Waters™

Applikationsbericht

Automated Sample Preparation Using Andrew+ Pipetting Robot for UPLC-MS^E Identification and UPLC-MS/MS Quantification of Bovine Milk Proteins

Li Yan Chan, Adabelle Ong

Waters Corporation

Abstract

Milk proteins of high biological and nutritional value show technological functionalities when applied to food products. A full suite of solutions were developed to identify and quantify two classifications of bovine milk proteins, casein, and whey. The ProteinWorks Auto-eXpress Digest Kit is a flexible and broadly applicable sample preparation kit containing pre-weighed, lot traceable reagents optimized for users to achieve an accurate, precise, and robust LC-MS quantification of proteins via the surrogate peptide approach. The ACQUITY Premier Column used in this application is based on the MaxPeak High Performance Surfaces (HPS) Technology that provides a highly effective surface barrier that mitigates undesired interactions with metal surfaces and improves the analysis of phosphopeptides found in casein. Progenesis QI for Proteomics workflow facilitates the discovery of signature marker peptides, which is required for subsequent UPLC-MS/MS quantification of the milk proteins. Automating the approach of generating calibration standards and stepwise digestion protocol using the Andrew+ Pipetting Robot, with cloud-native OneLab Software significantly reduces the time taken to manually pipette and allows the lab analyst to perform other tasks simultaneously, thus maximizing productivity.

Benefits

- Improve speed and completeness in digestion – accelerated tryptic digestion of proteins using *Rapi*Gest SF Surfactant
- No missing peptides – enhanced phosphopeptide detection with MaxPeak High Performance Surfaces Technology
- Maximizing productivity – ease-of-use and time saving automated protocol with less repetitive pipetting

Introduction

Milk is recognized to have health benefits and contains valuable nutrients. It is also used to produce dairy products such as cheese, butter, and yogurt. The important components of milk include proteins, lipids, fatty acids, and carbohydrates, and their content differs in milk derived from different animal species.[1] Increased interest in dairy alternatives stem from many drivers, including the avoidance of dairy allergens; desire for clean label products; compatibility with vegetarian, vegan, and flexitarian lifestyles; and concerns about carbon footprint, sustainability and animal welfare. It is important to ensure such milk alternatives have the required nutritional content.

Amongst the milk components, proteins are of particular interest as research suggests that they provide primary nutrients and protection against infections via antimicrobial and immune-modulatory activities. They are also known to play a crucial role in the flavor of the dairy products as they have an impact on the mouthfeel and texture which are part of the sensory experience. However, the same sensory attributes and functionalities found in dairy milk proteins are still limited and less desirable in alternative dairy proteins, and therefore pose as a challenge.[2] Current development in dairy alternatives include cell-based dairy products or isolated dairy ingredients like vegan casein where there is a demand to characterize new matrices and properties that affect product development for food technology.

There are numerous studies on milk proteins and the existing analytical methods include gel electrophoresis, capillary electrophoresis, liquid chromatography, and immunological techniques. However, these methods may have some inaccuracies, lower reproducibility, and sample processing procedures are laborious and time-consuming. The development of a robust and reliable quantitative method for these proteins is therefore

beneficial to both research and manufacturing applications[3] and can also be applied to dairy alternative proteins. The bovine milk proteins selected below were used to demonstrate proof of performance for transfer of the method to characterize other alternative proteins.

The objective of this application note, was to determine the concentration of five bovine milk proteins of interest, α-casein (α-CN), β-casein (β-CN), κ-casein (κ-CN), α-lactalbumin (α-LA), and β-lactoglobulin (β-LG) from liquid milk samples. The solution includes (i) tryptic digestion using the ProteinWorks Auto-eXpress Digest Kit; (ii) bottom-up proteomics approach to discover signature peptides for each protein; (iii) UPLC-MS/MS method in multiple reaction monitoring (MRM) mode to quantify the five milk proteins; and (iv) automated protocol using the Andrew+ Pipetting Robot to streamline the process.

Experimental

Sample Preparation

Five stock solutions of α-CN, β-CN, κ-CN, α-LA, and β-LG at a concentration of 10 mg/mL were individually prepared by dissolving the standards in digestion buffer obtained from the ProteinWorks Auto-eXpress Low Digest Kit (p/n: 176004078 <https://www.waters.com/nextgen/us/en/shop/application-kits/176004078-proteinworks-auto-express-low-5-digest-kit.html>). An appropriate volume of the individual stock solution was aliquoted and combined to give a mixed standards solution with initial concentrations of 3,500 µg/mL for α-CN and β-CN, 1,000 µg/mL for κ-CN and β-LG, and 500 µg/mL for α-LA. A serial dilution was then performed by the Andrew+ Pipetting Robot to give a series of the remaining calibration standards. Five commercial milk samples were first diluted ten times with the digestion buffer before proceeding with the digestion protocol. Using 28 µL of the calibration standards or diluted milk samples, the ProteinWorks Auto-eXpress Low Digest Kit, 5-step digestion protocol was followed. Both the calibration series preparation and digestion protocol were fully executed by the Andrew+ Pipetting Robot.

LC Conditions

LC System:	ACQUITY UPLC I-Class PLUS FTN

Column:	ACQUITY Premier CSH C_{18} Column 1.7 µm, 2.1 x 100 mm (p/n: 186009461)
Column temp.:	55 °C
Sample temp.:	4 °C
Injection volume:	1 µL
Flow rate:	0.15 mL/min
Mobile phase A:	0.1% formic acid in water
Mobile phase B:	0.1% formic acid in acetonitrile
Gradient: (UPLC-QTof-MS^E)	Initial to 1 min: 1% B; 1 to 15 min: 1 to 40% B; 15 to 15.5 min: 40 to 90% B; 15.5 to 17 min: 90% B; 17 to 17.1 min: 90 to 1% B; 17.1 to 20 min: 1% B
Gradient: (UPLC-MS/MS)	Initial to 1 min: 5% B; 1 to 7.5 min: 5 to 50% B; 7.5 to 8 min: 50 to 90% B; 8 to 9 min: 90% B; 9 to 9.1 min: 90 to 5% B; 9.1 to 11 min: 5% B

MS Conditions for UPLC-QTof-MS^E

MS system:	Xevo G2-XS QTof
Ionization mode:	ESI+
Acquisition mode:	Tof MS^E
Acquisition range:	50–2000 Da
Capillary voltage:	2.5 kV
Collision energy ramp:	15–40 eV
Cone voltage:	40 V
Lockmass:	Glu Fibrinopeptide B (2+, *m/z* 785.8426)

MS Conditions for UPLC-MS/MS

MS system:	Xevo TQ-S micro
Ionization mode:	ESI+
Capillary voltage:	2.2 kV
Cone voltage:	30 V
Source temp.:	120 °C
Desolvation temp.:	600 °C
Desolvation gas flow:	1000 L/hr

Cone gas flow:	50 L/hr

Data Management

Chromatography software:	MassLynx v4.2
MS software:	MassLynx v4.2
Informatics:	TargetLynx v4.2 Progenesis QI for Proteomics v4.0 EZInfo v3.0 OneLab

Results and Discussion

The use of *Rapi*Gest SF Surfactant improves the speed and completeness of enzymatic digestions with minimal or no post-digestion sample preparation, and therefore simplifies the preparation protocols (p/n: 720003102 < https://www.waters.com/nextgen/us/en/library/application-notes/2009/in-solution-enzymatic-protein-digestions-with-rapiGest.html>). Triplicates of five bovine milk protein standards were first individually digested and subjected to UPLC-QTof-MS^E analysis. The raw data from the pooled QC sample was then imported into Progenesis QI for Proteomics, which performed chromatographic alignment, data normalization, and peak picking. A total of 1,050 peptides were identified using the Ion Accounting identification workflow with the settings: FDR less than 1%, fixed modification (carbamidomethylation of C), and variable modifications (oxidation of M and phosphoryl of STY). The UniProt FASTA file loaded contained database that was filtered to include bovine reviewed sequence only.

The principal components analysis (PCA) is one of the key features present in Progenesis QI for Proteomics to allow clear visualization of the data in order to determine the presence of any outliers, and also how well the samples are grouped. The unsupervised PCA score plot in Figure 1 shows clear clustering of the five proteins. All

pooled QC samples were found to be tightly located and close to the center of the PCA score plot which demonstrated good reproducibility of the method and the absence of any bias introduced during the processing of the data. The clusters for β-CN and κ-CN were found closely together, indicating that the differences between the two caseins were minimal.

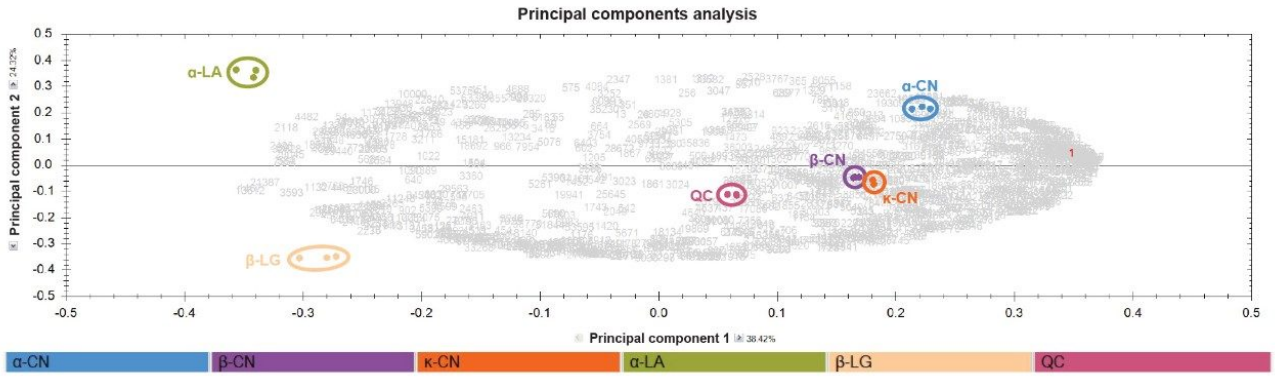

Figure 1. PCA score plot of five milk proteins from Progenesis QI for Proteomics.

The peptide data was exported to EZInfo, whereby a supervised multivariate analysis was performed using orthogonal partial least square discriminant analysis (OPLS-DA). OPLS-DA is a useful statistical tool to help identify signature marker peptides. Figure 2A shows a representative OPLS-DA score plot of β-CN versus the other four proteins. The S-plot was created (Figure 2B) to highlight those characteristic features responsible for the differences found in β-CN from the four other proteins. The features with the highest confidence and importance (highlighted in blue) were selected and imported back into Progenesis QI for Proteomics, for verification and further evaluation of identity (Figure 2C). The selected peptides in this example, are unique peptide sequences of β-CN, and are not found in the other four proteins. The process was repeated for all five milk proteins and a list of signature marker peptides was shortlisted. For each protein, a peptide was then chosen for quantification, and at least one more peptide was monitored for confirmatory purposes (Table 1). The selection criteria for the peptides included good chromatographic resolution, signal intensity and signal-to-noise ratio.

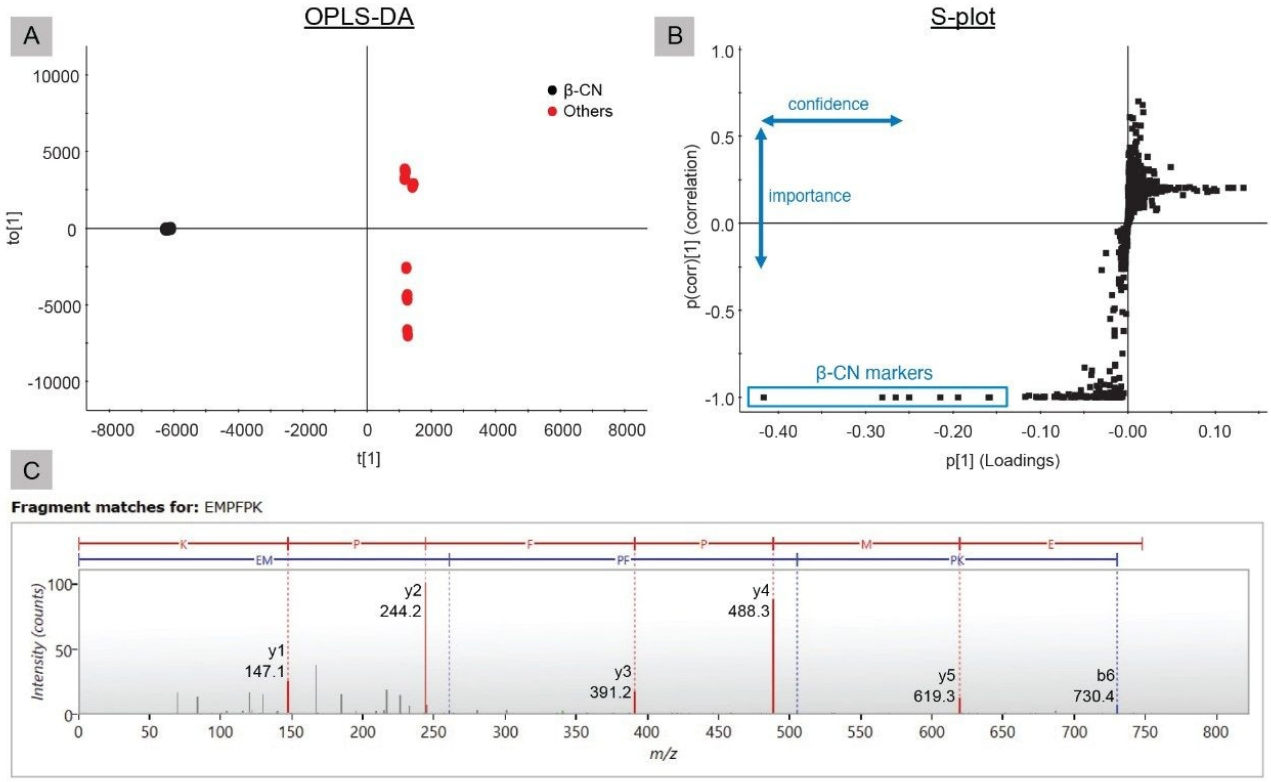

Figure 2. A. OPLS-DA score plot; B. S-plot of β-CN versus the other four proteins; and C. Identification example of the peptide EMPFPK.

Protein	Peptide	MRM	Collision (eV)
α-Casein (α-CN)	HQGLPQEVLNENLLR	587.3>758.4	20
		587.3>790.4	20
		587.3>889.5	20
	YKVPQLEIVPNSAEER[a]	651.3>784.4	20
		651.3>882.3	20
		651.3>971.6	20
	FFVAPFPEVFGK	692.9>920.5	20
		692.9>991.5	20
		692.9>1090.6	20
β-Casein (β-CN)	EMPFPK[a]	374.7>244.2	10
		374.7>391.2	10
		374.7>488.3	10
	VLPVPQK	390.8>275.2	20
		390.8>372.2	20
		390.8>568.3	10
κ-Casein (κ-CN)	YIPIQYVLSR[a]	626.4>637.4	20
		626.4>765.4	20
		626.4>975.6	20
	SPAQILQWQVLSNTVPAK	660.7>315.2	20
		660.7>716.4	20
		660.7>738.4	20
α-Lactalbumin (α-LA)	CEVFR[a]	355.8>175.4	20
		355.8>322.4	20
		355.8>421.5	20
	VGINYWLAHK	400.9>468.3	20
		400.9>654.4	20
		400.9>817.4	20
β-Lactoglobulin (β-LG)	VLVLDTDYKK[a]	398.6>654.3	20
		398.6>769.4	20
		398.6>882.5	20
	ALPMHIR	419.2>327.2	20
		419.2>425.3	20
		419.2>653.4	20

Table 1. MRM conditions of the signature peptides for the respective milk proteins. [a]Peptide chosen for

quantification.

The MRM conditions for the selected signature peptides are summarized in Table 1. All the MRM transitions with a minimum signal-to-noise ratio of three must be present for a peptide to be detected, while at least two peptides must be determined for protein assignment.[4] In the previously reported application note (p/n: 720007025 <https://www.waters.com/nextgen/us/en/library/application-notes/2020/maximizing-phosphopeptide-recovery-in-lc-ms-studies-with-maxpeak-high-performance-surfaces-technology.html>), the benefits of using MaxPeak High Performance Surfaces (HPS) Technology in the analysis of phosphopeptides was described with improvement in the detection of the peptide YKVPQLEIVPNSAEER.

The preparation of calibration standards is a requirement in all quantitative analysis, but these steps are repetitive, time-consuming, and prone to human error. Similarly, the protein digestion protocol requires several repetitive additions of reagents, making both ideal for automating the task. The calibration curve preparation and the ProteinWorks 5-Step Digestion Protocol were generated in OneLab, the cloud-native software used for both designing and executing protocols on the Andrew+ Pipetting Robot (Figure 3).

Figure 3. Andrew+ Deck Layout for sample preparation of standards and samples using Andrew+ Pipettes and Domino Blocks: 1. Tip insertion system (10–300 µL Optifit tips); 2. Tip insertion system (0.1–10 µl Optifit tips); 3. Microplate; 4. Storage plate; 5. Microtube; 6. 8-channel pipette reservoir; and 7. 96-PCR Plate Peltier+.

Comparing the pipetting hours (heating hours during digestion are excluded) required between Andrew+ Pipetting Robot and a typical lab analyst, Andrew+ was able to complete the pipetting task at a rate >5.5 times faster (Figure 4). Automating the process also allows the lab analyst to perform other tasks simultaneously, streamlines the sample preparation procedure, reduces the potential of human error, and ensures consistent analytical method performance throughout the process.

Figure 4. Comparing pipetting hours required for Andrew+ and manual pipetting.

The standard calibration curve created by Andrew+ Pipetting Robot shows excellent linear correlation coefficient of r^2 >0.99 using simple 1/x weighting and covering a range of concentration (Table 2).

Peptide marker	Calibration curve	Linear fit (r^2)	Range (µg/mL)
α-CN (YKVPQLEIVPNSAEER)	44.564x + 353.050	0.996	35–3,500
β-CN (EMPFPK)	211.776x + 283.947	0.998	35–3,500
κ-CN (YIPIQYVLSR)	318.841x + 2063.050	0.994	10–1,000
α-LA (CEVFR)	1181.280x + 940.877	0.999	5–500
β-LG (VLVLDTDYKK)	27.041x – 146.845	0.997	10–1,000

Table 2. Calibration curves for the respective five proteins.

In order to provide a representative sample set, a range of commercial milk products including, full cream, low-fat, chocolate flavored, and strawberry flavored milk, were analyzed using the semi-automated method developed. Figure 5 shows the different concentrations of identified milk proteins found in the various milk samples. The results obtained are in good agreement to previously reported literature.[5] Among the five proteins of interest, we observed that β-CN has the highest concentration (41–49%) among all the milk samples, followed by α-CN (27–34%), κ-CN (11–15%), β-LG (10–12%), and α-LA (3%). For absolute quantification analysis, isotope

labelled peptides as internal standards should be introduced. Further method validation, such as matrix effect and recovery are also recommended for testing.

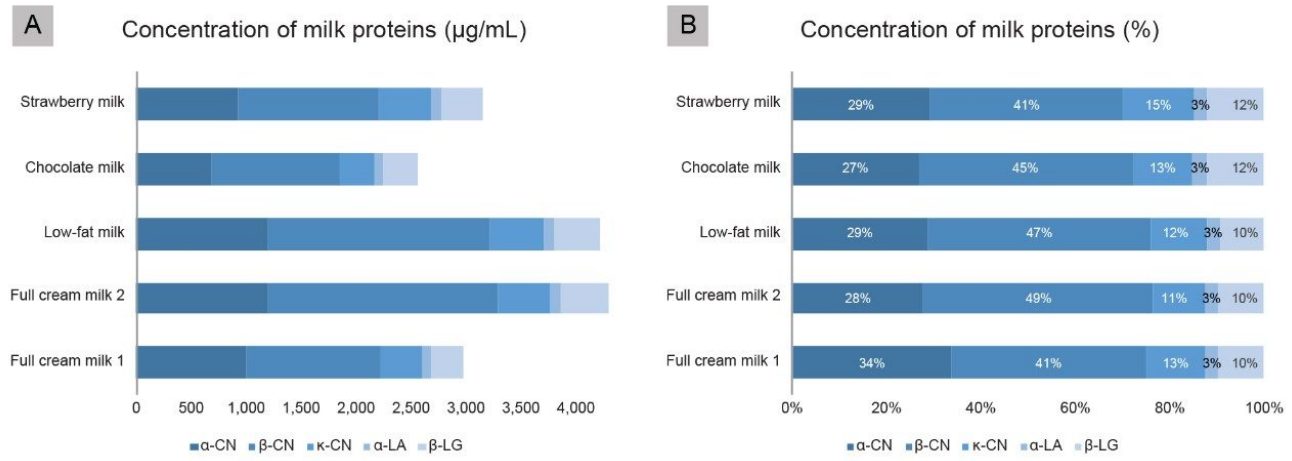

Figure 5. Concentration of the five proteins of interest in commercial milk products expressed as A. µg/mL; and B. percentage.

Conclusion

Bottom-up proteomics is a useful approach for the discovery of signature peptides and has been demonstrated to work for bovine milk proteins. Progenesis QI for Proteomics, together with the statistical tools in EZInfo, has streamlined the challenging procedures into an easy-to-use workflow. The signature marker peptides discovered were successfully developed into a quantification analysis, and a range of commercial milk products were analyzed to demonstrate the performance and effectiveness of this method. The automated method created with OneLab, and performed on the Andrew+ Pipetting Robot, greatly reduced the pipetting hours by >5.5 times, while ensuring consistency. This application note illustrated a full suite of solutions and the same workflow can be applied to analyze alternative milk proteins.

References

1. Lucey J, Otter D, Horne D. A 100-Year Review: Progress on The Chemistry of Milk and Its Components. *J Dairy Sci.* (2017), 100: 9916–9932.
2. McClements D, Newman E, McClements I. Plant-based Milks: A Review of the Science Underpinning Their Design, Fabrication, and Performance. *Comprehensive Reviews in Food Science and Food Safety* (2019), 18: 2047–2067.
3. Tremblay L, Laporte M, Léonil J, Dupont D, Paquin P. (2003) Quantitation of Proteins in Milk and Milk Products. In: Fox P, McSweeney P. (eds) Advanced Dairy Chemistry—1 Proteins. Springer, Boston, MA; 49–138.
4. Song E, Gao Y, Wu C, Shi T, Nie S, Fillmore T L, Schepmoes A A, Gritsenko M A, Qian W, Smith R D, Rodland K D, Liu T. Targeted Proteomic Assays For Quantitation of Proteins Identified by Proteogenomic Analysis of Ovarian Cancer. *Sci. Data* (2017), 1–13.
5. Bär C, Mathis D, Neuhaus P, Dürr D, Bisig W, Egger L, Portmann R. Protein Profile of Dairy Products: Simultaneous Quantification of Twenty Bovine Milk Proteins. *Int. Dairy J.* (2019), 97: 167–175.

Acknowledgements

This work was conducted at Waters International Food and Water Research Centre (IFWRC, Singapore). The authors would like to thank Turtletree Labs Pte. Ltd. for the protein standards and expertise in the area of novel milk.

Featured Products

ACQUITY UPLC I-Class PLUS System <https://www.waters.com/134613317>

Xevo G2-XS QTof Quadrupole Time-of-Flight Mass Spectrometry <https://www.waters.com/134798222>

Xevo TQ-S micro Triple Quadrupole Mass Spectrometry <https://www.waters.com/134798856>

Automated Liquid Handling for LC & LC-MS Sample Preparation Workflows <https://www.waters.com/waters/nav.htm?cid=135070059>

Progenesis QI for Proteomics <https://www.waters.com/waters/nav.htm?cid=134790665>

720007335, August 2021