BOTTOM-UP AND TOP-DOWN ANALYSES OF MILK PROTEINS USING LC-MS/MS

Delphine Vincent [1], Aaron Elkins [1], Mark Condina [2], Benjamin G Cocks [1 3], Simone Rochfort [1 3]

1. DEDJTR, Biosciences Research Division, AgriBio, Bundoora, VIC, Australia
2. Bruker Pty Ltd, Preston, VIC, Australia
3. School of Applied Systems Biology, La Trobe University, Bundoora, VIC, Australia

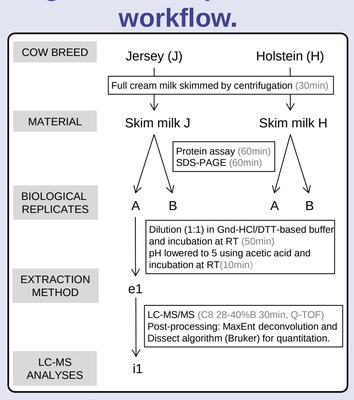
TOP-DOWN (TD)

TD INTRODUCTION

The development of robust analytical methods for the separation and quantification of variants of milk proteins is of interest.

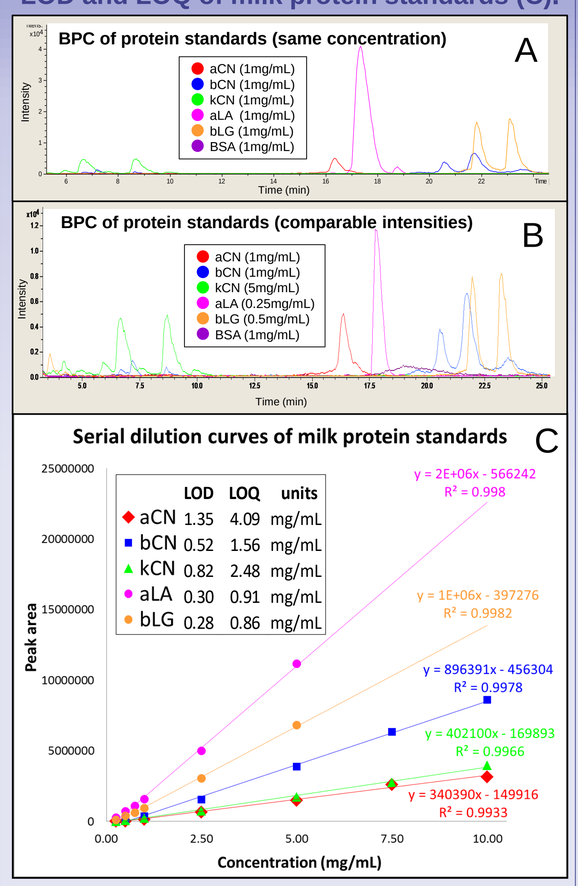
This work aims at establishing a LC-MS-based method to study intact major proteins (caseins and whey) from cow's skim milk. High resolution spectra subsequently allowed accurate mass determination and top-down sequencing (TDS) of major proteins. This LC-MS workflow enables a high-throughput profiling of abundant proteins of cow's milk with minimal sample preparation.

Figure TD1: Experimental workflow.



Skim milk samples were prepared according to Bobe *et al*. (J. Agric. Food Chem. 1998, 46, 458-463).

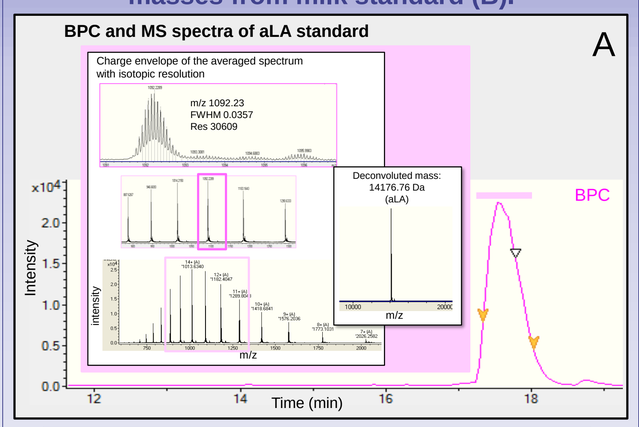
Figure TD2: Ionisation efficiency (A-B), and LOD and LOQ of milk protein standards (C).



The LC-MS analysis of milk protein standards showed little co-elution (A, B). The ionisation efficiency varied across standards (A), yet it was not concentration-dependent (B).

The limit of detection (LOD) ranged from 0.28 (bLG) to 1.35 (ACN) mg/mL and the limit of quantitation (LOQ) ranged from 0.86 (bLG) to 4.09 (aCN) mg/mL (C).

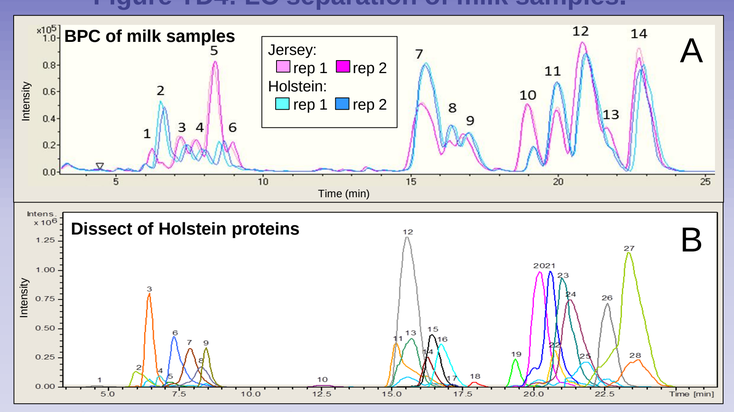
Figure TD3: Mass spectrum deconvolution of aLA elution peak (A) and list of accurate masses from milk standard (B).



Standard	RT [min]	Area	Intensity	S/N	Max. m/z	Observed Monoisotopic Mass	Theoretical Monoisotopic Mass	Δ mass	ID
aCN	7.48	136791	4604	52.1	870.897	25212.977	25213.343	0.367	aS2CN A
aCN	7.98	108967	3007	33.2	938.301	25292.957	25293.343	0.387	aS2CN A+1P
aCN	15.20	176712	9394	108.0	942.344	23520.243			-
aCN	15.65	2010064	51802	601.3	984.932	23600.429	23600.472	0.043	aS1CN B
aCN	16.44	611913	17906	75.2	1031.197	23680.218	23680.472	0.254	aS1CN B+1P
bCN	6.61	167481	7147	22.7	1002.930	19025.524			kCN A
bCN	8.57	159037	7902	25.0	1118.968	18993.569			kCN B
bCN	19.11	476421	18201	59.1	1048.456	24077.247	24077.386	0.139	bCN B
bCN	20.57	2004618	54506	178.4	1001.940	24008.203	24008.317	0.114	bCN A1
bCN	21.70	4148764	79791	261.5	1043.719	23968.216	23968.311	0.095	bCN A2
bCN	23.51	203175	3893	11.9	983.861	-			-
kCN	5.77	382815	16133	50.4	955.0662	11308.799			-
kCN	6.52	1285882	44873	172.1	1002.972	19026.517	19026.542	0.025	kCN A
kCN	7.47	294608	10048	36.1	1074.712	16096.108			-
kCN	8.53	1486543	42681	164.5	1001.299	18994.565	18994.589	0.024	kCN B
kCN	9.82	334147	10647	39.6	1072.556	16064.145			-
aLA	16.20	3533.6	394	3.4	1083.603	-			-
aLA	17.34	1517537	41053	498.5	1014.279	14176.763	14176.798	0.035	aLA
aLA	18.75	37368.5	2232	13.8	0.000	14158.755			-
bLG	21.58	28242	2014	20.1	1017.322	18286.360	18285.440	-0.920	bLG J
bLG	21.78	417689	16710	212.7	1016.593	18269.384	18269.409	0.025	bLG B
bLG	23.11	474983	16838	101.9	1021.388	18356.424	18355.446	-0.978	bLG A
BSA	19.49	-	992	-	696.853	-	-	-	BSA

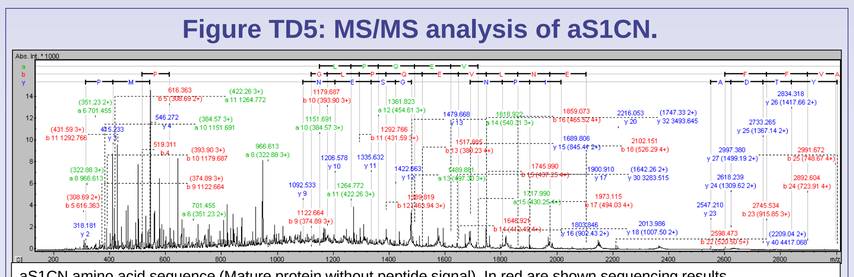
Mass spectra were isotopically resolved (A) which allowed for accurate monoisotopic mass determination by MaxEnt deconvolution followed by SNAP annotation of the standards. Each standard eluted several peaks which were successfully deconvoluted into variants and quantified (B).

Figure TD4: LC separation of milk samples.



Milk samples were reproducible and produce 14 chromatographic peaks (A) containing 28 overlapping proteins (B) whose abundance varied across breeds.

Figure TD5: MS/MS analysis of aS1CN.



aS1CN amino acid sequence (Mature protein without peptide signal). In red are shown sequencing results.

RPKHPIKHQGLPQEVLNENLLRFFVAPFPEVFGKEKVNELSKDIGSESTEDQAMEDIKQMEAESISSSEEIVPNSVEQKHIQKEDVPSERYLGYLEQLLRLKKY
KVPQLEIVPNSAEERLHSMKEGIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWYYVPLGTQYTDAPSFSDIPNPIGSENSEKTTMPLW

MS/MS fragmentation using in-source cell-induced-collision (ISCID) was successfully obtained on standards, such as aCN and milk proteins. Resulting a-, y- and b- ion series obtained from IsCID spectra were annotated against the AA sequence using BioTools software (Bruker), as illustrated here for aCN.

TD CONCLUSIONS

In this top-down study, we were able to perform quantitative MS and MS/MS analyses of proteins from Sigma standards and cow's milk samples. MS spectra were of high resolution and accurate protein masses were obtained, thus identifying milk protein isoforms. Top-down sequencing results achieved partial amino acid coverage of C- and N-termini for protein isoform characterisation. This method will be applied in a high-throughput fashion to hundreds of milk samples. Protein variants and their quantities will be related to milk properties and cow phenotype/genotype, and incorporated to dairy cow breeding programs.

ABSTRACT

Cow's milk is a complex fluid whose proteome displays a diverse set of proteins of high abundance such as caseins (CNs) and medium to low abundance whey proteins such as beta-lactoglobulin (bLG), alpha-lactabumin (aLA), lactoferrin, immunoglobulins, glycoproteins, peptide hormones and enzymes. CNs represent 80% of milk proteins and possess important nutritional and functional properties, whilst bLG (largely unknown function) and aLA (involved in lactose biosynthesis) represents approximately 10 and 4%, respectively. As genetic variants of CNs and whey proteins and post-translational modifications (PTMs) influence many properties of milk that are essential to the dairy industry.

This work had two purposes:

1/ Optimising a bottom-up (BU) proteomics procedure to identify as many proteins as possible in a high-throughput fashion.

2/ Establishing a top-down (TD) proteomics method to study intact major proteins, along with the identification of post-translational modifications (PTMs).

Applying these bottom-up and top-down methods on bulk milk samples from Jersey and Holstein cows, we demonstrate the complementarity of both approaches.

BU INTRODUCTION

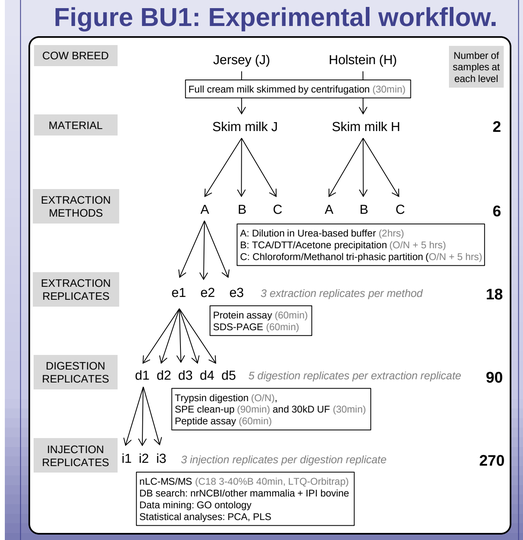
A sample preparation method that enables high reproducibility and throughput is key in reliably identifying proteins present in milk.

Using skim milk samples from Jersey and Holstein cows, we compared three extraction procedures which have not previously been applied to samples of cows' milk.

Method A (urea) involved a simple dilution of the milk in a urea-based buffer, method B (TCA/acetone) involved a TCA/acetone precipitation and method C (methanol/chloroform) involved a tri-phasic partition method in chloroform/methanol solution.

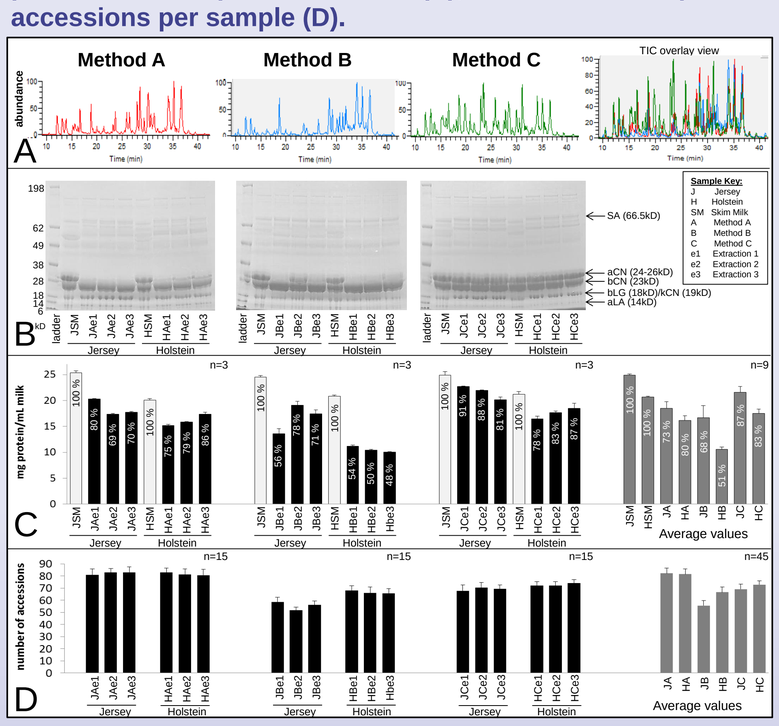
Protein assays, SDS-PAGE profiling, and trypsin digestion followed by nanoHPLC-electrospray ionisation-tandem mass spectrometry (nLC-ESI-MS/MS) analyses were performed to assess their efficiency.

Figure BU1: Experimental workflow.



Technical replicates were used during the extraction, digestion, as well as randomised injection steps to assess the reproducibility.

Figure BU2: Comparison of TICs (A), SDS-PAGE patterns (B), protein amount per mL of milk (C), and number of protein accessions per sample (D).



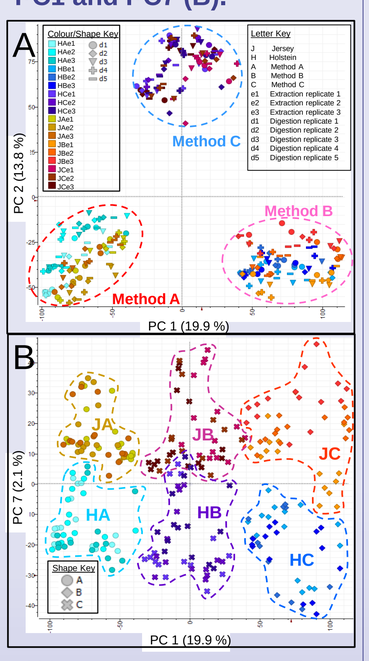
The 3 methods produced distinct TICs (A), with method C displaying peaks with higher resolution than methods A and B.

SDS-PAGE patterns (B) were similar from one extraction method to another, albeit extracts C displayed the best resolution with the sharpest bands. In particular extracts C were the only ones consistently resolving the very intense 24-26 kD band corresponding to alpha-caseins, and were therefore more comparable to skim milk profiles than extracts A and B.

Method C yielded the highest protein concentrations (C) substantiated by the highest recovery rate, followed by method A, while method B resulted in the lowest concentrations particularly for Holstein breed.

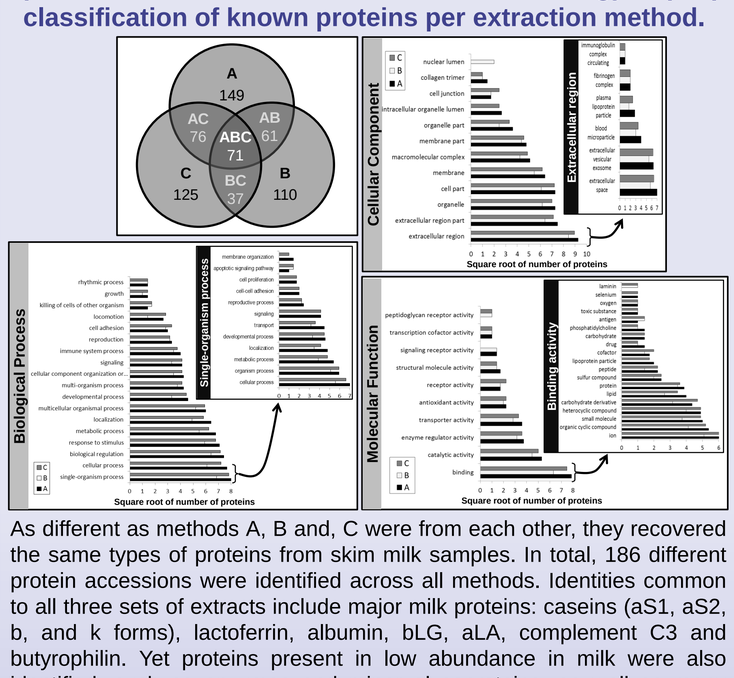
Extracts C generated less unique IDs (D) than extracts A, albeit more than B.

Figure BU3: PCA plots along PC1 and PC2 (A), and PC1 and PC7 (B).



Within each method, all replicates clustered together whether it be at the extraction, digestion or injection levels. Within methods, cow breeds did not cluster together.

Figure BU4: Venn diagram of the number of unique protein accessions and Gene Ontology (GO) classification of known proteins per extraction method.



As different as methods A, B and, C were from each other, they recovered the same types of proteins from skim milk samples. In total, 186 different protein accessions were identified across all methods. Identities common to all three sets of extracts include major milk proteins: caseins (aS1, aS2, b, and k forms), lactoferrin, albumin, bLG, aLA, complement C3 and butyrophilin. Yet proteins present in low abundance in milk were also identified, such as enzymes and minor glycoproteins, as well as many immunoglobulins (Igs), antibodies and antigens.

All considered, GO classifications were very similar across methods, with method B generally displaying the smallest number of proteins per category. As expected the most prominent protein category in the "Cellular Component" classification was the "extracellular region" as most milk proteins are secreted.

BU CONCLUSIONS

Overall 186 unique accessions, major and minor proteins, were identified across methods. Method C yielded the best resolved SDS-patterns and highest protein recovery rates, method A yielded the greatest number of accessions, and, of the three procedures, method B was the least compatible of all. Our results also highlighted breed differences between the proteins in milk of Jersey and Holstein cows. This bottom-up method will be applied in a high-throughput fashion to hundreds of milk samples in order to better understand the dynamic nature of the milk proteome, with potential applications for milk quality and processing (e.g. cheese, health).

BOTTOM-UP (BU)

Victoria The Place To Be

LA TROBE UNIVERSITY

biosciences research