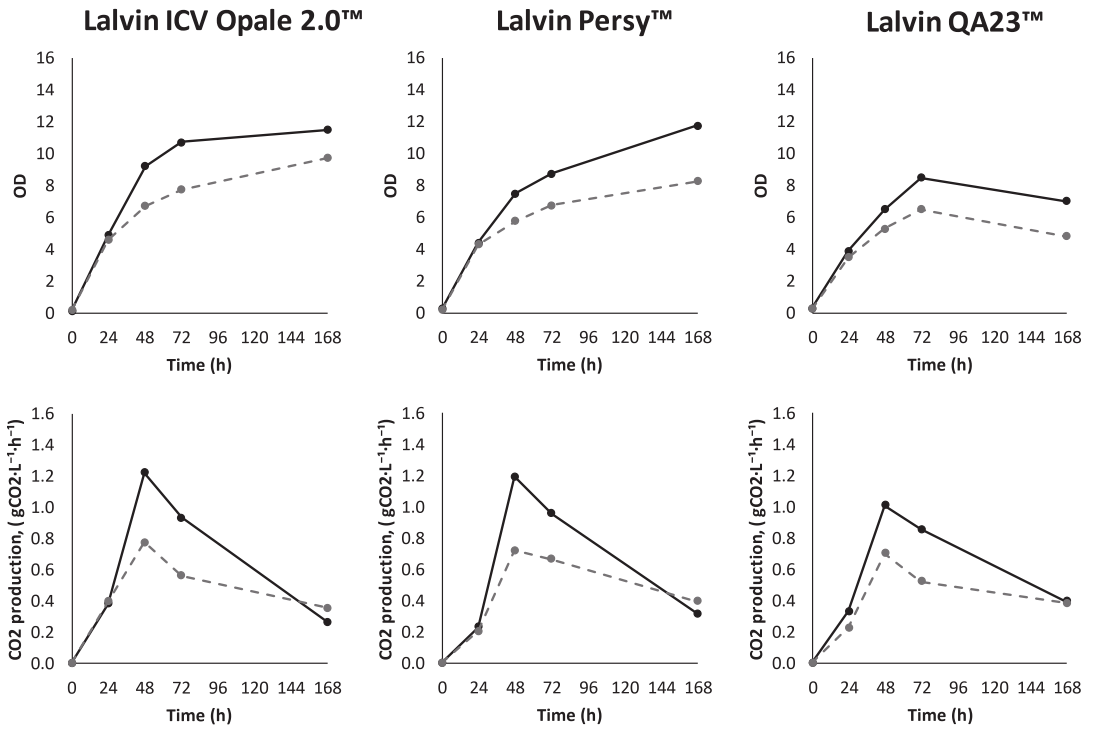


Figure 2. The optical density (OD) and CO₂ production rate in fermentations with Lalvin ICV Opale 2.0™, Lalvin Persy™ and Lalvin QA23™ of the modified synthetic must with (solid line) and without (dashed line) added BSA hydrolysate as the peptide source.

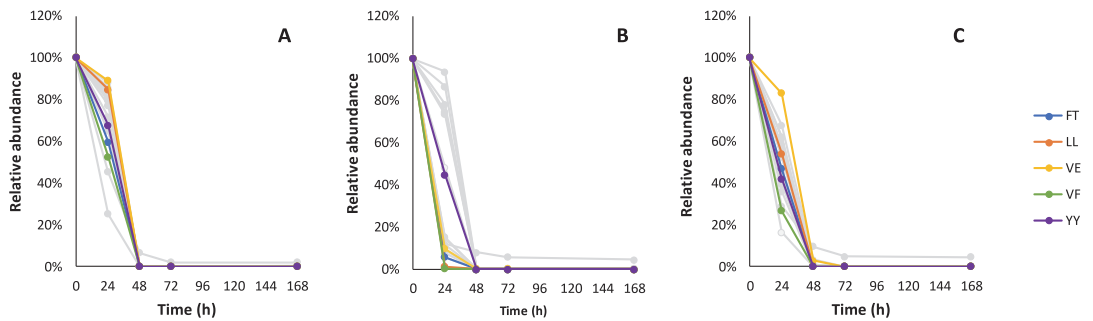


Figure 3. The relative consumption trends of 19 dipeptides in fermentation with Lalvin ICV Opale 2.0™ (A), Lalvin Persy™ (B) and Lalvin QA23™ (C) of the modified synthetic must with added BSA hydrolysate as the peptide source. The five dipeptide candidates with the highest abundance are shown in colour. Other dipeptide candidate sequences and their observed relative consumption trends are shown in Table A4.

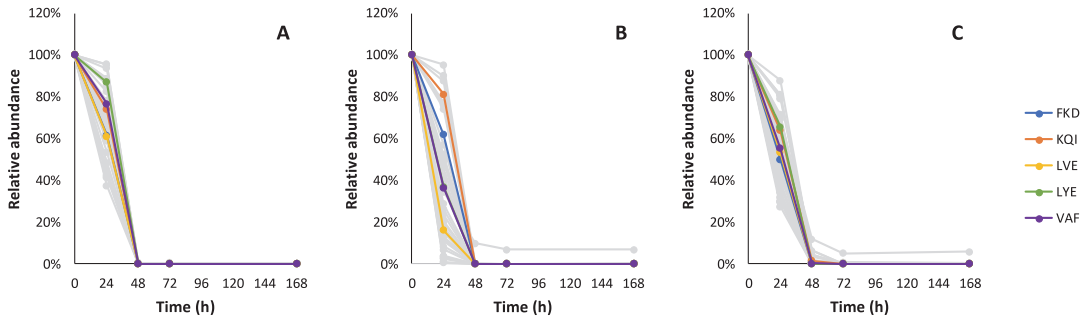


Figure 4. The relative consumption trends of 41 tripeptides in fermentation with Lalvin ICV Opale 2.0™ (A), Lalvin Persy™ (B) and Lalvin QA23™ (C) of the modified synthetic must with added BSA hydrolysate as the peptide source. The five tripeptide candidates with the highest abundance are shown in colour. Other tripeptide candidate sequences and their observed relative consumption trends are shown in Table A5.

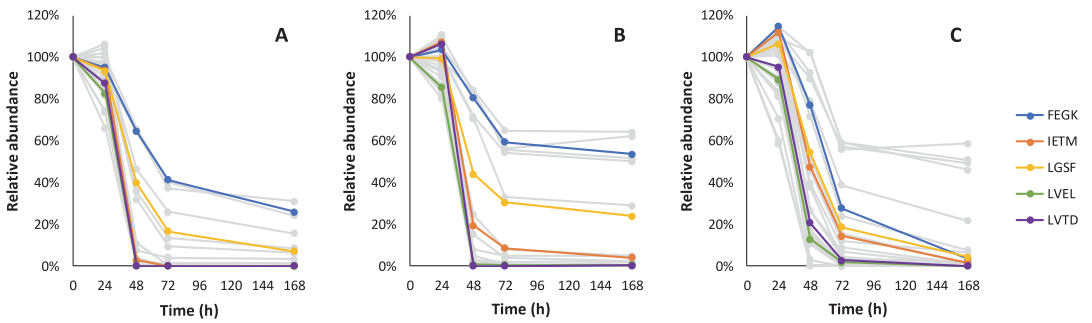


Figure 5. The relative consumption trends of 31 tetrapeptides in fermentation with Lalvin ICV Opale 2.0™ (A), Lalvin Persy™ (B) and Lalvin QA23™ (C) of the modified synthetic must with added BSA hydrolysate as the peptide source. The five tetrapeptide candidates with the highest abundance are shown in colour. Other tetrapeptide candidate sequences and their observed relative consumption trends are shown in Table A6.

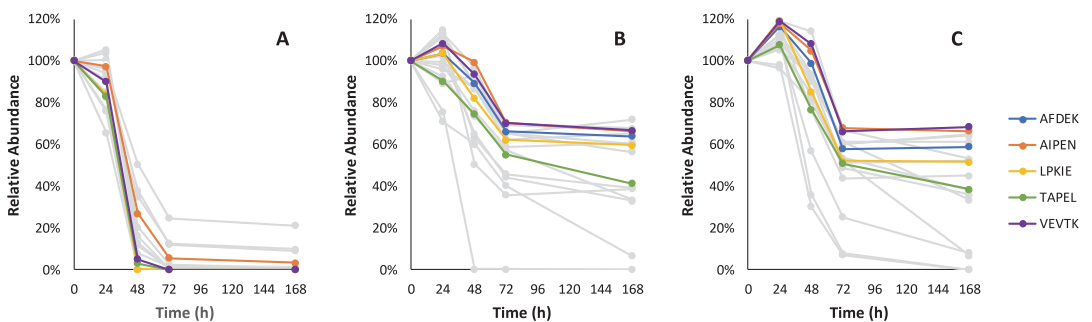


Figure 6. The relative consumption trends of 19 pentapeptides in fermentation with Lalvin ICV Opale 2.0™ (A), Lalvin Persy™ (B) and Lalvin QA23™ (C) of the modified synthetic must with added BSA hydrolysate as the peptide source. The five pentapeptide candidates with the highest abundance are shown in colour. Other pentapeptide candidate sequences and their observed relative consumption trends are shown in Table A7.

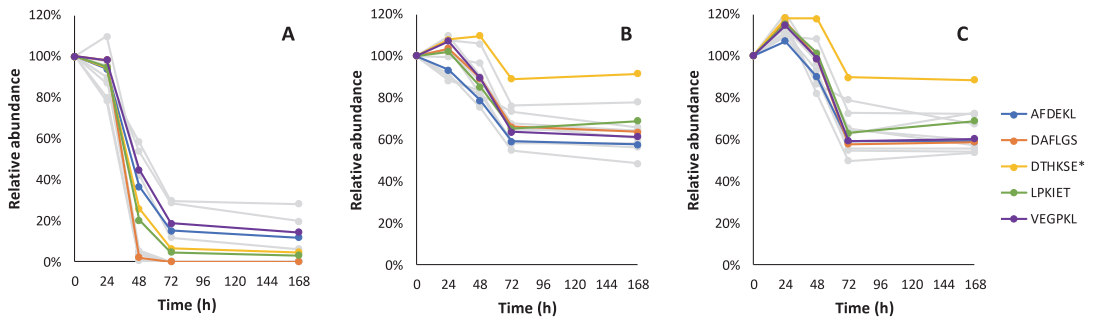


Figure 7. The relative consumption trends of 13 hexapeptides in fermentation with Lalvin ICV Opale 2.0™ (A), Lalvin Persy™ (B) and Lalvin QA23™ (C) of the modified synthetic must with added BSA hydrolysate as the peptide source. The five hexapeptide candidates with the highest abundance are shown in colour. Other hexapeptide candidate sequences and their observed relative consumption trends are shown in Table A8. * DTHKSE elutes within the void volume.

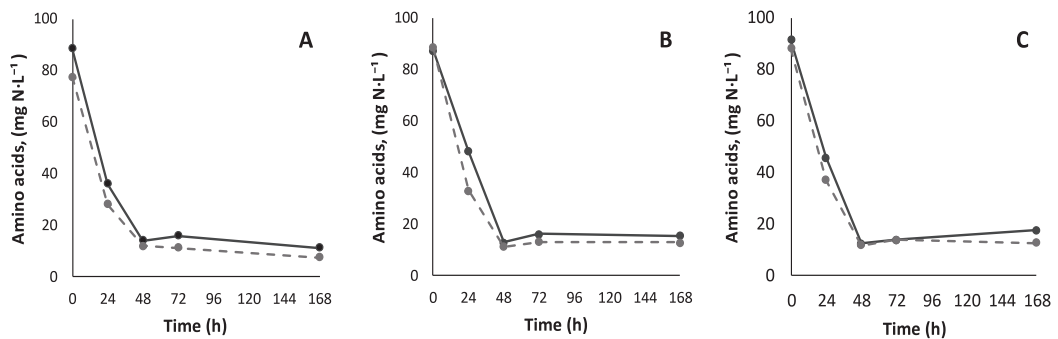


Figure 8. Free amino acid consumption trends in fermentations with Lalvin ICV Opale 2.0™ (A), Lalvin Persy™ (B) and Lalvin QA23™ (C) of the modified synthetic must with (solid line) and without (dashed line) added BSA hydrolysate as the peptide source.

Unlike di- and tripeptides, not all tetra- to hexapeptides were fully depleted (signal to noise ratio < 3:1), and their uptake ceased when the growth of cells entered the stationary phase (Lalvin ICV Opale 2.0™ and Lalvin QA23™) or slowed down remarkably (Lalvin Persy™) after 72 h. The differences in peptide uptake between strains were most notable for penta- and hexapeptides. Thus, the relative abundances of pentapeptides at 168 h decreased on average by 49% for Lalvin Persy™ and 58% for Lalvin QA23™, and hexapeptides decreased by 35% for both strains. The respective values for Lalvin ICV Opale 2.0™ were 98% for pentapeptides and 93% for hexapeptides.

3.4. Untargeted Peptide Analysis

The results of the untargeted data analysis (using the PQI_CS_Peptides database) were compared against the results obtained by peptide mapping to assess the identification accuracy of the untargeted workflow.

Figure 9 highlights the ability of the untargeted approach to identify all di- to tetrapeptides in the BSA hydrolysate in comparison to the peptide mapping. A trend emerged highlighting how the number of peptides assigned identically by both search engines depends on the identification specificity. Only 27.4% of the peptides were matched on the level of absolute amino acid sequence and 37.5% on the level of “J Absolute” where leucine (L) and isoleucine (I) were not differentiated and annotated as “J”. On the level of “J

Composition”, where the order of the amino acids in the sequence was ignored and only the amino acid compositions of the peptide was considered, 71.4% of peptides were matched. Lastly, on the level of “Length”, where only the length of the peptide was considered, 90.5% of peptides were matched.

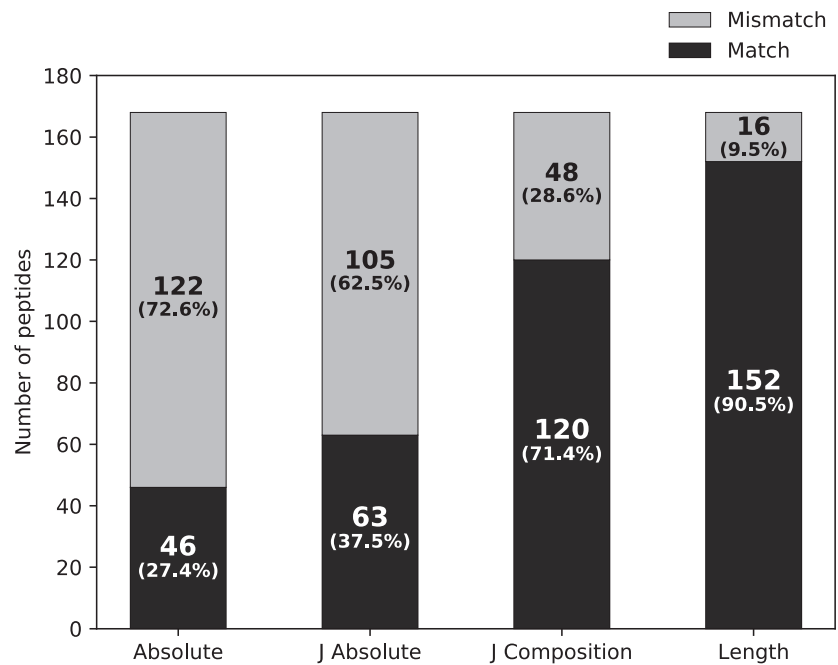


Figure 9. The dependence of the peptide matching specificity criteria on the peptide identification assignment discrepancy: absolute amino acid sequence match (Absolute), amino acid sequence match, where leucine (L) and isoleucine (I) were not differentiated and are annotated as “J” (J Absolute), amino acid composition match (J Composition) and peptide length match (Length).

Figure 10 displays the identification discrepancies between the two methods regarding the peptide length assignment for all di- to tetrapeptides. For dipeptides, the match was 100%: all 29 dipeptides were assigned by the untargeted method and the peptide mapping. The tripeptide length assignment displayed minimal discrepancy: out of 62 tripeptides identified by the untargeted method, only one (1.6%) was assigned as a tetrapeptide by the peptide mapping.

A larger discrepancy was observed for tetrapeptides: out of 77 tetrapeptides assigned by the untargeted method, three (3.9%) were assigned as tripeptides by peptide mapping, and 12 (15.6%) were assigned as pentapeptides by peptide mapping. The larger mismatch for tetrapeptides can be partly explained by the fact that the PQI_CS database contains sequences only up to tetrapeptides. While the peptide mapping approach resulted in a higher number of total identifications (not limited to di- to hexapeptides), a consistently low number (<10%) of unique (picked only by PQI) peptides were detected by the untargeted approach.

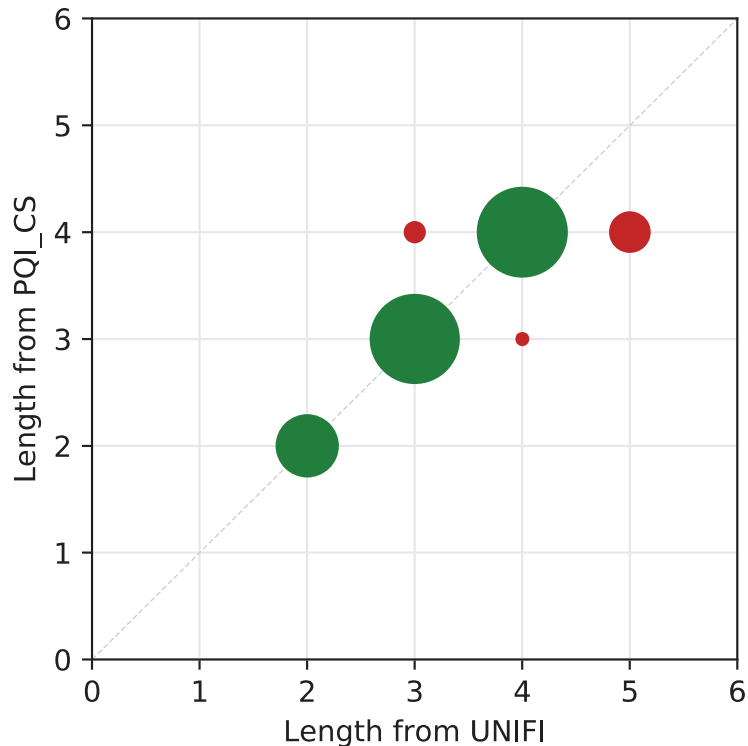


Figure 10. Length assignment discrepancies by the untargeted method in relation to peptide mapping. Green circles represent peptide identification overlap between the two identification methods, whereas red circles represent mismatches between peptide identifications. Circle areas are scaled to represent the ratios between matched and mismatched peptides.

4. Discussion

In this study, we explored the application of an in-house produced BSA protein hydrolysate as a cost-efficient model mixture of peptides for yeast studies. The results indicate the potential of this method for screening yeast strains for their ability to consume different peptides during alcoholic fermentation, for studying peptide transporters, as well as for evaluating the effect of peptides as a nitrogen source on fermentation kinetics. The hydrolysis of BSA with COROLASE® 7089 resulted in 123 di- to hexapeptides, followed by their consumption by different wine yeast strains during fermentation experiments.

The results highlight differences in penta- and hexapeptide consumption trends, indicating the need for further research. Until now, the studies of peptide consumption by *S. cerevisiae* have focused on di- to pentapeptides, mainly in single peptide-based systems and small volumes, mostly due to the higher cost of synthetic peptides. An example of such a system is the Biolog (Hayward, CA, USA) Phenotype MicroArrays (PM) used by Becerra-Rodríguez et al. [10] to characterize FOT transporter knock-out strains for their ability to utilize 270 dipeptides and 14 tripeptides as nitrogen sources using a 96-well plate system. The ability of yeast to consume an individual peptide was assessed based on the growth in the medium containing a single peptide as the nitrogen source.

Despite being a high-throughput method for screening, this approach, due to its set-up (one single peptide per experiment) and very small volume, imposes several limitations for more in-depth studies of peptide consumption. These limitations exclude the possibility to determine fermentation and peptide uptake kinetics as well as peptide transporter

expression analysis in the continuously changing environment of fermentation processes. The application of in-house-prepared protein hydrolysates allows for a more versatile and cost-effective approach to studying yeasts and other microorganisms for their ability to utilize various peptides in environments with more complex peptide compositions, which resemble more natural fermentation environments.

For this study, COROLASE[®] 7089 was chosen due to its high and broad endoproteolytic activity, which resulted in a higher number of di- to hexapeptides, which is important when peptide consumption by yeast is studied. However, the endoprotease and the protein source could be easily substituted with alternatives producing hydrolysates with different peptide compositions, allowing the tailor-made production of peptide mixtures based on various (micro-)organism-specific requirements [32–35].

For peptide composition analysis, it is typically recommended to hydrolyze a single protein with a known amino acid sequence. The identification of peptides against a known protein sequence allows for an additional degree of confidence in the identification accuracy. However, a protein sequence is not always available, and thus de novo identification should be used. The challenge of characterizing peptide mixtures using industrial proteases is often accompanied by unelucidated cleavage specificity of the proteases. This is also evident when peptide composition in natural fermentation feedstocks is to be analyzed, emphasizing the importance of the identification accuracy.

The results generated in this work by the untargeted analysis suggest that the method does not allow for small peptide unambiguous absolute sequence identification, which might be crucial for studying peptide transporters specificities. However, the method allows for peptide identification at the level of peptide length (“Length”, 90.5%) and amino acid composition, where isoleucine or leucine are not differentiated (“J Composition”, 71.4%), suggesting its potential for characterization of the composition of unelucidated natural peptide-containing fermentation matrices, such as grain mash, wort or grape must.

Even though the TWIMS used in this study increased the peak capacity and allowed for a shorter gradient time, the IMS resolution of the system was insufficient to reliably distinguish between isobaric peptides of the same length and different sequences. However, the collision cross section values were found to provide an extra qualifier when aligning features between two peak picking methods.

Quantitative representation of the content and consumption of all peptides remains a topic of further research. Due to the variation in the ionization efficiency of different peptides [36], consumption trends could only be used for tracking individual peptides across different samples, setting a limitation regarding the analysis of specific consumption rates of individual peptides by yeast in complex environments. The absolute quantitation of peptides remains reliant on labelled isotope dilution or tandem-mass-tag labelling [37–39].

5. Conclusions

We developed a methodology for characterization of the peptide composition in enzymatically produced hydrolysates of bovine serum albumin and implemented high-throughput monitoring of the peptide consumption trends of yeast during the alcoholic fermentation of a BSA-hydrolysate-supplemented synthetic grape must.

A correlation between the identification results of peptide mapping and untargeted analysis of di- to tetrapeptides (limited by the PQL_CS_Peptides database) suggests the potential of the untargeted analysis for following the peptide composition in unelucidated matrices, such as grain mash, wort or grape must. The method can additionally be used for profiling various enzymatically produced protein hydrolysates. However, consumption trends can only be used for tracking individual peptides across different samples. For the absolute quantitation of short peptides, the applications of tandem-mass-tag labelling must be further studied. The proposed methods, in combination with other analytical techniques, such as gene or protein expression analysis, can be a useful tool for different metabolic studies related to the consumption of complex nitrogen sources by yeast or other microorganisms.

Author Contributions: Conceptualization, G.A., H.Y.B. and I.N.; methodology, G.A. and H.Y.B.; software, G.A.; validation, G.A., H.Y.B. and T.L.; formal analysis, G.A. and H.Y.B.; investigation, G.A. and H.Y.B.; resources, I.N.; data curation, G.A. and H.Y.B.; writing—original draft preparation, G.A. and H.Y.B.; writing—review and editing, T.L. and I.N.; visualization, H.Y.B. and T.L.; supervision, I.N.; project administration, G.A.; funding acquisition, G.A., H.Y.B. and I.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by “TUT Institutional Development Program for 2016–2022” Graduate School in Biomedicine and Biotechnology receiving funding from the European Regional Development Fund under program ASTRA 2014–2020.4.01.16-0032 in Estonia and the Estonian Research Council via project RESTA13.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: The authors would like to acknowledge Signe Saarmets for performing the fermentations.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Sugars and ethanol were measured using the Waters Alliance 2695 system (Waters Corporation, Milford, MA, USA). The flow rate used was 0.6 mL/min. The mobile phase was isocratic for 25 min on 100% *v/v* 0.5 mmol H₂SO₄ in MilliQ water. The column used for analyses was Bio-Rad HPX-87H column (Bio-Rad Laboratories Inc, Hercules, CA, USA) with the dimensions 7.8 × 300 mm and 9 μm particle size, coupled to a Micro-Guard Cation H Cartridge pre-column (Bio-Rad Laboratories Inc, Hercules, CA, USA). The analytes were detected by the refractive index detector.

Table A1. Glucose, fructose and ethanol measurements at the start and end of fermentation with Lalvin ICV Opale 2.0TM, Lalvin PersyTM and Lalvin QA23TM of the modified synthetic must. ‘Control’ is without added peptides. ‘BSA’ is with added peptides.

Strain		Glucose Consumed (g·L ⁻¹)	Fructose Consumed (g·L ⁻¹)	Ethanol Produced (g·L ⁻¹)
Lalvin ICV	Control	97.3	72.4	77.5
Opale 2.0 TM	BSA	102.8	97.3	95.4
Lalvin Persy TM	Control	97.4	76.8	81.4
	BSA	104.4	101.2	95.7
Lalvin QA23 TM	Control	90.1	66.9	75.5
	BSA	100.0	92.1	96.0

Appendix B

Free amino acids and ammonium chloride in a modified synthetic must

Table A2. Free amino acid and ammonium chloride concentrations in a modified synthetic must.

Compound	Amount (mg·L ⁻¹)	Compound	Amount (mg·L ⁻¹)
Tryptophan	38.7	Alanine	31.4
Phenylalanine	8.2	Threonine	16.4
Isoleucine	7.1	Glutamic acid	26.0

Table A2. *Cont.*

Compound	Amount (mg·L ⁻¹)	Compound	Amount (mg·L ⁻¹)
Leucine	10.5	Aspartic acid	9.6
Valine	9.6	Glycine	4.0
Methionine	6.8	Arginine	80.8
Tyrosine	4.0	Glutamine	109.1
Lysine	3.7	Serine	17.0
Cysteine	2.8	Asparagine	0.0
Proline	132.2	Histidine	7.1
NH ₄ CL	102.5		

Appendix C

Observed nitrogen consumption

Table A3. The observed nitrogen consumption of free amino acids (FAA), bound amino acids (BAA) and ammonia during fermentation with Lalvin ICV Opale 2.0™, Lalvin Persy™ and Lalvin QA23™ of the modified synthetic must. ‘Control’ is without added peptides, ‘BSA’ is with added peptides. Standard deviations represent measurements in three technical replicates.

Strain		N from FAA (mg·L ⁻¹)	N from BAA (mg·L ⁻¹)	N from NH ₄ (mg·L ⁻¹)	N Total (mg·L ⁻¹)
Lalvin ICV Opale 2.0™	Control	70.20 ± 0.07	-	18.08 ± 6.87	88.2 ± 6.87
	BSA	77.44 ± 0.22	74.4 ± 2.12	22.49 ± 2.23	174.3 ± 3.08
Lalvin Persy™	Control	75.8 ± 0.17	-	26.96 ± 6.35	102.8 ± 6.65
	BSA	72.0 ± 0.18	57.3 ± 0.53	25.41 ± 5.07	154.7 ± 5.10
Lalvin QA23™	Control	75.50 ± 0.20	-	25.13 ± 6.19	100.6 ± 6.49
	BSA	73.79 ± 0.16	51.7 ± 5.87	25.41 ± 5.07	150.9 ± 7.76

Appendix D

The proposed peptide candidates and their relative consumption by yeast

Table A4. The proposed dipeptide sequences and observed relative consumption during fermentation with Lalvin ICV Opale 2.0™, Lalvin Persy™ and Lalvin QA23™ of the modified synthetic must. Peptide sequences were acquired by peptide mapping to bovine serum albumin protein sequence (UniProtKB-P02769 (ALBU_BOVIN)). Additionally presented are the retention time, observed *m/z*, collision cross section values, as well as the averaged peptide intensities at the starting point and associated relative standard deviations.

Peptide Candidate	Retention Time (min)	Observed <i>m/z</i>	Charge	CCS (Å ²)	Average Original Intensity	RSD	Lalvin ICV Opale 2.0™ (Consumption, %)	Lalvin Persy™ (Consumption, %)	Lalvin QA23™ (Consumption, %)
AH	2.68	227.1134	1	150.41	807	4.86	99.97	100.00	99.98
AW	5.40	276.1338	1	160.10	1914	5.03	100.00	100.00	100.00
AY	4.56	253.1177	1	157.67	1470	6.44	100.00	100.00	100.00
FT	3.72	267.1335	1	163.99	5487	3.53	100.00	100.00	100.00
FW	7.03	352.1652	1	156.33	2617	5.18	100.00	100.00	100.00
FY	5.38	329.1487	1	179.32	698	3.61	97.94	95.49	95.50
JT	3.18	233.1492	1	156.95	2875	4.95	99.99	100.00	100.00
KQ	2.39	275.1709	1	160.14	1164	2.32	99.99	99.82	99.87
KT	1.40	248.1601	1	156.18	2201	8.75	99.97	99.98	100.00

Table A4. Cont.

Peptide Candidate	Retention Time (min)	Observed <i>m/z</i>	Charge	CCS (Å ²)	Average Original Intensity	RSD	Lalvin ICV Opale 2.0™ (Consumption, %)	Lalvin Persy™ (Consumption, %)	Lalvin QA23™ (Consumption, %)
LL	5.84	245.1856	1	166.83	20,475	3.04	100.00	100.00	100.00
LY	4.88	295.1648	1	171.64	4360	3.64	100.00	100.00	100.00
TF	4.78	267.1335	1	160.50	1507	3.99	100.00	99.87	100.00
TK	2.51	248.1600	1	156.18	2070	2.46	100.00	100.00	100.00
VE	2.14	247.1285	1	154.51	6009	6.03	99.96	99.98	99.95
VF	5.50	265.1542	1	164.08	5044	4.39	100.00	100.00	100.00
VN	1.37	232.1288	1	150.14	1012	2.78	99.98	100.00	100.00
VR	3.30	274.1866	1	167.20	1039	5.77	100.00	100.00	100.00
YE	3.72	311.1229	1	172.79	1091	4.24	99.67	100.00	100.00
YY	4.55	345.1441	1	184.22	5445	5.16	100.00	100.00	100.00

Table A5. The proposed tripeptide sequences and observed relative consumption during fermentation with Lalvin ICV Opale 2.0™, Lalvin Persy™ and Lalvin QA23™ of the modified synthetic must. Peptide sequences were acquired by peptide mapping to bovine serum albumin protein sequence (UniProtKB–P02769 (ALBU_BOVIN)). Additionally presented are the retention time, observed *m/z*, collision cross section values, as well as the averaged peptide intensities at the starting point and associated relative standard deviations.

Peptide Candidate	Retention Time (min)	Observed <i>m/z</i>	Charge	CCS (Å ²)	Average Original Intensity	%RSD	Lalvin ICV Opale 2.0™ (Consumption, %)	Lalvin Persy™ (Consumption, %)	Lalvin QA23™ (Consumption, %)
AEF	5.27	366.1656	1	181.69	5873	4.08	100.00	100.00	100.00
AWS	4.68	363.1658	1	181.79	2470	4.50	100.00	100.00	100.00
DAF	5.21	352.1496	1	176.71	1100	2.70	100.00	100.00	100.00
DLL	5.76	360.2124	1	189.23	2953	6.90	100.00	100.00	100.00
DQF	4.88	409.1712	1	195.14	728	3.28	100.00	100.00	100.00
ELT	4.01	362.1919	1	187.32	3543	2.99	99.98	99.62	99.92
FKD	3.53	205.1072	2	249.17	7981	3.00	100.00	100.00	100.00
FKG	3.33	176.1043	2	242.20	2685	2.41	100.00	100.00	100.00
FLG	5.65	336.1910	1	182.71	734	3.52	100.00	100.00	100.00
FQE	4.32	423.1866	1	198.49	658	3.50	100.00	100.00	100.00
FSA	4.73	324.1549	1	174.09	3223	2.23	100.00	100.00	100.00
FTF	7.23	414.2020	1	194.99	3313	6.62	100.00	100.00	100.00
FVE	4.88	394.1970	1	190.00	3242	3.60	100.00	100.00	100.00
GSF	4.64	310.1391	1	167.50	1014	6.12	100.00	100.00	100.00
IAE	3.72	332.1811	1	175.60	2444	4.33	100.00	100.00	100.00
IAF	6.63	350.2071	1	180.40	5291	6.60	100.00	100.00	100.00
IAH	2.66	340.1973	1	182.57	1628	7.50	100.00	100.00	100.00
IAR	3.08	359.2396	1	185.57	2132	6.89	100.00	100.00	100.00
ISL	6.42	332.2170	1	182.85	933	5.48	100.00	100.00	100.00
IVR	3.31	194.1387	2	253.10	4260	4.53	99.92	99.99	99.80
KIE	3.26	195.1228	2	243.74	2444	6.39	99.61	99.86	99.84
KQI	2.40	194.6309	2	246.86	8652	6.83	100.00	100.00	100.00
LAK	2.50	331.2332	1	181.06	482	3.99	100.00	100.00	100.00
LEE	3.90	390.1876	1	188.27	1557	5.17	100.00	93.10	94.29
LFT	5.92	380.2172	1	190.43	706	3.45	100.00	100.00	100.00
LLF	8.00	392.2539	1	197.53	1399	5.91	100.00	100.00	100.00
LSQ	3.30	347.1919	1	178.69	3179	5.20	100.00	100.00	100.00
LVE	4.37	360.2126	1	183.71	18,437	5.06	100.00	100.00	100.00
LVN	3.77	345.2125	1	180.57	990	4.59	100.00	100.00	100.00
LYE	4.68	424.2076	1	196.58	9328	3.77	100.00	100.00	100.00
LYY	6.08	458.2284	1	207.03	3200	6.61	100.00	99.99	100.00
SQY	3.78	397.1714	1	186.22	3568	5.93	100.00	100.00	100.00
SVL	5.63	318.2013	1	176.11	265	7.45	100.00	100.00	99.94
TEF	5.23	396.1756	1	189.94	620	3.45	100.00	100.00	100.00
TLV	4.44	332.2174	1	181.03	1307	2.71	100.00	100.00	100.00
VAF	6.01	336.1915	1	175.46	10,184	2.70	100.00	99.96	100.00
VNE	3.12	361.1714	1	178.22	5809	4.62	100.00	100.00	100.00

Table A5. Cont.

Peptide Candidate	Retention Time (min)	Observed <i>m/z</i>	Charge	CCS (Å ²)	Average Original Intensity	%RSD	Lalvin ICV Opale 2.0™ (Consumption, %)	Lalvin Persy™ (Consumption, %)	Lalvin QA23™ (Consumption, %)
VTF	5.96	366.2019	1	179.87	1740	6.69	100.00	100.00	100.00
VTK	1.40	347.2286	1	180.50	3469	6.77	100.00	100.00	100.00
YEY	5.06	474.1869	1	208.53	3085	5.50	99.98	100.00	99.99
YNG	3.07	353.1448	1	183.95	737	3.00	100.00	100.00	100.00

Table A6. The proposed tetrapeptide sequences and observed relative consumption during fermentation with Lalvin ICV Opale 2.0™, Lalvin Persy™ and Lalvin QA23™ of the modified synthetic must. Peptide sequences were acquired by peptide mapping to bovine serum albumin protein sequence (UniProtKB-P02769 (ALBU_BOVIN)). Additionally presented are the retention time, observed *m/z*, collision cross section values, as well as the averaged peptide intensities at the starting point and associated relative standard deviations.

Peptide Candidate	Retention Time (min)	Observed <i>m/z</i>	Charge	CCS (Å ²)	Average Original Intensity	%RSD	Lalvin ICV Opale 2.0™ (Consumption, %)	Lalvin Persy™ (Consumption, %)	Lalvin QA23™ (Consumption, %)
ALVE	5.13	431.2496	1	200.15	3448	7.14	100.00	98.07	99.34
FAKT	3.49	233.6360	2	266.06	2319	7.68	100.00	100.00	100.00
FEGK	5.44	269.1441	2	279.98	9697	5.25	74.05	46.42	96.43
FEKL	5.27	268.6570	2	276.78	1106	8.84	96.81	95.23	93.37
FHAD	3.66	245.1077	2	262.19	5008	5.24	100.00	100.00	99.99
FKDL	5.26	261.6491	2	277.18	3826	7.25	100.00	100.00	99.40
FLGS	5.62	423.2233	1	198.49	1214	1.92	100.00	100.00	100.00
FSQY	5.07	544.2400	1	222.40	5451	4.56	99.89	99.65	99.77
GERA	1.71	216.6131	2	254.59	646	5.90	84.37	71.04	78.31
IETM	5.11	493.2326	1	215.78	10,279	5.05	100.00	95.85	98.39
IKQN	1.91	251.6520	2	261.81	745	8.35	100.00	95.20	98.59
ISSK	1.99	217.6335	2	257.65	1951	10.49	100.00	99.71	99.07
KDAF	4.45	480.2451	1	204.55	4169	6.61	93.70	37.87	41.44
LEKS	2.52	238.6388	2	262.58	3371	5.37	100.00	100.00	99.97
LGSF	6.08	423.2236	1	196.61	7284	4.72	93.04	76.01	95.77
LILN	6.24	472.3125	1	218.28	2581	5.87	100.00	100.00	100.00
LIVR	4.82	250.6806	2	277.82	1165	3.78	75.70	49.87	50.76
LLEK	4.13	251.6648	2	271.33	4134	3.64	91.38	48.34	49.20
LRET	3.30	259.6495	2	267.68	915	7.21	100.00	100.00	100.00
LTAD	3.81	419.2132	1	194.85	2055	6.78	100.00	100.00	100.00
LTEF	6.28	509.2602	1	217.34	2520	6.27	100.00	100.00	100.00
LVEL	6.85	473.2970	1	214.35	7655	5.36	100.00	99.36	99.86
LVNE	4.15	474.2558	1	208.53	6118	6.06	99.95	99.71	99.98
LVTD	4.33	447.2446	1	201.61	8109	3.72	100.00	99.72	99.85
MENF	5.75	540.2122	1	222.49	2536	6.75	100.00	100.00	100.00
TQTA	2.58	420.2083	1	196.70	1160	6.33	100.00	100.00	100.00
VASL	2.93	195.1228	2	253.02	1457	5.55	99.77	99.68	99.85
VEVS	4.02	433.2284	1	198.21	2150	6.29	98.70	97.71	97.77
VFDK	4.02	254.6413	2	267.97	1976	5.98	69.07	35.82	53.99
VSEK	1.95	231.6310	2	259.87	2212	7.45	99.97	97.52	92.49
VVST	3.70	405.2340	1	191.53	3913	3.23	100.00	100.00	100.00

Table A7. The proposed pentapeptide sequences and observed relative consumption during fermentation with Lalvin ICV Opale 2.0™, Lalvin Persy™ and Lalvin QA23™ of the modified synthetic must. Peptide sequences were acquired by peptide mapping to bovine serum albumin protein sequence (UniProtKB–P02769 (ALBU_BOVIN)). Additionally presented are the retention time, observed *m/z*, collision cross section values, as well as the averaged peptide intensities at the starting point and associated relative standard deviations.

Peptide Candidate	Retention Time (min)	Observed <i>m/z</i>	Charge	CCS (Å ²)	Average Original Intensity	RSD	Lalvin ICV Opale 2.0™ (Consumption, %)	Lalvin Persy™ (Consumption, %)	Lalvin QA23™ (Consumption, %)
AFDEK	3.91	305.1472	2	274.91	15,458	6.56	99.93	36.21	41.18
AIPEN	4.24	543.2772	1	218.50	8777	4.98	96.76	33.49	33.68
ALVEL	7.32	544.3338	1	230.33	912	7.41	99.99	40.31	48.36
FDEKL	5.50	326.1703	2	300.05	5931	5.12	100.00	58.69	61.54
FLGSF	7.55	570.2920	1	231.74	934	6.74	99.96	33.44	31.64
FYAPE	5.74	626.2814	1	244.72	2299	4.18	100.00	66.49	66.40
GFQNA	4.50	536.2458	1	220.62	1214	7.16	78.76	34.47	38.77
KFWGK	5.00	333.1909	2	296.41	878	4.19	100.00	99.98	100.00
LAKEY	4.23	312.1727	2	277.79	4602	2.52	100.00	61.22	91.98
LFGDE	6.00	580.2609	1	227.53	825	7.83	99.79	28.10	35.07
LGEYG	4.95	538.2504	1	218.61	831	3.51	90.99	32.33	35.78
LILNR	5.33	314.7094	2	287.40	797	7.07	99.75	30.78	47.93
LPKIE	5.42	300.1912	2	284.85	8843	6.91	100.00	60.78	99.87
LVELL	7.98	586.3803	1	241.49	707	5.49	90.22	33.34	64.02
LVEVS	5.21	546.3130	1	224.32	4239	0.63	99.26	39.13	54.87
TAPEL	6.26	592.2976	1	235.28	15,546	5.29	100.00	93.48	93.12
TVFDK	4.61	305.1650	2	284.60	2748	4.80	99.98	66.96	61.71
VEVTK	3.65	288.1730	2	275.74	7041	8.04	98.83	39.91	48.04
VVSTQ	3.53	533.2928	1	218.72	3638	5.90	98.97	43.60	46.92

Table A8. The proposed hexapeptide sequences and observed relative consumption during fermentation with Lalvin ICV Opale 2.0™, Lalvin Persy™ and Lalvin QA23™ of the modified synthetic must. Peptide sequences were acquired by peptide mapping to bovine serum albumin protein sequence (UniProtKB–P02769 (ALBU_BOVIN)). * Elution within void volume. Additionally presented are the retention time, observed *m/z*, collision cross section values, as well as the averaged peptide intensities at the starting point and associated relative standard deviations.

Peptide Candidate	Retention Time (min)	Observed <i>m/z</i>	Charge	CCS (Å ²)	Average Original Intensity	%RSD	Lalvin ICV Opale 2.0™ (Consumption, %)	Lalvin Persy™ (Consumption, %)	Lalvin QA23™ (Consumption, %)
AFDEKL	5.70	361.6892	2	288.60	36,673	6.08	88.2	42.5	40.9
DAFLGS	3.91	305.1472	2	274.91	15,458	6.56	99.9	36.2	41.2
DTHKSE *	1.19	358.6637	2	292.00	4537	2.90	95.4	8.3	11.6
KDAIPE	4.55	336.6810	2	286.40	952	7.69	71.7	35.6	27.2
KFGERA	3.66	354.1946	2	298.79	1133	3.60	93.9	22.1	27.9
KFPKAE	4.01	360.2070	2	308.56	2291	5.69	100.0	39.4	44.3
LFGDEL	7.43	693.3452	1	253.69	902	10.18	100.0	34.2	32.5
LLPKIE	6.10	356.7331	2	315.47	1720	7.65	100.0	41.7	46.1
LPKIET	5.60	350.7149	2	292.34	7905	6.27	96.9	31.0	31.1
NLPPLT	6.57	654.3815	1	248.25	1517	7.63	80.2	51.5	43.0
TVFDKL	6.40	361.7072	2	295.17	2718	5.65	100.0	35.9	40.2
VEGPKL	5.33	321.6941	2	280.58	9864	5.33	85.6	38.7	39.6
VSTPTL	6.14	617.3495	1	238.79	674	5.56	100.0	43.6	46.0

References

1. Cruz, S.H.; Cilli, E.M.; Ernandes, J.R. Structural Complexity of the Nitrogen Source and Influence on Yeast Growth and Fermentation. *J. Inst. Brew.* **2002**, *108*, 54–61. [CrossRef]
2. Lei, H.; Zhao, H.; Zhao, M. Proteases supplementation to high gravity worts enhances fermentation performance of brewer’s yeast. *Biochem. Eng. J.* **2013**, *77*, 1–6. [CrossRef]
3. Verbelen, P.J.; Delvaux, F.R. Brewing yeast in action: Beer fermentation. *Appl. Mycol.* **2009**, *7*, 110–135. [CrossRef]

4. Nisbet, M.A.; Martinson, T.E.; Mansfield, A.K. Accumulation and Prediction of Yeast Assimilable Nitrogen in New York Winegrape Cultivars. *Am. J. Enol. Vitic.* **2014**, *65*, 325–332. [[CrossRef](#)]
5. Gibson, B.R.; Lawrence, S.; LeClaire, J.P.R.; Powell, C.; Smart, K.A. Yeast responses to stresses associated with industrial brewery handling: Figure 1. *FEMS Microbiol. Rev.* **2007**, *31*, 535–569. [[CrossRef](#)]
6. Malfeito-Ferreira, M. Yeasts and wine off-flavours: A technological perspective. *Ann. Microbiol.* **2010**, *61*, 95–102. [[CrossRef](#)]
7. Vilanova, M.; Ugliano, M.; Varela, C.; Siebert, T.; Pretorius, I.S.; Henschke, P.A. Assimilable Nitrogen Utilisation and Production of Volatile and Non-Volatile Compounds in Chemically Defined Medium by *Saccharomyces Cerevisiae* Wine Yeasts. *Appl. Microbiol. Biotechnol.* **2007**, *77*, 145–157. [[CrossRef](#)]
8. Bely, M.; Rinaldi, A.; Dubourdieu, D. Influence of Assimilable Nitrogen on Volatile Acidity Production by *Saccharomyces Cerevisiae* during High Sugar Fermentation. *J. Biosci. Bioeng.* **2003**, *96*, 507–512. [[CrossRef](#)]
9. Becerra-Rodríguez, C.; Marsit, S.; Galeote, V. Diversity of Oligopeptide Transport in Yeast and Its Impact on Adaptation to Winemaking Conditions. *Front. Genet.* **2020**, *11*, 602. [[CrossRef](#)]
10. Becerra-Rodríguez, C.; Taghouti, G.; Portier, P.; Dequin, S.; Casal, M.; Paiva, S.; Galeote, V. Yeast Plasma Membrane Fungal Oligopeptide Transporters Display Distinct Substrate Preferences despite Their High Sequence Identity. *J. Fungi* **2021**, *7*, 963. [[CrossRef](#)]
11. Ganapathy, V.; Leibach, F.H. Proton-coupled solute transport in the animal cell plasma membrane. *Curr. Opin. Cell Biol.* **1991**, *3*, 695–701. [[CrossRef](#)]
12. Marsit, S.; Mena, A.; Bigey, F.; Sauvage, F.-X.; Couloux, A.; Guy, J.; Legras, J.-L.; Barrio, E.; Dequin, S.; Galeote, V. Evolutionary Advantage Conferred by an Eukaryote-to-Eukaryote Gene Transfer Event in Wine Yeasts. *Mol. Biol. Evol.* **2015**, *32*, 1695–1707. [[CrossRef](#)] [[PubMed](#)]
13. Payne, J.W.; Smith, M.W. Peptide Transport by Micro-Organisms. *Adv. Microb. Physiol.* **1994**, *36*, 1–80. [[CrossRef](#)] [[PubMed](#)]
14. Damon, C.; Vallon, L.; Zimmermann, S.; Haider, M.Z.; Galeote, V.; Dequin, S.; Luis, P.; Fraissinet-Tachet, L.; Marmeisse, R. A novel fungal family of oligopeptide transporters identified by functional metatranscriptomics of soil eukaryotes. *ISME J.* **2011**, *5*, 1871–1880. [[CrossRef](#)]
15. Hauser, M.; Donhardt, A.M.; Barnes, D.; Naider, F.; Becker, J.M. Enkephalins Are Transported by a Novel Eukaryotic Peptide Uptake System. *J. Biol. Chem.* **2000**, *275*, 3037–3041. [[CrossRef](#)]
16. Lubkowitz, M.A.; Barnes, D.; Breslav, M.; Burchfield, A.; Naider, F.; Becker, J.M. *Schizosaccharomyces pombe* isp4 encodes a transporter representing a novel family of oligopeptide transporters. *Mol. Microbiol.* **2002**, *28*, 729–741. [[CrossRef](#)]
17. Lubkowitz, M.A.; Hauser, L.; Breslav, M.; Naider, F.; Becker, J.M. An oligopeptide transport gene from *Candida albicans*. *Microbiology* **1997**, *143*, 387–396. [[CrossRef](#)]
18. Wiles, A.M.; Cai, H.; Naider, F.; Becker, J.M. Nutrient regulation of oligopeptide transport in *Saccharomyces cerevisiae*. *Microbiology* **2006**, *152*, 3133–3145. [[CrossRef](#)]
19. Harscoat-Schiavo, C.; Nioi, C.; Ronat-Heit, E.; Paris, C.; Vanderesse, R.; Fournier, F.; Marc, I. Hydrophilic properties as a new contribution for computer-aided identification of short peptides in complex mixtures. *Anal. Bioanal. Chem.* **2012**, *403*, 1939–1949. [[CrossRef](#)]
20. Huang, Y.-P.; Dias, F.F.G.; de Moura, J.M.L.N.; Barile, D. A complete workflow for discovering small bioactive peptides in foods by LC-MS/MS: A case study on almonds. *Food Chem.* **2021**, *369*, 130834. [[CrossRef](#)]
21. Le Maux, S.; Nongonierma, A.B.; Murray, B.; Kelly, P.M.; FitzGerald, R.J. Identification of short peptide sequences in the nanofiltration permeate of a bioactive whey protein hydrolysate. *Food Res. Int.* **2015**, *77*, 534–539. [[CrossRef](#)]
22. Piovesana, S.; Capriotti, A.L.; Cerrato, A.; Crescenzi, C.; La Barbera, G.; Laganà, A.; Montone, C.M.; Cavaliere, C. Graphitized Carbon Black Enrichment and UHPLC-MS/MS Allow to Meet the Challenge of Small Chain Peptidomics in Urine. *Anal. Chem.* **2019**, *91*, 11474–11481. [[CrossRef](#)] [[PubMed](#)]
23. Sonsmann, G.; Römer, A.; Schomburg, D. Investigation of the influence of charge derivatization on the fragmentation of multiply protonated peptides. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 47–58. [[CrossRef](#)]
24. Hernández-Mesa, M.; Escourrou, A.; Monteau, F.; Le Bizec, B.; Dervilly-Pinel, G. Current applications and perspectives of ion mobility spectrometry to answer chemical food safety issues. *TrAC Trends Anal. Chem.* **2017**, *94*, 39–53. [[CrossRef](#)]
25. Arju, G.; Taivosalo, A.; Pismennoi, D.; Lints, T.; Vilu, R.; Daneberga, Z.; Vorslova, S.; Renkonen, R.; Joenvaara, S. Application of the UHPLC-DIA-HRMS Method for Determination of Cheese Peptides. *Foods* **2020**, *9*, 979. [[CrossRef](#)]
26. Campuzano, I.D.G.; Giles, K. Historical, Current and Future Developments of Travelling Wave Ion Mobility Mass Spectrometry: A Personal Perspective. *TrAC Trends Anal. Chem.* **2019**, *120*, 115620. [[CrossRef](#)]
27. Fiechter, G.; Mayer, H. UPLC analysis of free amino acids in wines: Profiling of on-lees aged wines. *J. Chromatogr. B* **2011**, *879*, 1361–1366. [[CrossRef](#)]
28. Salmon, J.-M.; Barre, P. Improvement of Nitrogen Assimilation and Fermentation Kinetics under Enological Conditions by Derepression of Alternative Nitrogen-Assimilatory Pathways in an Industrial *Saccharomyces cerevisiae* Strain. *Appl. Environ. Microbiol.* **1998**, *64*, 3831–3837. [[CrossRef](#)]
29. Bell, S.-J.; Henschke, P.A. Implications of Nitrogen Nutrition for Grapes, Fermentation and Wine. *Aust. J. Grape Wine Res.* **2008**, *11*, 242–295. [[CrossRef](#)]
30. Bely, M.; Sablayrolles, J.-M.; Barre, P. Automatic detection of assimilable nitrogen deficiencies during alcoholic fermentation in oenological conditions. *J. Ferment. Bioeng.* **1990**, *70*, 246–252. [[CrossRef](#)]

31. Adler-Nissen, J. *Enzymic Hydrolysis of Food Proteins*; Elsevier Science Limited: London, UK, 1986.
32. Pasupuleti, V.K.; Demain, A.L. *Protein Hydrolysates in Biotechnology*; Springer: Cham, Switzerland, 2010. [[CrossRef](#)]
33. Aloo, S.O.; Oh, D.-H. The Functional Interplay between Gut Microbiota, Protein Hydrolysates/Bioactive Peptides, and Obesity: A Critical Review on the Study Advances. *Antioxidants* **2022**, *11*, 333. [[CrossRef](#)] [[PubMed](#)]
34. Colla, G.; Hoagland, L.; Ruzzi, M.; Cardarelli, M.; Bonini, P.; Canaguier, R.; Roupheal, Y. Biostimulant Action of Protein Hydrolysates: Unraveling Their Effects on Plant Physiology and Microbiome. *Front. Plant Sci.* **2017**, *8*, 2202. [[CrossRef](#)] [[PubMed](#)]
35. Van Loon, L.J.; Kies, A.K.; Saris, W.H. Protein and protein hydrolysates in sports nutrition. *Int. J. Sport Nutr. Exerc. Metab.* **2007**, *17*, S1–S4. [[CrossRef](#)] [[PubMed](#)]
36. Liigand, P.; Kaupmees, K.; Kruve, A. Influence of the amino acid composition on the ionization efficiencies of small peptides. *Biol. Mass Spectrom.* **2019**, *54*, 481–487. [[CrossRef](#)] [[PubMed](#)]
37. Hutchinson-Bunch, C.; Sanford, J.A.; Hansen, J.R.; Gritsenko, M.A.; Rodland, K.D.; Piehowski, P.D.; Qian, W.-J.; Adkins, J.N. Assessment of TMT Labeling Efficiency in Large-Scale Quantitative Proteomics: The Critical Effect of Sample pH. *ACS Omega* **2021**, *6*, 12660–12666. [[CrossRef](#)]
38. Waliczek, M.; Kijewska, M.; Rudowska, M.; Setner, B.; Stefanowicz, P.; Szewczuk, Z. Peptides Labeled with Pyridinium Salts for Sensitive Detection and Sequencing by Electrospray Tandem Mass Spectrometry. *Sci. Rep.* **2016**, *6*, 37720. [[CrossRef](#)]
39. Zecha, J.; Satpathy, S.; Kanashova, T.; Avanesian, S.C.; Kane, H.; Clauser, K.; Mertins, P.; Carr, S.A.; Kuster, B. TMT Labeling for the Masses: A Robust and Cost-efficient, In-solution Labeling Approach. *Mol. Cell. Proteom.* **2019**, *18*, 1468–1478. [[CrossRef](#)]

Curriculum vitae

Personal data

Name: Georg Arju
Date of birth: 5 April 1991
Place of birth: Estonia
Citizenship: Estonian

Contact data

E-mail: aryu.georg1991@gmail.com

Education

2016–2022 Tallinn University of Technology, Estonia – PhD
2014–2015 University of Glasgow, Scotland – MSc (MMed)
2010–2014 University of Abertay Dundee, Scotland – BSc (Hons)

Employment

2020–Present Center of Food and Fermentation Technologies
Senior Application Manager

2017–2020 Center of Food and Fermentation Technologies
Advanced LC-MS Application Manager

2015–2017 Center of Food and Fermentation Technologies
Research Analyst

2019–Present DermTipp Nahakliinik
Chief Technical Officer

2017–Present Faneks Eesti
Field Service Engineer and Application Specialist

Elulookirjeldus

Isikuandmed

Nimi: Georg Arju
Sünniaeg: 5. Aprill, 1991
Sünnikoht: Eesti
Kodakondsus: Eesti

Kontaktandmed

E-post: aryu.georg1991@gmail.com

Hariduskäik

2016–2022 Tallinna Tehnikaülikool, Eesti – PhD
2014–2015 University of Glasgow, Šotimaa – MSc (MMed)
2010–2014 University of Abertay Dundee, Šotimaa – BSc (Hons)

Teenistuskäik

2020–... Toidu- ja Fermentatsioonitehnoloogia Arenduskeskus
MS Vanem Rakendusjuht

2017–2020 Toidu- ja Fermentatsioonitehnoloogia Arenduskeskus
LC-MS Rakenduspetsialist

2015–2017 Toidu- ja Fermentatsioonitehnoloogia Arenduskeskus
Teadur

2019–... DermTipp Nahakliinik
Tehniline Direktor

2017–... Faneks Eesti
Hooldusinsener

ISSN 2585-6901 (PDF)
ISBN 978-9949-83-944-5 (PDF)