SLUTRAPPORT/FINAL REPORT
NR. 2024-192

Søren Balling Engelsen:

# MilkStreamValue – Forbindelser med lav molekylvægt i mælke- og mejeriprodukter - en potentiel ny kilde til ”added value” produkter

## MilkStreamValue – Low molecular weight compounds in milk and dairy streams - a potential new source for value added products

Mejeribrugets
ForskningsFond

December 2024

# Final report

## for collaborative projects funded via the Danish Dairy Research Foundation (DDRF)

## 1. Title of the project

**Danish:** Forbindelser med lav molekylvægt i mælke- og mejeriprodukter - en potentiel ny kilde til ”added value”produkter (MilkStreamValue)

**English:** Low molecular weight compounds in milk and dairy streams - a potential new source for value added products (MilkStreamValue)

## 2. Project manager

Professor Søren Balling Engelsen, University of Copenhagen, Department of Food Science, Rolighedsvej 26 1958 Frederiksberg C, DIR +45 35 33 32 05/ MOB +45 20 20 00 64, se@food.ku.dk

## 3. Other project staff

Assoc. Prof. Bekzod Khakimov, University of Copenhagen, Department of Food Science
PhD student Paraskevi Tsermoula, University of Copenhagen, Department of Food Science
Mie Rostved Bechshøft, Scientist, R&D – Discover, Arla Foods Ingredients Group P/S

## 4. Sources of funding

Danish Dairy Research Foundation, University of Copenhagen and Arla Foods Ingredients

## 5. Project period

**Project period with DDRF funding:** January/2020, January/2023

## 6. Project summary

**Danish:**

Målet med MILKSTREAMVALUE var at etablere følsomme og high-throughput analytiske metoder til bestemmelse af lavmolekylære forbindelser i mejeri produkter og matricer. Metoderne inkluderede proton kerne magnetisk resonans (1H NMR) spektroskopi og gaskromatografi-massespektrometri (GC-MS) i samspil med optimerede standard operationsprocedurer (SOP'er) og avancerede multivariate dataanalysemetoder. De udviklede metoder blev anvendt og valideret ved at analysere prøver fra en pilot skala laktoseproduktion hos Arla Foods Ingredients (AFI), på kommerciel komælk samt på kolostrum fra et fodringsforsøg.

Fokus i MILKSTREAMVALUE var de lavmolekylære forbindelser, der bidrager til non-protein nitrogen (NPN) fraktionen som en potentiel kilde til værdifulde forbindelser. Analysen af NPN-profilen af 10 prøver fra forsøg med laktoseproduktion ved AFI resulterede i identifikation og kvantificering af 37 NPN-forbindelser, herunder 13 aminosyrer, syv aminosyrederivater, fire aminoalkoholer, to organiske syrer, tre methylaminer, to nukleobaser, et nukleotid, et nukleosid, et aminosukker, et vitamin, urinstof og ammoniak. Kvantificering af disse NPN-forbindelser blev brugt til at illustrere virkningen af forskellige enhedsoperationer på de identificerede NPN-forbindelser og til at beregne procentdelen af genvundet nitrogen. Urinstof, ammoniak, glycerofosfocholin, kreatin, kreatinin, orotic acid og cholin var de mest dominerende forbindelser i det oprindelige materiale, der blev anvendt til laktoseproduktion. Forsøget viste også at laktosepulver indeholdt væsentlige (relative) mængder af N-acetylglucosamin, phosphocholin og orotic acid, mens den koncentrerede flydende væske efter laktosekrystallisation indeholdt den højeste koncentration af NPN. Kvælstoffet i NPN-forbindelserne summede op til 57–99 % af det samlede kvælstof, afhængigt af prøvetypen. Den højeste nitrogenrecovery blev fundet i prøver med høj NPN-procent, mens den laveste blev fundet for laktosepulver.

Ovennævnte resultater blev derefter udvidet og anvendt til en generel undersøgelse af lavmolekylære stoffer på tværs af en laktose krystalliseringsproces. I alt blev 110 forbindelser identificeret, og 49 blev kvantificeret. Denne undersøgelse viste hvorledes processerne i laktosefremstillingen indvirker på de lavmolekylære forbindelser der er tilstede i vallepermeat. Lavmolekylære komponenter (f.eks. methanol) og nogle få lidt større molekyler (f.eks. fedtsyrer) var i stand til at penetrere den omvendt osmose membranen, mens 23 forbindelser (f.eks. hydroxypyruvic acid, malonic acid, gluconic acid and ribonic acid) ko-krystalliserede med laktose (og dermed laktose pulveret).

Endelig blev en hurtig, effektiv og reproducerbar metode til behandling af mælke 1H NMR-spektre udviklet ved at etablere et ”chemical shift” bibliotek, der inkluderer 63 lavmolekylære mælkeforbindelser, og kombinere det med avancerede dataanalysemetoder. Denne metode bruger som input de rå 1H NMR-spektre og genererer en metabolittabel med kvantitative data. Metoden blev etableret ved at analysere kommerciel komælk og valideret på kolostrumprøver indsamlet fra malkekøer (omkring kælvning), hvilket muliggjorde absolut kvantificering af 63 mælkeforbindelser.

I løbet af dette projekt blev der etableret forbedrede og standardiserede protokoller til analyse af lavmolekylære forbindelser i prøver fra kommerciel mælk og forsøg med laktoseproduktion. Ved at sammenligne prøvernes molekylære profil under laktoseproduktion afslørede virkningen af de forskellige enhedsoperationer på deres sammensætning. Profilering af disse prøver giver derfor indsigt i hvorledes man kan optimere laktoseproduktionen og forbedre laktosepulverets sammensætning. Samtidig giver disse resultater indsigt i hvorledes man kan øge valorisering af sidestrømme gennem udvinding af gavnlige NPN-forbindelser til produktion af højværdifødevareingredienser og sikre dokumenteret kvalitet i mejerifødevareproduktionen.

**English:**

The aim of this project was to establish high-throughput analytical methods using proton nuclear magnetic resonance ($^1$H NMR) spectroscopy and gas chromatography-mass spectrometry (GC-MS), along with optimized standard operating procedures (SOPs) and advanced multivariate data analysis methods for reproducible measurement of low-molecular weight compounds in dairy matrices. The developed methods were validated by analyzing samples derived from trial lactose productions at Arla Foods Ingredients (AFI), commercial cow’s milk and cow colostrum from a feeding trial, which allowed the investigation of the low-molecular weight compounds present in these samples.

In the present project, special attention was given to the low-molecular weight compounds contributing to the non-protein nitrogen (NPN) fraction as a potential source of valuable compounds. The analysis of the NPN profile of 10 samples derived from trial lactose productions at AFI resulted in the identification and quantification of 37 NPN compounds including 13 amino acids, seven amino acid derivatives, four amino alcohols, two organic acids, three methylamines, two nucleobases, one nucleotide, one nucleoside, one amino sugar, one vitamin, urea and ammonia. Quantification of the NPN compounds facilitated the study of the effect of unit operations on the identified NPN compounds and the calculation of the percentage of recovered nitrogen from each sample type. Urea, ammonia, glycerophosphocholine, creatine, creatinine, orotic acid and choline were the most dominant compounds in the initial material used for

lactose production. Lactose powder had substantial relative amounts of N-acetylglucosamine, phosphocholine and orotic acid, while the concentrated liquid material after lactose crystallization and removal had the highest concentration of NPN. The nitrogen in the NPN compounds added up to 57–99% of the total nitrogen, depending on the sample type. The highest nitrogen recovery was found in samples with high NPN percentage, whereas the lowest was found for lactose powder.

The above-mentioned results were further expanded and the global foodome (complete set of chemical compounds found in food) of the same samples was investigated. In total, 110 compounds were identified and 49 were quantified. This study revealed the impact of lactose processing on the low-molecular weight compounds present in whey permeate. Small molecules (e.g. methanol) and a few larger molecules (e.g. fatty acids) permeated the reverse osmosis membrane, while 23 compounds (e.g. hydroxypyruvic acid, malonic acid, gluconic acid and ribonic acid) co-crystallized with lactose and ended up in lactose power.

Finally, a rapid, efficient and reproducible method for processing milk $^{1}$H NMR spectra was developed by establishing a chemical shift library including 63 milk low-molecular weight compounds and combining it with advanced data analysis methods. This method utilizes as input the raw $^{1}$H NMR spectra and generates a metabolite table with quantitative or semi-quantitative data. The method was established by analyzing commercial cow milk and validated on cow colostrum samples collected from periparturient dairy cows, enabling the absolute quantification of 63 compounds.

During this project, improved and standardized protocols were established for the analysis of low-molecular weight compounds in samples from trial lactose productions and milk. Linking the molecular profile of samples to processing steps during lactose production revealed the impact of unit operations to their composition. Therefore, profiling of these samples provides insights for optimizing lactose production and improving lactose powder composition. At the same time these results highlight possibilities for increasing valorization of side-streams through extraction of beneficial NPN compounds for the production of high-value food ingredients and ensure documented quality in dairy food production.

## 7. Project aim

**Danish:**

Projektet har til formål at etablere analytiske metoder til high-throughput analyse af mindre bestanddele (lav koncentration) i mælkeprodukter, mælke- og vallefraktioner samt at beskrive kompositionen af udvalgte produkter og mælke- og vallefraktioner. Mange mindre bestanddele i mejeriprodukter er stadig ukendte og deres potential for at bidrage til værdisætning af produkter og ingredienser er derfor ukendt. Projektet vil implementere nye high-throughput målemetoder, der er solidt forankret på de analytiske foodomics platforme (GC-MS, LC-MS og LC-NMR/MS), Design of Experiment (DoE), optimerede standard operating procedures (SOP'er) og avanceret multivariativ dataanalyse. Det er målet at etablere præcise og reproducerbare målinger af mindre mælkekomponenter, herunder hidtil ukendte fraktioner såsom ikke-protein nitrogen (NPN). Derudover vil projektet etablere den første omfattende mælke-metabolom-database, indeholdende NPN og andre ukendte stoffer, som vil lette en fremtidig hurtig og omfattende identifikation af mælkerelaterede kemiske komponenter.

**English:** The project aims to extend current knowledge on the unknown substances of milk products and establish high-throughput and robust analytical methodologies for detailed screening of minor components of milk derived products including whey powders.

The project has three objectives:

(1). Development and establishment of Standardized and Optimized Foodomics proTocols for MILK (SOFTMILK). The SOFTMILK platform will cover milk process streams, milk and whey powders and will utilize the state-of-the-art methods developed within analytical chemistry, design of experiments (DoE), multivariate data analysis, and metabolomics fields.

(2). Development of a comprehensive Milk Foodome Database (MFD) covering small molecular size substances that contribute to the Non-Protein Nitrogen (NPN) and ash parts of milk and milk products. MFD will cover not only natural small metabolites but also potentially deleterious and process related substances.

(3). Extend the current knowledge on small molecular components of milk products through continues screening of milk, whey powders, and casein fractions in various production lines at Arla Foods Ingredients (Denmark) during the period of 12 months. This will allow investigation of dynamics of the milk NPN substances, and possible deleterious substances as a function of milk processing steps including milk fractionation.

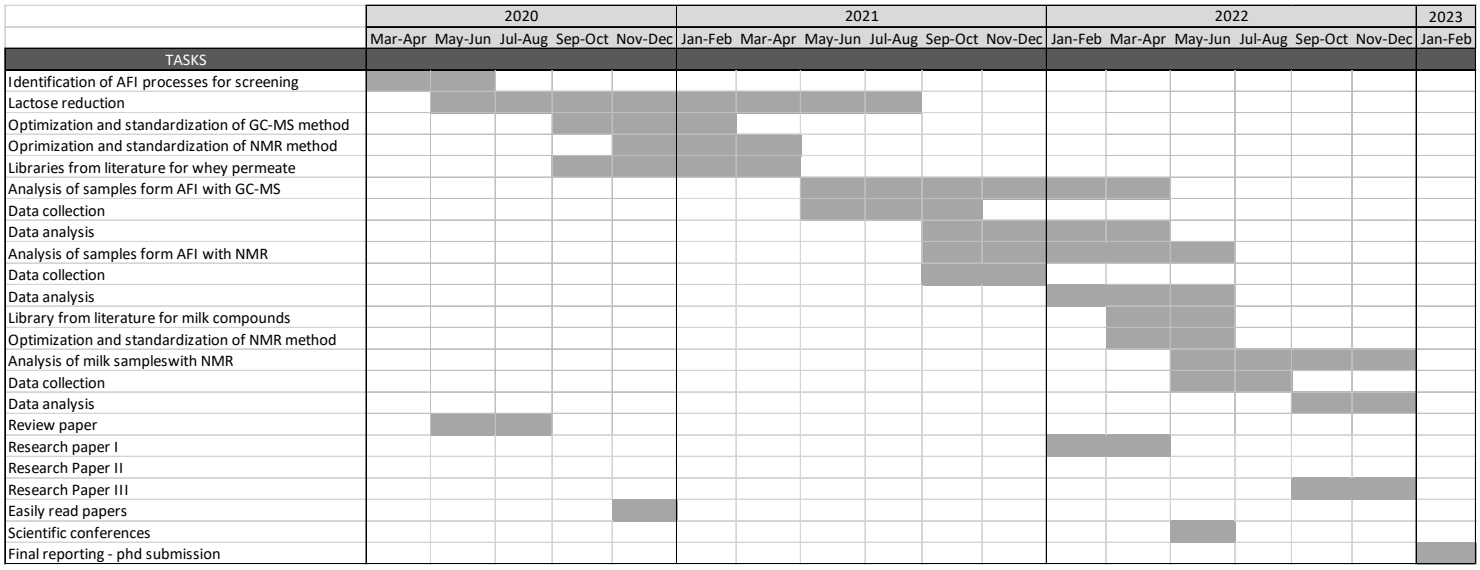
## 8. Background for the project

The biggest share of the milk production worldwide (36%) is used for cheese manufacturing. During this process, whey is expelled, which is further processed into whey-based products and ingredients. Whey is the largest volume co-product of the dairy industry, and while significant research has been conducted on its chemical composition, the content of low-molecular weight components has been overlooked. With low-molecular weight components in whey remaining largely unknown, their potential valorization for the production of new value-added products remains limited.

Processing of whey and production of whey protein ingredients result in the generation of whey retentate, primarily used for the production of whey protein ingredients, and whey permeate, which is mainly used for lactose production due to its high lactose content (> 80% on dry matter basis). It is assumed that the water-soluble low-molecular weight fraction of milk and whey ends up in whey permeate, but its detailed composition has never been systematically investigated. Instead, the composition of most dairy products is determined by using standardized classical analytical methods due to their simplicity. Such methods usually target one or few compounds of interest at a time, while other chemical substances remain undetected. Research on whey permeate composition has revealed its potential as a valuable source of oligosaccharides and bioactive peptides with demonstrated bioactivities, which intensify the interest in analyzing the non-protein nitrogen (NPN) fraction of whey permeate.

Today's demands for whey-derived ingredients require the accurate and comprehensive documentation of whey constituents for safety and standardizing quality of the end products. Although the molecular level details of whey proteins, lipids and carbohydrates present in whey have been investigated to a degree, whey also contains low-molecular weight compounds that have been overlooked. In order to identify and quantify the low-molecular weight components in whey and whey process streams, high-throughput analytical technologies and powerful data analysis capabilities are required. The dairy industry is actively seeking to add value to specific milk fractions and components for the production of high-value products such as infant formulas and medical foods. The low-molecular weight fraction of whey permeate requires detailed characterization to fully document it and unlock its potential value.

## 9. Sub-activities in the entire project period




## 10. Deviations

No significant deviations

## 11. Project results

During the present project, sample preparation, GC-MS and $^{1}$H NMR analysis methods for the untargeted analysis of milk and whey process samples were optimized and standardized. For the extraction of information after obtaining the results, advanced data analysis tools were employed. This facilitated the establishment of the Standardized and Optimized Foodomics proTocols for MILK (SOFTMILK) platform that covers milk and whey process streams and lactose powder which is the first objective of the expression of interest to DDRF.

The second aim of the project was to extend the current knowledge of the unknown substances of milk products with the objective to develop a comprehensive Milk Foodome Database (MFD) covering small molecular size substances including natural small metabolites but also potentially deleterious and process related substances. A Milk Foodome database containing 110 low-molecular weight (LMW) compounds, including 37 non-protein nitrogen (NPN) compounds has been developed, expanding the current knowledge of whey permeate composition. This database includes the identity of the compounds, their presence and concentration in different whey process samples and milk.

The standardized methods and the developed database from the first and second objective were utilized to screen a total of 54 samples from whey permeate processing and 6 samples of the final product (i.e. lactose powder) from three independent trial lactose productions performed at Arla Foods Ingredients Group P/S (Nr. Vium, Denmark). This helped to extend a current knowledge on small molecular components in these samples and investigate the effect of unit operations on the compounds dynamics as outlined in the third objective in the expression of interest to DDRF.

The schematic flow of the trial lactose productions at Arla Foods Ingredients (AFI) is presented in Fig. 1. In brief, permeate (UFP) was concentrated using a combination of reverse osmosis (RO) and evaporation. The permeate stream of RO filtration (ROP) was filtered using RO polisher membrane for the recovery of water into the permeate (ROPP) and retentate (ROPR) streams, while the ROPR was added to the RO feed as a diafiltration buffer. In sequence, reverse osmosis retentate (ROR) was evaporated and the evaporation concentrate (EVC) was then fed to a batch crystallization tank where lactose crystals were formed during a gradual cooling process. Mother liquor (ML) was decanted and lactose crystals were washed with water (WW) to remove non-lactose compounds. Lactose powder (LP) was produced by drying the purified lactose crystals in a fluid bed and ML was evaporated to obtain mother liquor concentrate (MLC).

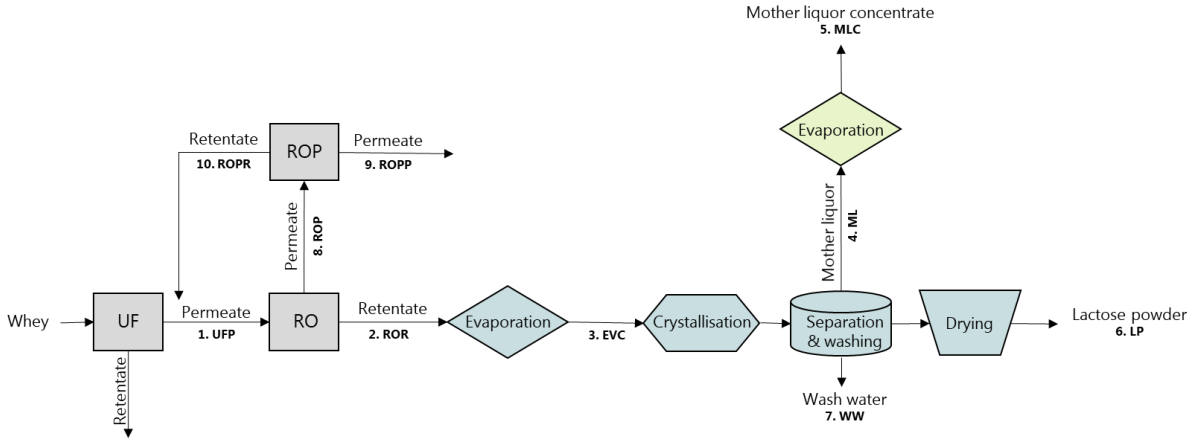



*Figure 1. Schematic diagram of trial lactose production process from whey permeate, including ultrafiltration (UF), reverse osmosis (RO), evaporation, and crystallization. The scheme also shows the recovery and processing of RO permeate with RO polisher (ROP) and mother liquor recovery and evaporation. The sampling points of the process streams are marked from 1 to 10. The samples taken are: 1. UFP (whey ultrafiltration permeate), 2. ROR (RO retentate), 3. EVC (evaporation concentrate), (4) mother liquor (ML), (5) mother liquor concentrate (MLC), (6) lactose powder (LP), (7) wash water (WW), (8) RO permeate (ROP), (9) ROP permeate (ROPP), (10) ROP retentate (ROPR).*

Whey ultrafiltration permeate (UFP), which serves as the starting material for lactose production, contains numerous LMW compounds. In this project, 110 of these compounds were identified and 51 quantified. These compounds have the potential to interact with lactose and interfere with the crystallization process or limit the run times of unit operations such as evaporation and RO filtration. Therefore, profiling these compounds during lactose production can provide insight into the effect of unit operations on their dynamics and the overall impact on the lactose production process.

In the present study, urea emerged as one of the most interesting molecules that showed to readily permeating RO membranes. Among the identified compounds, urea exhibits high similarity to water, which might also explain the high degree of RO membranes permeation. Other organic molecules were also found to permeate RO membrane. From the NMR-based identification and quantification, 30 compounds were detected in RO permeate (ROP) samples, mostly at low concentrations (below 7 µM). Among these compounds, lactose, lactic acid, citric acid and acetic acid were detected at concentrations 542 ± 125, 54.7 ± 19.8, 36 ± 9 and 44.5 ± 12.8 µM, respectively. Special attention is required for dimethylamine and creatine 1-phosphate, as they were not detected in either ROP or RO retentate (ROR) samples. Among the GC-MS based identified compounds, eight were found to permeate the RO membrane. These compounds include hydroxyisocaproic acid, pyruvic acid, itaconic acid, ribonic acid, pyroglutamic acid, glyceric acid, hydroxyisovaleric acid and succinic acid. Although the second RO filtration (RO polisher) proved to be efficient in purifying ROP, 21 compounds were detected in RO polisher permeate (ROPP) at concentrations below 5 µM. However, choline, acetic acid, carnitine and lactose were detected at higher concentrations, 26.4 ± 12.2, 18.6 ± 7.3, 59.9 ± 22.6 and 126 ± 12µM, respectively.

As expected, during water evaporation, volatile compounds such as urea, methanol, acetone, acetic acid and some fatty acid derivatives, namely methyl-hydroxybutyric acid, hydroxyisocaproic acid and hydroxyisovaleric acid, evaporate as

well. Two other compounds affected by evaporation were creatine and creatinine. A decrease of creatine with a concomitant increase of creatinine levels was observed in the evaporation concentrates (EVC). After crystallizing the EVC samples and removing lactose crystals, mother liquor (ML) was obtained. The composition of mother liquor varies depending on the whey permeate composition and the processing steps used during lactose production. To date, 91% of the average dry matter of mother liquor has been identified and quantified (Durham et al., 2022), and the present project expands the list of the identified and quantified compounds in mother liquor to 110 and 51, respectively. The composition of ML samples were similar to EVC samples with the only exception being lactose, which had lower concentration in the ML samples as observed in Fig. 2. In Fig. 2, the scores (a) and loadings (b) of the principal components analysis (PCA) model of the dataset obtained from the NMR analysis, where ML samples from batch one are closely related to EVC samples, possibly due to the presence of lactose fines.

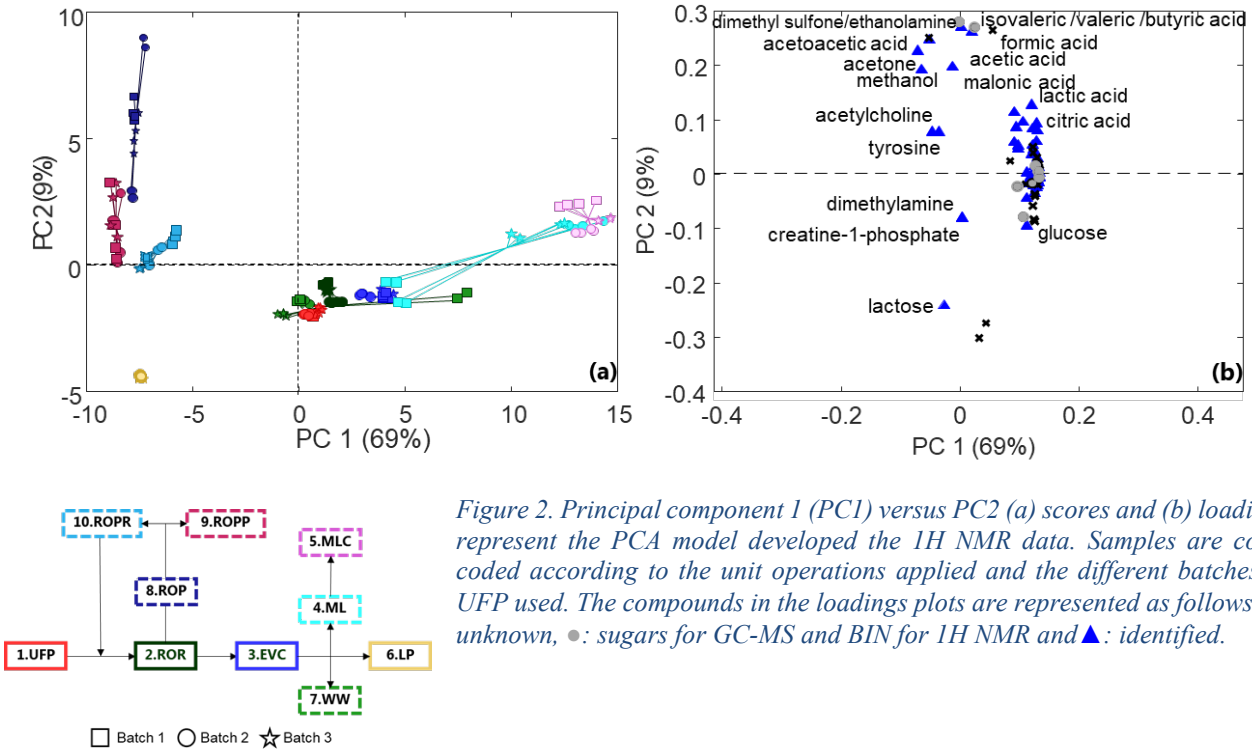



*Figure 2. Principal component 1 (PC1) versus PC2 (a) scores and (b) loadings represent the PCA model developed the 1H NMR data. Samples are color coded according to the unit operations applied and the different batches of UFP used. The compounds in the loadings plots are represented as follows: ×: unknown, ●: sugars for GC-MS and BIN for 1H NMR and ▲: identified.*

As far as it concerns the molecular profile of wash water (WW), it consists a diluted counterpart of mother liquor and reflects the effectiveness of lactose crystals washing. This is particularly evident in one of the replicates of batch one, which is closely related to ML samples in Fig. 2. The compositional data of WW samples revealed high levels of ML and lactose, which can be attributed to the generation of small lactose crystals, leading to challenges in separating the ML from the lactose crystals and resulting in losses of lactose fines.

Finally, in lactose powder (LP) some compounds were detected at low/trace amounts, while others were found in higher levels. Therefore, it was hypothesized that impurities found in LP samples can be there either because of (a) remaining/entrapped mother liquor in the lactose crystals or (b) affinity of certain compounds in whey permeate with lactose in solution and/or lactose crystals. Any remaining/entrapped mother liquor would be expected to have a small contribution to the total solids of lactose powder and thus, their concentration would be very small. Accordingly, adenine, phosphocholine, aspartic acid, creatinine, choline, orotic acid, creatine, acetic acid, carnitine, citric acid and lactic acid which were detected in very small amounts are expected to be embedded into lactose crystals as part of the entrapped mother liquor. On the other hand, galactose, glucose, hydroxypyruvic acid, malonic acid, gluconic acid and ribonic acid, which were found in lactose powder at high levels (their concentration in LP was similar to those in the EVC samples)

are probably interacting with lactose during crystallization. Until today, only the interaction of lactose phosphate, glucose 6-phosphate, and glucose with lactose has been investigate, which opens a space for the investigation of the interaction of these compounds with lactose during processing and crystallization.

One of the objectives of this project was to identify and quantify the non-protein nitrogen (NPN) fraction. Thirty-seven different NPN compounds were identified and quantified in UFP, lactose powder and process streams from the trial lactose productions. The percentage of the nitrogen recovered was calculated from the nitrogen of the quantified NPN relative to the NPN measured by Kjeldahl. The percentage of the quantified nitrogen per sample type is presented in Fig. 3. As expected the primary nitrogen contributor in NPN fraction is urea, with the only exception being lactose powder where urea was not detected. Following urea, amino acid derivatives (e.g. creatine, pyroglutamic acid), amino alcohols (e.g. choline), organic acids (e.g. orotic acid) and amino acids substantially contributed to nitrogen content. The total quantified nitrogen ranges from 57% to 99% depending on the sample type. Samples with high NPN content showed better nitrogen recovery, probably because most of the NPN compounds were above the detection and quantification limit of GC-MS and $^{1}$H NMR. The samples with the highest nitrogen recovery were ROR, MLC, WW and ROP, followed by EVC, ML, UFP and ROPP. Although ROP samples have low nitrogen content, the percentage of the quantified nitrogen was 99% due to the fact that 91% of the quantified nitrogen is urea. For lactose powder samples the recovery of nitrogen was as low as 57%. In samples containing 0.004 ± 0.0002% (w/w) nitrogen, it is expected that some compounds will be below the limit of detection and quantification of $^{1}$H NMR and GC–MS.

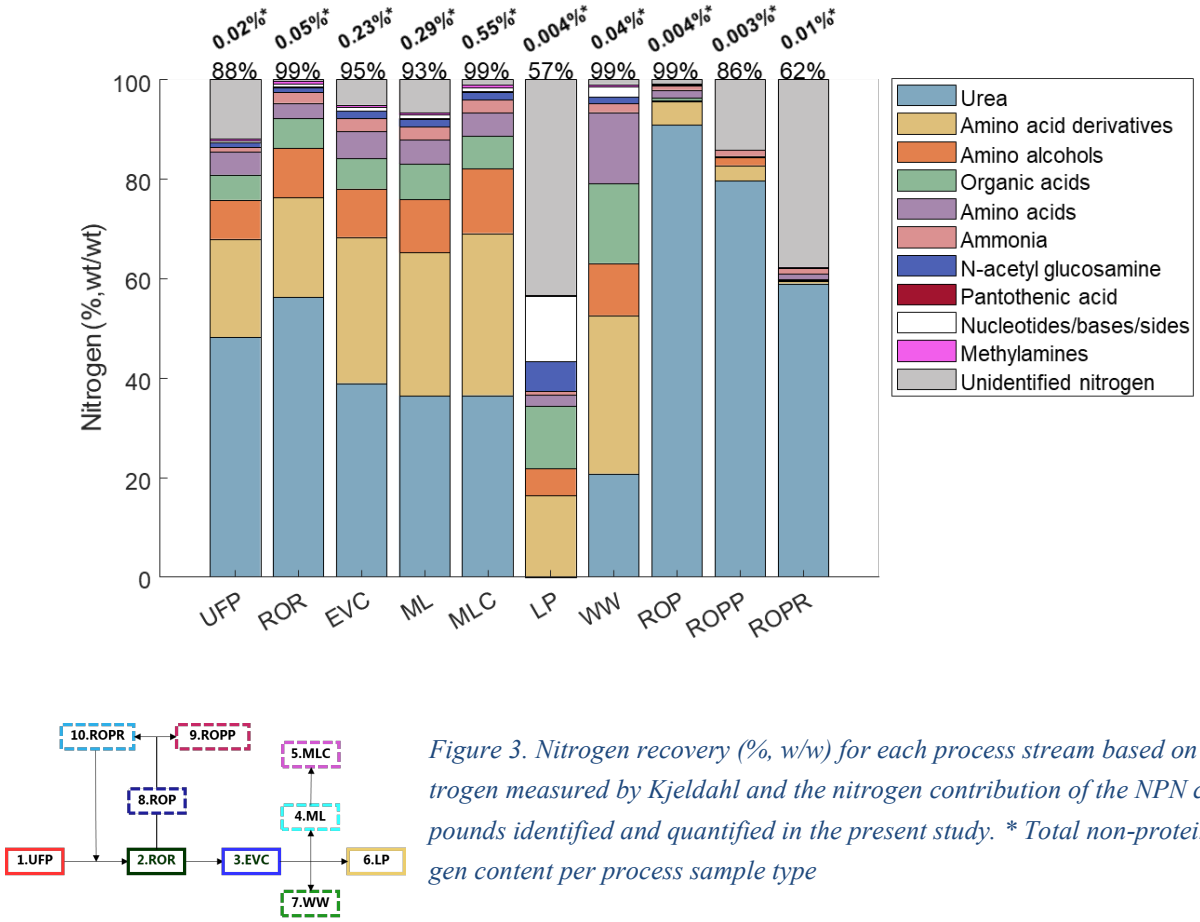






*Figure 3. Nitrogen recovery (%, w/w) for each process stream based on the nitrogen measured by Kjeldahl and the nitrogen contribution of the NPN compounds identified and quantified in the present study. * Total non-protein nitrogen content per process sample type*

As far as it concerns milk, a chemical shift library of 63 milk metabolites representing eight chemical classes was established and implemented in the open access Signature Mapping (SigMa) software, to facilitate rapid and unambiguous identification and quantification of the milk metabolome from $^{1}$H NMR spectra. SigMa is a spectral analysis tool for

converting complex milk $^{1}$H NMR spectra into an informative and quantitative metabolite table in a reproducible way across laboratories. The SigMa milk methodology was tested on whole milk, skim milk and ultrafiltered milk. Analysis of ultrafiltered milk using $^{1}$H NMR offers high metabolite coverage. Whole and skim milk represent the intact biochemical fingerprint of milk with proteins and lipids or without lipids, respectively. Thus, $^{1}$H NMR spectra of whole or skim milk can be used for fingerprinting purposes where the goal is to differentiate milks or to find unique biochemical characteristics. This study showed that 33 metabolites and 5 functional groups can be detected in whole milk, 53 metabolites in skim milk and 63 in ultrafiltered milk. However, not all of the detected metabolites can be quantified accurately. In whole milk it was possible to quantify three metabolites, while in skim milk 46 metabolites. The remaining metabolites were not quantitative due to masked and overlapping signals and, metabolites-protein complex formation or metabolites at very low concentrations in the samples. In particular, the low number of metabolites that can be accurately quantified in whole milk (i.e. carnitine, choline and lactose), limits its application in metabolomics studies. In ultrafiltered milk all 63 metabolites can be detected and quantified.

In this study, it was also demonstrated that metabolite quantification from the SigMa metabolite table using gallic acid as an internal standard has performance similar to the reference quantification method (i.e. standard addition calibration curve). This allows simple, swift and efficient quantification of large number of metabolites in big metabolomics studies. Finally, the performance of the methodology was demonstrated by analyzing cow colostrum from 88 cows subjected to four different feeding regimes. The experiment revealed that SigMa is suitable for untargeted metabolomics by converting the spectral data to one metabolite table facilitating their interpretation.

## 12. The relevance of the results, including relevance for the dairy industry

Increasingly demanding consumers are concerned about the traceability, origin and nutritional information of dairy products. Although advanced technologies have already been applied in dairy science, a fraction of low-molecular weight compounds in milk still remains unknown, and little information is available about the detailed composition of the processing side-streams like whey permeate. This project helped towards mapping the low-molecular weight compounds that end-up in milk and whey permeate streams with a view to offer better documentation of dairy products such as lactose powder.

From the dairy industry perspective, a current trend is by-products recovery and utilization. In this context, sustainable management of dairy streams, limited waste and optimized use of by-products should be taken into account. The complete molecular profile of dairy matrices can augment the exploitation of their low-molecular-weight fraction. More specifically, from the results of the present project the dairy industry can utilize the molecular composition of wash water, lactose powder and mother liquor to further control lactose crystallization and secure increased yield of crystals with minimum loss in mother liquor and during washing. The detailed composition of underutilized side-streams, such as mother liquor concentrate, may help their valorization and the development of new ingredients/products. Finally, knowledge about the dynamics of permeate compounds during lactose production can help to manipulate operating conditions for increased lactose powder yield and quality.

Although low-molecular weight compounds in the dairy supply chain have been largely overlooked, their identification, and to a lesser extent their quantification, has gained increasing attention in the last twenty years. Key compounds have been identified as potential biomarkers for predicting cows' health and milk quality and technological properties. By extending the molecular profiling of dairy matrices throughout the entire dairy production chain, from raw milk to end products and side-streams, a better understanding of the impact of biological and technological factors on the dynamics of low-molecular weight compounds in dairy matrices will be achieved. This, in turn, can be correlated with the quality and processing performance of dairy matrices. Overall, such practices will contribute to better compositional documentation of dairy products and ingredients and will pave the road for the valorization of underutilized process streams. In

addition, the documented composition of dairy matrices has the potential to act as a decision tool for the selection of raw materials for the production of targeted products/ingredients.

## 13. Communication and knowledge sharing about the project

**Peer-reviewed papers in international journals:**
Tsermoula, P.; Khakimov, B.; Nielsen, J.H.; Engelsen, S.B. (2021). WHEY—The waste-stream that became more valuable than the food product. *Trends in Food Science and Technology,* 118, pp. 230-241.

Paraskevi Tsermoula, Mie Rostved Bechshøft, Christoffer Friis, Søren Balling Engelsen and Bekzod Khakimov, (2023). Screening of non-protein nitrogen compounds in lactose refining streams from industrial whey permeate processing. Food Chemistry, 405, 134716.

Paraskevi Tsermoula, Mie Rostved Bechshøft, Christoffer Friis, Søren Balling Engelsen and Bekzod Khakimov, (2023). Molecular profiling of whey permeate reveals new insights into molecular affinities related to industrial unit operations during lactose production. Food Chemistry, 420, 136060.

Tsermoula, P., Kristensen, N. B., Mobaraki, N., Engelsen, S. B. and Khakimov, K. (2024). Efficient Quantification of Milk Metabolites from $^{1}$H NMR Spectra using the Signature Mapping (SigMa) approach: Chemical shift library development for cows' milk and colostrum. Analytical Chemistry, 96, 1861-1871.

**Easily read papers:**
Bekzod Khakimov, Paraskevi Tsermoula, Kristine Blans, Mie Rostved Bechshøft, Jacob Holm Nielsen, Søren Balling Engelsen (2020). The dark matter of whey: Low molecular weight compounds in milk and dairy streams - a potential new source for value added products (MilkStreamValue). Mælkeritidende, 4, 4-5.

Paraskevi Tsermoula, Mie Rostved Bechshøft, Christoffer Friis, Bekzod Khakimov, Søren Balling Engelsen (2025). Low molecular weight compounds in lactose refining streams from whey - a potential new source for value added products (MilkStreamValue). Mælkeritidende, in press.

**Student theses:**
PhD, Paraskevi Tsermoula. Screening and valorization of minor components in refined milk fractions (defended June 2nd 2023).

**Oral presentations at scientific conferences, symposiums etc.:**
Paraskevi Tsermoula. SigMa Milk: Quantification of metabolites in milk and milk products from $^{1}$H NMR spectra using SigMa. MRFOOD2022 - 15th International Conference on the Applications of Magnetic Resonance in Food Science. Aarhus, Denmark, 8 June 2022.

**Oral presentations at meetings:**
Danish Dairy Research Foundation (DDRF) Coordination group "Technology & Safety", online. MilkStreamValue - Introduction. Tsermoula, P., Khakimov, B., Engelsen, S., November 2020.

Danish Dairy Research Foundation (DDRF) Coordination group "Technology & Safety". MilkStreamValue – Status report. Tsermoula, P., Khakimov, B., Engelsen, S., May 2021.

Danish Dairy Research Foundation (DDRF) Coordination group "Technology & Safety". MilkStreamValue – Final report. Tsermoula, P., Khakimov, B., Engelsen, S., May 2023.

**Other:**
FOOD Research day, Copenhagen, Denmark. The dark matter of whey. Tsermoula, P., Khakimov, B., Engelsen, S., November 2020.

## 14. Contribution to master and PhD education

The project has contributed to the education and PhD thesis of Paraskevi Tsermoula: Screening and valorization of minor components in refined milk fractions (defended June 2nd 2023; University of Copenhagen).

## 15. New contacts/projects

The establishment and demonstration of the SigMa methodology to whole milk, skim milk and ultrafiltered milk led to a collaboration with SEGES innovation as there was a need for milk samples which would have apparent compositional differences with commercial milk. This let us to demonstrate that the method is robust and reproducible. Therefore, the method was tested on ultrafiltered colostrum samples from dairy cows (n=88) to evaluate whether metabolic changes in colostrum may reflect the metabolic status of cows. The results obtained were promising in predicting the metabolic changes in colostrum of periparturient cows given different feedings and this was upscaled into a one year postdoc project.