2-28-2021

Quantitative mass spectrometry-based analysis of proteins related to cattle and their products – focus on cows' milk beta-casein proteoforms

James A. Broadbent

Mark R. Condina

Michelle L. Colgrave
Edith Cowan University

10.1016/j.ymeth.2020.09.011
This is an author's accepted manuscript of: Broadbent, J. A., Condina, M. R., & Colgrave, M. L. (2021). Quantitative mass spectrometry-based analysis of proteins related to cattle and their products–focus on cows' milk beta-casein proteoforms. *Methods, 186*, 112-118. https://doi.org/10.1016/j.ymeth.2020.09.011

Quantitative mass spectrometry-based analysis of proteins related to cattle and their products – focus on cows' milk beta-casein proteoforms

James A. Broadbent[1], Mark R. Condina[2], Michelle L. Colgrave[1,3]*

1. CSIRO Agriculture and Food, 306 Carmody Road, St Lucia, Queensland 4067 Australia
2. University of South Australia, Future Industries Institute, South Australia, 5000, Australia
3. ARC Centre of Excellence in Peptide and Protein Science, Edith Cowan University, 270 Joondalup Drive, Joondalup, Western Australia 6027, Australia

*Corresponding author

Michelle L. Colgrave

A: CSIRO Agriculture and Food, St Lucia, Queensland 4067 Australia

E: Michelle.colgrave@csiro.au

P: +61 7 3214 2697

KEYWORDS

LC-MS; protein quantitation; cattle; ruminants; milk; beta-casein

ABSTRACT

Modern mass spectrometers can accurately measure thousands of compounds in complex mixtures over a given liquid chromatograph method, depending on desired outcome and method duration. This stream of analytical chemistry has wide ranging application across food, pharma, environmental, forensics, clinical and research. With consistent pressure on both the ruminant production and product industries to face new and substantial challenges, liquid chromatography-mass spectrometry (LC-MS) is an ideal tool to identify, detect and quantify markers of breeding, production and adaption to support both research and industry to overcome these challenges. Herein, we provide a description of the theoretical basis and framework for LC-MS as a rapidly developing technique and highlight its application in measuring cattle and cattle product traits through protein quantitation with specific focus on beta-casein proteoforms.


MILK
Response
m/z
P
H
Concentration
A1
A2

1.0 DESCRIPTION OF THEORETICAL BASIS AND FRAMEWORK FOR LIQUID CHROMATOGRAPHY-MASS SPECTROMETRY (LC-MS)

1.1 An introduction to LC-MS

Liquid chromatography-mass spectrometry (LC-MS) is an analytical chemistry technique that permits the time-resolved separation of compounds (e.g. proteins or peptides) in complex mixtures as well as the simultaneous detection of specific compounds based on their elemental and structural properties. Through carefully considered experimental design and controls, the technique permits accurate and precise chemical measurements [1], for instance in order to quantify a specific compound (e.g. beta-casein [β-CN] abundance in milk) or define a physical trait (e.g. a specific animal's carcass quality [2]).

In principle, liquid chromatography (LC) involves exploiting the balance between a compound's propensity to interact with a stationary chemical structure and a competing mobile phase in order to concentrate and resolve a compound in time. For peptides and proteins, the stationary phase typically consists of chains of 18 carbons in length bound to a silica resin, called C18 (although other chemistries such as C8 and C4 are also common); whereas, the mobile phase is typically a mixture of acidified aqueous and organic solvent. Compounds are resolved by passing a gradient of the organic solvent through a column of the stationary phase, thereby sequentially eluting each as their proclivity to bind to the stationary phase is outweighed by their affinity for a specific concentration of the organic solvent mobile phase. While LC has a well-characterised ability to resolve compounds in complex matrices, the specificity afforded by this method can be limited in terms of quantitative purposes due to co-elution of compounds. In this respect, LC is commonly coupled to mass spectrometry (MS) to provide a combination of on-column concentration and separation with specific detection.

Mass spectrometry is an approach for the separation and detection of ionised gas-phase compounds by their mass to charge ratios (m/z) [3]. When coupled to LC, the liquid sample exiting from the LC column enters the instrument's ion source where a combination of voltage, temperature and inert gas are applied to the sample. These factors are used to aerosolize, guide and de-solvate the sample while concomitantly imparting charge from the de-solvating liquid droplets to the compounds within during the process [4]. The resulting gas phase ions can then be drawn into the mass spectrometer by means of a voltage gradient. Therein, the ions are manipulated using electric or magnetic fields, voltage and gas gradients for the purposes of focusing, transmission, isolation, — in some cases, chemical reactions [5] — or fragmentation. In terms of fragmentation mechanisms, typically used approaches rely on low or high energy collisional induced dissociation (CID) [6], employing argon, helium or nitrogen as a collision gas, or fragmentation using electron-based gas dissociation approaches, which rely on interactions between gas phase ions and electrons inducing fragmentation [7]. Electron-based techniques include electron transfer dissociation (ETD) [8], electron capture dissociation (ECD) [9], hot ECD (HECD) [10], electron-detachment dissociation (EDD) [11], electronic-excitation dissociation (EED) [12] and electron induced dissociation (EID) [13]. The fragmentation approach applied will be dependent on the analyte being measured and the type of analyser. Through combinations of these gas phase manipulations, a selection of ionised compounds will transit the mass spectrometer to ultimately reach the instrument's detector and produce a response. This combination of ion source, analyser and detector (or combined analyser/detector) are the key features of all mass spectrometers.

There are a range of LC-MS configurations used in modern analytical chemistry. In terms of LC, there are two principal instrument types: high performance liquid chromatography and ultra-performance liquid chromatography, with the latter offering enhancements in terms of speed, resolution and sensitivity [14]. LC systems can also operate across several typical solvent flow rates, the lowest being nano flow (~10–2500 nL/min) offering the highest sensitivity but lowest robustness and peak

resolution; the intermediate flow rates designated capillary flow (~1-10 µL/min) and micro flow (~10-100 µL/min) offering a balance of robustness and sensitivity and moderate peak resolution; and, analytical flow (~500-2000 µL/min) offering ultimate robustness and peak resolution but lowest sensitivity [14]. The requirement of routine testing laboratories to produce reproducible results in a timely fashion sees analytical flow often adopted, while research laboratories will typically use lower flow rates due to their sensitivity gains.

In terms of MS systems, there are three principal commercial instrument configurations used in quantitative analytical chemistry: low-resolution triple quadrupole (QqQ), high-resolution quadrupole-quadrupole-time-of-flight (QqTOF) and orbital trapping instruments. Additionally, there are a range of approaches to compound manipulation and detection afforded by these configurations to specifically obtain the quantitative signal for a compound (or analyse data for the specific post-acquisition extraction of a compound's signal from complex data).

1.2 Data acquisition approaches

In terms of QqQ type instruments, the principle approach to specific compound detection is by multiple reaction monitoring (MRM) [15]. This approach relies on the pre-selection of target precursor ions and their specific product ions prior to data acquisition. In this respect, as a sample's ions traverse the mass spectrometer, the instrument's first quadrupole analyser isolates the target precursor ions (or, more precisely, regions of m/z space containing the target precursor ion and perhaps other non-target precursor ions). These target ions are then fragmented by collision-induced dissociation in the second quadrupole, which functions as a collision cell, to produce product ions. Subsequently, individual target product ions are isolated (one at a time) in the third quadrupole to produce a response at the detector that is specific for the combination of the target compound's transition from a precursor to a product ion.

Akin to MRM, high resolution MS2 (MS/MS) approaches rely on the reproducible detection of the product ions that result from peptide fragmentation [16,17]. Here a compound of interest is isolated and fragmented, after which a full spectrum of high-resolution product ions is collected using the TOF or orbital trap. Specific product ion responses are then extracted from the resulting full spectrum MS/MS data to computationally determine a quantitative measurement post-acquisition, rather than on-the-fly. Due to its similarity to MRM, this method is referred to as MRM-high resolution (MRM-HR) or parallel reaction monitoring (PRM).

The high resolution afforded by QqTOF and orbital trapping mass spectrometers may also be exploited to resolve one compound from another, thereby providing requisite specificity (when combined with a known LC elution time) to identify and quantify target peptides. Such an approach is referred to as MS1 filtering [18].

High resolution MS analysis may also be directed towards intact protein detection and quantification [19,20]. Specifically, when referring to the analysis of intact proteins by MS1, this approach is referred to as "top-down intact" analysis [21]. Where these intact proteins are fragmented and sequenced is referred to as "top-down sequencing" analysis [22].

1.3 New technologies

In addition to MRM, intact and MRM-HR, one additional quantitative acquisition approach — data independent acquisition (DIA) — has gained vast popularity since its re-introduction at the HUPO world congress in 2010 [23]. Nowadays, DIA in the proteomics field describes a substantial suit of methods whereby fragment ion data for all detectable peptides is collected in such a fashion that, in

principle, specific quantitative data is obtainable for every detectable feature in each sample [24]. This approach involves targeting a band of the mass-to-charge space (rather than a single ion worth, which is about 0.7 Th), thereby co-isolating and co-fragmenting multiple ions to generate a high-resolution convoluted product ion spectrum. In wide band DIA analysis, such as sequential window acquisition of all theoretical ions (SWATH), this mass-to-charge band is stepped across the selected mass range to enable chromatographic resolution of all detectable peptides' product ions [23]. Alternatively, the entire mass range can targeted to acquire data at low and high collision energy, to detect all co-eluting fragments in a single scan in the style of MS^E acquisition [25]. Akin to MRM-HR, specific product ions are then computationally extracted post-acquisition to determine quantitative measurements. Importantly, when combined with internal standards, this acquisition approach has shown promise for absolute quantitation across the proteome [26,27]. See Zhang et al. for further reading [24].

Ion mobility spectrometry is a recently emerging technique in the proteomics field that is rapidly growing in popularity, but yet to find its way to routine protein or peptide quantitation methods. Ion mobility spectrometry makes use of collisional cross section [28] (the space around a compound in which the center of another compound must occupy in order for a collision to occur) or dipole moment [29] (the product of the magnitude of the charge and the distance between the centres of the positive and negative charges of a polar compound) to resolve compounds in time. These resolved compounds are then typically passed to the entrance of a mass spectrometer for conventional MRM, intact, MRM-HR or DIA analysis, thereby enabling ion mobility as an added concentration and resolving coordinate in additional to retention time, precursor and product ion masses. Such an advantage may lead to improvements in both specificity and acquisition time and should be closely watched.

1.4 Considerations for quantitative LC-MS

Mass spectrometers are not inherently quantitative instruments. In order to produce quantitative results using a matrix matched calibration curve-based approach, mass spectrometric assays should:

- compare the responses of unknown samples to external synthetic calibration standards that precisely mimic the target analyte;
- calibration standards should be constituted in sample matrix that is devoid of the target analyte but closely matches the target matrix (or matrix effects regularly quantified);
- use the calibration standards to quantify the linear range of the compound's response;
- preferably include isotopically labelled internal standards, that precisely mimic the target analyte in all but mass, in all standards and unknowns as early on in the sample processing as possible to enable the correction of matrix and technical effects;
- internal standards should also be at a concentration comparable to that expected in the unknown samples; and,
- quantify the upper and lower limits of quantitation (ULOQ and LLOQ), precision, accuracy and recovery [1].

Such considerations, while not exhaustive, are fundamental to generating robust quantitative results and should be followed wherever practicable.

In addition to the matrix matched external calibration, additional calibration strategies are available, including: standard additional internal calibration and internal calibration by direct addition. The former of these involves the preparation of a series of calibration samples where the matrix is the sample itself. A calibration curve is then generated and the difference between the zero point on the concentration axis and the intersection of the calibration line to the same axis is taken as the concentration estimate [30]. The latter method, also commonly referred to as isotope dilution mass

spectrometry, typically involves comparing the response of a known amount of an internal spike of an isotopologue to that of the native target compound in order to estimate the target compound concentration by means of their ratio [31]. For further reading regarding reliable LC-MS measurements see the Guide to achieving reliable quantitative LC-MS measurements [1].

2.0 APPLICATIONS IN CATTLE

2.1 LC-MS assays for cattle proteins

To date there have been relatively few published papers outlining the development of quantitative LC-MS assays to measure proteins in cattle tissues or products. These studies have included the assessment of beef gelatine origin [32], leather origin [33], meat quality [2] and low abundant serum protein quantitation [34]. Where the determination of gelatine and leather origin and tenderness has clear application in food quality/cultural requirement and fraud detection, the measurement of low abundance serum proteins is desirable to measure animal traits via a protein proxy. Such traits may include residual feed intake through growth factor measurement [35,36] or lactation transition outcome through cytokine measurement [37,38].

In addition to the abovementioned cattle tissues and products, cow's milk proteins have also been the subject of substantial investigation by quantitative LC-MS.

2.2 Cows' milk A1 and A2 β-CN proteoforms

Cows' milk is a significant source of calories and nutrition for a segment of the population, with the total milk production for 2017 estimated at 823,966 kt (approximately 749,060 ML) [39]. While a critical source of calories, cows' milk also poses a health risk to sub-populations in terms of allergy (particularly in infants) and intolerance to milk components, namely protein and lactose. In this respect, the prevalence of cows' milk allergy by skin prick test is observed in the order of 0.2% to 2.5% whereas serum IgE diagnosis ranges from 2% to 9% [40]. Global lactose intolerance is observed in the order of 68%, with the maximum observed level at 100% for the South Korean population [41]. In addition to these well-known milk-associated digestive issues, another potential and lesser-known impact of cows' milk on human health has been under investigation in the last few decades.

Substantial research effort has been placed in the investigation of the role that β-CN proteoforms may play in human health and development, where the term proteoform refers to "all of the different molecular forms in which the protein product of a single gene can be found, including changes due to genetic variations, alternatively spliced RNA transcripts and post-translational modifications" [42]. The drive for this research effort concerns the genetically recent appearance of the A1 β-CN protein sequence variant and the inability of human digestive systems to have adapted to this dietary protein within the same timeframe. Detail around this maladaptation primarily concerns a bioactive peptide that is liberated from A1 β-CN during digestion but to a much lesser degree from A2 β-CN [43]. This peptide, designated beta casomorphin-7 (BCM7), and the A1 protein in general, have been associated with a slew of disease states [6,7,8,9,10]. For this reason, substantial selective breeding and genetic testing programs have been undertaken to develop a cows' milk product that contains only the A2 proteoform of β-CN [49]. Perhaps more importantly, the process of detecting β-CN genotypes in milk-producing cows was patented and is now licensed by the A2 Milk Company for use in the segregation of animals (WO2003100074-A2), such that an A2 β-CN-only product — "A2 milk" — can be produced and sold exclusively through the A2 Milk Company brand.

The production of A2-milk relies on the separation of herd animals such that A2 homozygous animals are milked separately to A2 non-homozygous animals. While the science that supports such a practice

is sound, there are several circumstances that may lead to a final product reaching supermarket shelves that is not free from other β-CN proteoforms. For example, inadvertent cross-contamination through use of the same milking or freight equipment for A2 and other milk products; human error in the dairy supply chain; and, purposeful adulteration of milk product (as A2 milk attracts a premium wholesale price). In this regard, beyond the genetic separation of milk producing cows, the abundance of A1 and A2 β-CN proteoforms should be measured and reported where claims of A2 protein-only milk are made.

A1 and A2 β-CN proteoforms from cow are identical in primary sequence with the exception of a Pro>His substitution at amino acid position 82 in A1 β-CN (also found in B, C, F and G proteoforms) [50]. Such a relatively small difference in chemical composition renders several common analytical techniques impractical, though not impossible [51], due to specificity or throughout, such as: ELISA and SDS gel electrophoresis. Importantly, the detection and precise measurement of this minute difference is ideally suited to mass spectrometric analysis; however, key experimental requirements are essential to such measurement, as outlined above under **section 1.4**. Below we look at efforts to detect and quantify A1 and A2 β-CN in cows' milk and highlight their adherence to the above quantitation requirements as well as their shortcomings, where relevant.

2.3 A1 and A2 β-CN proteoform quantitation by intact protein acquisition

Quantitative [52–54] and qualitative [55–58] mass spectrometric analysis of intact β-CN proteoforms have been described for cows' milk. The earliest report of a quantitative LC-MS approach described the use of a C18-HPLC-QqTOF LC-MS to quantify β-CN A1, A2, B and C as a proportion of total caseins — as measured by mid-infrared spectroscopy [52]. This analytical approach does not report the use of internal standards to correct for matrix effects and instrument performance, or external standards for measurement of each protein against like compounds in equivalent matrix. The authors make assumptions that the mass spectrometric (or UV) response of each of the four casein proteins are equivalent and that UV peaks do not contain more than one component. Though the methods may have benefited from additional controls, this report of intact β-CN analysis was the first to demonstrate detection and relative quantitation of intact β-CN proteoforms by LC-MS.

A more recent report of cow's milk protein quantitation details the use of a C8-UPLC-QqTOF LC-MS to quantify alpha-casein; β-CNs A1, A2, B, I; kappa-casein; alpha-lactalbumin; beta-lactoglobulin; bovine serum albumin; and, a selection of their proteoforms resulting from phosphorylation [53]. The authors improve upon the previous work and describe the use of a surrogate internal standard protein to assess matrix effects, assess protein response linearity using calibration curves and quantify the matrix effects from multiple cows' milk samples. In terms of the calibration curves, the authors describe the use of a single β-CN standard for measuring A1, A2, B and I β-CN, where the sum signal of these four proteoforms is used to determine curve parameters over an eight-point series. In doing so, the authors demonstrate the linear range for accurate quantitation. When measuring unknown milk samples, the authors are then able to use the external standards for the purpose of validating protein identity by mass and retention time matching. This approach allows profiling of various ruminant milk samples to provide *accurate* relative quantitative changes between samples; however, this approach is unable to provide absolute quantitation, primarily due to the absence of specific reference and internal standards for A1/A2 β-CN.

The most recent and final report of intact β-CN quantitation goes to great effort to consider the challenges of LC-MS quantitation of compounds in a complex food matrix [54]. Therein, the authors use C4-UPLC-Q-orbital trap LC-MS to: generate external calibration curves that are adjusted for standard purity; quantify matrix suppression; measure multiple proteoforms — up to two

lactosylation modifications per protein — and correct for signal suppression due to modification; and, compensate for instrument performance using external quality control injections. Through their thorough analysis of the β-CN standard, the authors also identify 31 β-CN proteoforms due to unanticipated N- and C-terminal truncation. In this regard, eight proteoforms were found to be specific for A2 β-CN, which appear to be summed for unmodified β-CN and glycosylated forms separately to be converted into final concentration measurements. This work represents the most comprehensive approach to β-CN quantitation in cows' milk reported to date, but also highlights the challenges pertaining to a lack of specific reference standards for each protein form, the measurement of multiple proteoforms and the effects on competing ionisation on achievable dynamic range within a complex food matrix.

Intact analysis and top down sequencing can aid in inferring proteoforms that arise from sequence modifications and truncations; however, there remains the challenge of understanding and quantifying the gas phase characteristics of each proteoform in order to obtain absolute measurements — as demonstrated elegantly by Fuerer *et al.* for β-CN. This problem is pertinent to β-CN from cow's milk where it has been observed that up to eight proteoforms can be resolved by 2-dimensional electrophoresis and LC-MS [54,59,60] and further, that β-CN is degraded as a function of temperature and time [61]. An alternative approach is to use the peptide products of proteolytic digestion to reduce several variable proteoforms down to a single peptide feature, thereby negating the combinatorial problem of protein modifications and truncations (Figure 1). Where intact analysis may benefit from simplified sample preparation and higher throughput, a proteolytic approach may simplify measurement and data output but also benefit from access to more mature software for data processing, interpretation, and reporting.

2.4 A1 and A2 β-CN proteoform quantitation by proteolytic digestion products

Given the challenges associated with unknown β-CN truncation and/or protein modifications in determining protein abundance by intact mass, proteolytic sample processing and analysis may provide a distinct advantage in addressing this combinatorial problem. In this respect, enzymatic digestion of β-CN by trypsin (or other proteolytic enzymes) may be used to reduce the various proteoforms to a single — or small number of peptides — that can be measured accurately (Figure 1). Indeed, there is already public evidence regarding the tryptic digestion of β-CN, which shows that there are four detectable peptides that carry the diagnostic amino acid at position 82 (http://www.peptideatlas.org/; Accession: ENSBTAP00000003409) [62]. Such a number is half that of the proteoforms observed by intact mass analysis in a β-CN standard [54]. Yet, no research papers have reported the quantitation of these four peptides in order to measure β-CN proteoforms; although, there are several examples of A1 and A2 β-CN detection by proteolytic peptide analysis.

In 2017, Duarte-Vázquez, described the gel separation of A1 and A2 beta casein proteoforms by SDS-PAGE analysis. The authors report the subsequent excision of the purported A1 and A2 β-CN bands and digestion of their protein contents using the proteolytic enzyme thermolysin [49]. The extracted peptides were then analysed by C18-nanoLC-orbital trap-Ion trap LC-MS/MS. Therein, the authors make use of a spectral counting approach to quantify the amounts of the individual β-CN proteoforms and express these values as mean relative abundance. Of note, the delta mass of the two casein proteoforms is 40 Da, which may be challenging to separate and accurately excise from a typical SDS-PAGE gel; however, no images of the gels in question are provided to make an objective assessment. This challenging separation is made more obvious in the authors' follow up paper that specifically describes the use of SDS-PAGE separation and identical LC-MS/MS conditions to analyse the two β-CN proteoforms [63]. The gel image provided shows minimal visual separation of the proteoforms. Therein, the authors demonstrate the detection of one peptide specific for A1 β-CN but do not

describe whether the peptide is found in both excised regions of the gel. Furthermore, the peptide sequence in question, PFPGPIHN, is not a specific product of thermolysin digestion at its N- or C-termini, rendering it a non-specific product that is unsuitable for quantitation experiments.

Work published in 2019 details efforts towards absolute quantitation of a panel of 20 milk proteins using bottom up analysis [64]. In this work the investigators use C18-HPLC-Linear Ion Trap LC-MS to quantify milk protein target peptides using MRM. While the authors do not specifically quantify A1 and A2 proteoforms of β-CN, they do report the use of isotopically labelled internal standards in order to account for matrix and technical variation as well as the determination of the upper and lower limits of detection and LOQ. Furthermore, the investigators report the monitoring of instrument performance through the interleaved analysis of a bovine serum albumin digest throughout the acquisition batch.

In related work, Nguyen and co-workers describe the quantitation of BCM-5 and BCM-7 — β-CN peptides that result from commercial milk processing and digestion in the gut [65]. Therein, the authors describe the use of synthetic peptide calibration curves and isotopically labelled internal standards combined with either MRM analysis on a C18-UPLC-QqQ LC-MS set up, or high resolution MS1 analysis on a C18-HPLC-OrbitalTrap LC-MS system. The authors demonstrate their adherence to the European Commission criteria (2002) for quantitative mass spectrometric detection, and in doing so, provide an example of the requirements for A1 and A2 β-CN quantitation in cows' milk.

Taken together, the published examples of A1 and A2 β-CN quantitation, and related studies, demonstrate that A1 and A2 β-CN are detectable by both intact and proteolytic digestion approaches. At present, none of the published works adhere to all key quantitative principals for A1/A2 β-CN quantitation, including the use of heavy labelled internal standards and external calibration curves (Supplementary table 1). What is also evident is that measurement of multiple proteoforms or peptides is a requirement for accurate quantitation, due to inherent genetic variability and industrial processing of cows' milk. Furthermore, it is clear is that multiple skilled laboratories are working towards a method for accurate quantitation of β-CN proteoforms and that, given access to reference standards, publication of such a method may be anticipated in the near future.

3.0 CONCLUSIONS AND OUTLOOK

Liquid chromatography-mass spectrometry is a technique that has found wide adoption for the measurement of proteins in complex mixtures in human clinical chemistry [66–70]. While there are examples of studies that describe the identification of biomarkers for ruminant traits and their transition to quantitative LC-MS assays (e.g. [2,71]), there remain limited papers demonstrating the development of robust LC-MS quantitation of cattle protein biomarkers. Even so, where methods have been developed, this technique has been able to provide information regarding cattle product origin, quality and consumer safety. Moreover, it may be expected that investigators will continue to identify and develop robust LC-MS assays to permit the accurate measurement or prediction of animal rearing and production traits from known protein biomarkers [72].

With respect to β-CN quantitation in cows' milk, the current literature demonstrates that there remains a gap between aspiration and practical absolute quantitation. While ideally suited to LC-MS protein quantitation, this application possesses a range of challenges, particularly in terms of the management of the complex proteoforms or proteolytic products arising from the biological, production and storage processes associated with cows' milk production. Both intact and proteolytic product analysis offer potential solutions to milk analysis. For instance, intact analysis offers the ability to provide quality control through the screening of bulk product for anomalies as well as accurate

relative quantitation for reporting the proportions of A1 and A2 proteoforms in products or bulk milk. Should absolute quantitation be a requirement, then analysts may look to measure proteolytic digestion products, where there is ready access to synthetic peptide calibration standards as well as isotopically labelled peptide internal standards. Both approaches have merit, depending on need, but also have their challenges. Regardless, the development of a robust LC-MS assay for β-CN proteoforms — or other informative bovine proteins — has the potential to provide both certainty for producers as well as confidence for the consumer, and as such is a task worthy of further pursuit.

ACKNOWLEDGMENTS

Mark Condina is supported by Bioplatforms Australia, the University of South Australia, and the State and Federal Governments, which co-fund the NCRIS-enabled Mass Spectrometry and Proteomics facility at the University of South Australia. The authors acknowledge the kind assistance of Dr Utpal Bose and Dr Omar Mendoza Porras for critical review of the manuscript as well as Dr Brad Patterson for valuable discussions.

Credit Author Statement

James A. Broadbent: Conceptualization, Investigation, Writing - Original Draft; **Mark R. Condina:** Investigation, Writing - Review & Editing; **Michelle L. Colgrave:** Conceptualization, Writing - Review & Editing

References

[1] M. Sargent (Ed.), Guide to achieving reliable quantitative LC-MS measurements, RSC Analytical Methods Committee, 2013.
[2] M. Bonnet, J. Soulat, J. Bons, S. Léger, L. De Koning, C. Carapito, B. Picard, Quantification of biomarkers for beef meat qualities using a combination of Parallel Reaction Monitoring- and antibody-based proteomics, Food Chemistry. 317 (2020) 126376. https://doi.org/10.1016/j.foodchem.2020.126376.
[3] F.A. Mellon, MASS SPECTROMETRY | Principles and Instrumentation, in: B. Caballero (Ed.), Encyclopedia of Food Sciences and Nutrition (Second Edition), Academic Press, Oxford, 2003: pp. 3739–3749. https://doi.org/10.1016/B0-12-227055-X/00746-X.
[4] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules, Science. 246 (1989) 64–71. https://doi.org/10.1126/science.2675315.
[5] B.L.J. Poad, H.T. Pham, M.C. Thomas, J.R. Nealon, J.L. Campbell, T.W. Mitchell, S.J. Blanksby, Ozone-Induced Dissociation on a Modified Tandem Linear Ion-Trap: Observations of Different Reactivity for Isomeric Lipids, Journal of the American Society for Mass Spectrometry. 21 (2010) 1989–1999. https://doi.org/10.1016/j.jasms.2010.08.011.
[6] J. Mitchell Wells, S.A. McLuckey, Collision-Induced Dissociation (CID) of Peptides and Proteins, in: Methods in Enzymology, Academic Press, 2005: pp. 148–185. https://doi.org/10.1016/S0076-6879(05)02005-7.
[7] J.W. Jones, C.J. Thompson, C.L. Carter, M.A. Kane, Electron-induced dissociation (EID) for structure characterization of glycerophosphatidylcholine: determination of double-bond positions and localization of acyl chains, Journal of Mass Spectrometry. 50 (2015) 1327–1339. https://doi.org/10.1002/jms.3698.

[8] J.E.P. Syka, J.J. Coon, M.J. Schroeder, J. Shabanowitz, D.F. Hunt, Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 9528–9533. https://doi.org/10.1073/pnas.0402700101.
[9] R.A. Zubarev, N.L. Kelleher, F.W. McLafferty, Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process, J. Am. Chem. Soc. 120 (1998) 3265–3266. https://doi.org/10.1021/ja973478k.
[10] F. Kjeldsen, K.F. Haselmann, B.A. Budnik, F. Jensen, R.A. Zubarev, Dissociative capture of hot (3–13 eV) electrons by polypeptide polycations: an efficient process accompanied by secondary fragmentation, Chemical Physics Letters. 356 (2002) 201–206. https://doi.org/10.1016/S0009-2614(02)00149-5.
[11] B.A. Budnik, R.A. Zubarev, MH2+· ion production from protonated polypeptides by electron impact: observation and determination of ionization energies and a cross-section, Chemical Physics Letters. 316 (2000) 19–23. https://doi.org/10.1016/S0009-2614(99)01256-7.
[12] M.L. Nielsen, B.A. Budnik, K.F. Haselmann, R.A. Zubarev, Tandem MALDI/EI ionization for tandem Fourier transform ion cyclotron resonance mass spectrometry of polypeptides, International Journal of Mass Spectrometry. 226 (2003) 181–187. https://doi.org/10.1016/S1387-3806(02)00970-3.
[13] J.R. Gord, S.R. Horning, J.M. Wood, R.G. Cooks, B.S. Freiser, Energy deposition during electron-induced dissociation, J. Am. Soc. Mass Spectrom. 4 (1993) 145–151. https://doi.org/10.1016/1044-0305(93)85070-E.
[14] M. de la Guardia, S. Armenta, Chapter 7 - Downsizing the Methods, in: M.D.L. Guardia, S. Armenta (Eds.), Comprehensive Analytical Chemistry, Elsevier, 2011: pp. 157–184. https://doi.org/10.1016/B978-0-444-53709-6.00007-0.
[15] P. Picotti, R. Aebersold, Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions, Nature Methods. 9 (2012) 555–566. https://doi.org/10.1038/nmeth.2015.
[16] A.C. Peterson, J.D. Russell, D.J. Bailey, M.S. Westphall, J.J. Coon, Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics, Molecular & Cellular Proteomics. 11 (2012) 1475–1488. https://doi.org/10.1074/mcp.O112.020131.
[17] S. Gallien, E. Duriez, C. Crone, M. Kellmann, T. Moehring, B. Domon, Targeted Proteomic Quantification on Quadrupole-Orbitrap Mass Spectrometer, Molecular & Cellular Proteomics. 11 (2012) 1709–1723. https://doi.org/10.1074/mcp.O112.019802.
[18] B. Schilling, M.J. Rardin, B.X. MacLean, A.M. Zawadzka, B.E. Frewen, M.P. Cusack, D.J. Sorensen, M.S. Bereman, E. Jing, C.C. Wu, E. Verdin, C.R. Kahn, M.J. Maccoss, B.W. Gibson, Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation, Mol. Cell Proteomics. 11 (2012) 202–214. https://doi.org/10.1074/mcp.M112.017707.
[19] Y. Chen, P. Mao, D. Wang, Quantitation of Intact Proteins in Human Plasma Using Top-Down Parallel Reaction Monitoring-MS, Anal. Chem. 90 (2018) 10650–10653. https://doi.org/10.1021/acs.analchem.8b02699.
[20] I. Ntai, K. Kim, R.T. Fellers, O.S. Skinner, A.D. Smith, B.P. Early, J.P. Savaryn, R.D. LeDuc, P.M. Thomas, N.L. Kelleher, Applying Label-Free Quantitation to Top Down Proteomics, Anal. Chem. 86 (2014) 4961–4968. https://doi.org/10.1021/ac500395k.
[21] J.D. Tipton, J.C. Tran, A.D. Catherman, D.R. Ahlf, K.R. Durbin, N.L. Kelleher, Analysis of Intact Protein Isoforms by Mass Spectrometry, J Biol Chem. 286 (2011) 25451–25458. https://doi.org/10.1074/jbc.R111.239442.
[22] W. Cui, H.W. Rohrs, M.L. Gross, Top-Down Mass Spectrometry: Recent Developments, Applications and Perspectives, Analyst. 136 (2011) 3854–3864. https://doi.org/10.1039/c1an15286f.
[23] L.C. Gillet, P. Navarro, S. Tate, H. Röst, N. Selevsek, L. Reiter, R. Bonner, R. Aebersold, Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new

concept for consistent and accurate proteome analysis, Mol. Cell Proteomics. 11 (2012) O111.016717. https://doi.org/10.1074/mcp.O111.016717.
[24] F. Zhang, W. Ge, G. Ruan, X. Cai, T. Guo, Data-Independent Acquisition Mass Spectrometry-Based Proteomics and Software Tools: A Glimpse in 2020, PROTEOMICS. n/a (n.d.) 1900276. https://doi.org/10.1002/pmic.201900276.
[25] M. Fioramonte, P.C. Guest, D. Martins-de-Souza, LC-MSE for Qualitative and Quantitative Proteomic Studies of Psychiatric Disorders, in: P.C. Guest (Ed.), Proteomic Methods in Neuropsychiatric Research, Springer International Publishing, Cham, 2017: pp. 115–129. https://doi.org/10.1007/978-3-319-52479-5_7.
[26] S.J. Rice, X. Liu, J. Zhang, C.P. Belani, Absolute Quantification of All Identified Plasma Proteins from SWATH Data for Biomarker Discovery, PROTEOMICS. 19 (2019) 1800135. https://doi.org/10.1002/pmic.201800135.
[27] J.C. Silva, M.V. Gorenstein, G.-Z. Li, J.P.C. Vissers, S.J. Geromanos, Absolute Quantification of Proteins by LCMSE : A Virtue of Parallel ms Acquisition, Molecular & Cellular Proteomics. 5 (2006) 144–156. https://doi.org/10.1074/mcp.M500230-MCP200.
[28] V. Gabelica, E. Marklund, Fundamentals of ion mobility spectrometry, Current Opinion in Chemical Biology. 42 (2018) 51–59. https://doi.org/10.1016/j.cbpa.2017.10.022.
[29] T.P.I. Lintonen, P.R.S. Baker, M. Suoniemi, B.K. Ubhi, K.M. Koistinen, E. Duchoslav, J.L. Campbell, K. Ekroos, Differential Mobility Spectrometry-Driven Shotgun Lipidomics, Anal. Chem. 86 (2014) 9662–9669. https://doi.org/10.1021/ac5021744.
[30] S.L.R. Ellison, M. Thompson, Standard additions: myth and reality, Analyst. 133 (2008) 992–997. https://doi.org/10.1039/b717660k.
[31] J. Villanueva, M. Carrascal, J. Abian, Isotope dilution mass spectrometry for absolute quantification in proteomics: Concepts and strategies, Journal of Proteomics. 96 (2014) 184–199. https://doi.org/10.1016/j.jprot.2013.11.004.
[32] B. jannat, K. Ghorbani, S. Kouchaki, N. Sadeghi, E. Eslamifarsani, F. Rabbani, S. Beyramysoltan, Distinguishing tissue origin of bovine gelatin in processed products using LC/MS technique in combination with chemometrics tools, Food Chemistry. 319 (2020) 126302. https://doi.org/10.1016/j.foodchem.2020.126302.
[33] Y. Kumazawa, Y. Taga, K. Iwai, Y. Koyama, A Rapid and Simple LC-MS Method Using Collagen Marker Peptides for Identification of the Animal Source of Leather, J. Agric. Food Chem. 64 (2016) 6051–6057. https://doi.org/10.1021/acs.jafc.6b02132.
[34] O. Mendoza-Porras, P.R.L. Pires, H. Goswami, F.V. Meirelles, M.L. Colgrave, G. Wijffels, Cytokines in the grass, a lesson learnt: Measuring cytokines in plasma using multiple reaction monitoring mass spectrometry, Rapid Commun. Mass Spectrom. (2020). https://doi.org/10.1002/rcm.8723.
[35] S.S. Moore, F.D. Mujibi, E.L. Sherman, Molecular basis for residual feed intake in beef cattle, J. Anim. Sci. 87 (2009) E41-47. https://doi.org/10.2527/jas.2008-1418.
[36] B.J. Wood, J.A. Archer, J.H.J. van der Werf, Response to selection in beef cattle using IGF-1 as a selection criterion for residual feed intake under different Australian breeding objectives, Livestock Production Science. 91 (2004) 69–81. https://doi.org/10.1016/j.livprodsci.2004.06.009.
[37] E. Trevisi, N. Jahan, G. Bertoni, A. Ferrari, A. Minuti, Pro-Inflammatory Cytokine Profile in Dairy Cows: Consequences for New Lactation, Italian Journal of Animal Science. 14 (2015) 3862. https://doi.org/10.4081/ijas.2015.3862.
[38] P. Brodzki, K. Kostro, L. Krakowski, J. Marczuk, Inflammatory cytokine and acute phase protein concentrations in the peripheral blood and uterine washings of cows with subclinical endometritis in the late postpartum period, Vet Res Commun. 39 (2015) 143–149. https://doi.org/10.1007/s11259-015-9635-4.
[39] FAOSTAT, (n.d.). http://www.fao.org/faostat/en/#data/FBS (accessed March 1, 2020).
[40] R.J. Rona, T. Keil, C. Summers, D. Gislason, L. Zuidmeer, E. Sodergren, S.T. Sigurdardottir, T. Lindner, K. Goldhahn, J. Dahlstrom, D. McBride, C. Madsen, The prevalence of food allergy: a

meta-analysis, J. Allergy Clin. Immunol. 120 (2007) 638–646. https://doi.org/10.1016/j.jaci.2007.05.026.
[41] C.L. Storhaug, S.K. Fosse, L.T. Fadnes, Country, regional, and global estimates for lactose malabsorption in adults: a systematic review and meta-analysis, The Lancet Gastroenterology & Hepatology. 2 (2017) 738–746. https://doi.org/10.1016/S2468-1253(17)30154-1.
[42] L.M. Smith, N.L. Kelleher, Proteoform: a single term describing protein complexity, Nat Methods. 10 (2013) 186–187. https://doi.org/10.1038/nmeth.2369.
[43] Y. Jinsmaa, M. Yoshikawa, Enzymatic release of neocasomorphin and beta-casomorphin from bovine beta-casein, Peptides. 20 (1999) 957–962. https://doi.org/10.1016/s0196-9781(99)00088-1.
[44] J.S.J. Chia, J.L. McRae, A.K. Enjapoori, C.M. Lefèvre, S. Kukuljan, K.M. Dwyer, Dietary Cows' Milk Protein A1 Beta-Casein Increases the Incidence of T1D in NOD Mice, Nutrients. 10 (2018). https://doi.org/10.3390/nu10091291.
[45] R.B. Elliott, D.P. Harris, J.P. Hill, N.J. Bibby, H.E. Wasmuth, Type I (insulin-dependent) diabetes mellitus and cow milk: casein variant consumption, Diabetologia. 42 (1999) 292–296. https://doi.org/10.1007/s001250051153.
[46] M. Laugesen, R. Elliott, Ischaemic heart disease, Type 1 diabetes, and cow milk A1 beta-casein, N. Z. Med. J. 116 (2003) U295.
[47] K.A. Tailford, C.L. Berry, A.C. Thomas, J.H. Campbell, A casein variant in cow's milk is atherogenic, Atherosclerosis. 170 (2003) 13–19. https://doi.org/10.1016/s0021-9150(03)00131-x.
[48] R. Cade, M. Privette, M. Fregly, N. Rowland, Z. Sun, V. Zele, H. Wagemaker, C. Edelstein, Autism and Schizophrenia: Intestinal Disorders, Nutr Neurosci. 3 (2000) 57–72. https://doi.org/10.1080/1028415X.2000.11747303.
[49] M.Á. Duarte-Vázquez, C. García-Ugalde, L.M. Villegas-Gutiérrez, B.E. García-Almendárez, J.L. Rosado, Production of Cow's Milk Free from Beta-Casein A1 and Its Application in the Manufacturing of Specialized Foods for Early Infant Nutrition, Foods. 6 (2017). https://doi.org/10.3390/foods6070050.
[50] CSN2 - Beta-casein precursor - Bos taurus (Bovine) - CSN2 gene & protein, (n.d.). https://www.uniprot.org/uniprot/P02666 (accessed March 1, 2020).
[51] A.M. Caroli, S. Savino, O. Bulgari, E. Monti, Detecting β-Casein Variation in Bovine Milk, Molecules. 21 (2016) 141. https://doi.org/10.3390/molecules21020141.
[52] I. Givens, P. Aikman, T. Gibson, R. Brown, Proportions of A1, A2, B and C β-casein protein variants in retail milk in the UK, Food Chemistry. 139 (2013) 549–552. https://doi.org/10.1016/j.foodchem.2013.01.115.
[53] D. Vincent, A. Elkins, M.R. Condina, V. Ezernieks, S. Rochfort, Quantitation and Identification of Intact Major Milk Proteins for High-Throughput LC-ESI-Q-TOF MS Analyses, PLoS ONE. 11 (2016) e0163471. https://doi.org/10.1371/journal.pone.0163471.
[54] C. Fuerer, R. Jenni, L. Cardinaux, F. Andetsion, S. Wagnière, J. Moulin, M. Affolter, Protein fingerprinting and quantification of β-casein variants by ultra-performance liquid chromatography–high-resolution mass spectrometry, Journal of Dairy Science. 103 (2020) 1193–1207. https://doi.org/10.3168/jds.2019-16273.
[55] N.A. Poulsen, H.B. Jensen, L.B. Larsen, Factors influencing degree of glycosylation and phosphorylation of caseins in individual cow milk samples, Journal of Dairy Science. 99 (2016) 3325–3333. https://doi.org/10.3168/jds.2015-10226.
[56] P.D. Frederiksen, K.K. Andersen, M. Hammershøj, H.D. Poulsen, J. Sørensen, M. Bakman, K.B. Qvist, L.B. Larsen, Composition and effect of blending of noncoagulating, poorly coagulating, and well-coagulating bovine milk from individual Danish Holstein cows, Journal of Dairy Science. 94 (2011) 4787–4799. https://doi.org/10.3168/jds.2011-4343.
[57] H.B. Jensen, N.A. Poulsen, K.K. Andersen, M. Hammershøj, H.D. Poulsen, L.B. Larsen, Distinct composition of bovine milk from Jersey and Holstein-Friesian cows with good, poor, or

noncoagulation properties as reflected in protein genetic variants and isoforms, J. Dairy Sci. 95 (2012) 6905–6917. https://doi.org/10.3168/jds.2012-5675.
[58] D.D. Nguyen, V.A. Solah, F. Busetti, G. Smolenski, T. Cooney, Application of ultra-high performance liquid chromatography coupled to high-resolution mass spectrometry (Orbitrap™) for the determination of beta-casein phenotypes in cow milk, Food Chemistry. 307 (2020) 125532. https://doi.org/10.1016/j.foodchem.2019.125532.
[59] O.I. Geicu, L. Stanca, A. Dinischiotu, A.I. Serban, Proteomic and immunochemical approaches to understanding the glycation behaviour of the casein and β-lactoglobulin fractions of flavoured drinks under UHT processing conditions, Scientific Reports. 8 (2018) 1–15. https://doi.org/10.1038/s41598-018-28943-4.
[60] I. Boggs, B. Hine, G. Smolenski, K. Hettinga, L. Zhang, T.T. Wheeler, Changes in the repertoire of bovine milk proteins during mammary involution, EuPA Open Proteomics. 9 (2015) 65–75. https://doi.org/10.1016/j.euprot.2015.09.001.
[61] G. Pierri, D. Kotoni, P. Simone, C. Villani, G. Pepe, P. Campiglia, P. Dugo, F. Gasparrini, Analysis of bovine milk caseins on organic monolithic columns: An integrated capillary liquid chromatography–high resolution mass spectrometry approach for the study of time-dependent casein degradation, Journal of Chromatography A. 1313 (2013) 259–269. https://doi.org/10.1016/j.chroma.2013.08.083.
[62] PeptideAtlas; Accession: ENSBTAP00000003409, (n.d.). http://www.peptideatlas.org/ (accessed February 21, 2020).
[63] M.A. Duarte-Vázquez, C.R. García-Ugalde, B.E. Álvarez, L.M. Villegas, B.E. García-Almendárez, J.L. Rosado, C. Regalado, Use of urea-polyacrylamide electrophoresis for discrimination of A1 and A2 beta casein variants in raw cow's milk, J Food Sci Technol. 55 (2018) 1942–1947. https://doi.org/10.1007/s13197-018-3088-z.
[64] C. Bär, D. Mathis, P. Neuhaus, D. Dürr, W. Bisig, L. Egger, R. Portmann, Protein profile of dairy products: Simultaneous quantification of twenty bovine milk proteins, International Dairy Journal. 97 (2019) 167–175. https://doi.org/10.1016/j.idairyj.2019.01.001.
[65] D.D. Nguyen, F. Busetti, S.K. Johnson, V.A. Solah, Identification and quantification of native beta-casomorphins in Australian milk by LC–MS/MS and LC–HRMS, Journal of Food Composition and Analysis. 44 (2015) 102–110. https://doi.org/10.1016/j.jfca.2015.08.009.
[66] M. Rauh, LC–MS/MS for protein and peptide quantification in clinical chemistry, Journal of Chromatography B. 883–884 (2012) 59–67. https://doi.org/10.1016/j.jchromb.2011.09.030.
[67] A.H.B. Wu, D. French, Implementation of liquid chromatography/mass spectrometry into the clinical laboratory, Clinica Chimica Acta. 420 (2013) 4–10. https://doi.org/10.1016/j.cca.2012.10.026.
[68] O.P. Bondar, D.R. Barnidge, E.W. Klee, B.J. Davis, G.G. Klee, LC-MS/MS Quantification of Zn-α2 Glycoprotein: A Potential Serum Biomarker for Prostate Cancer, Clin Chem. 53 (2007) 673–678. https://doi.org/10.1373/clinchem.2006.079681.
[69] D.R. Barnidge, M.K. Goodmanson, G.G. Klee, D.C. Muddiman, Absolute Quantification of the Model Biomarker Prostate-Specific Antigen in Serum by LC−MS/MS Using Protein Cleavage and Isotope Dilution Mass Spectrometry, J. Proteome Res. 3 (2004) 644–652. https://doi.org/10.1021/pr049963d.
[70] K.S.-Y. Leung, B.M.-W. Fong, LC–MS/MS in the routine clinical laboratory: has its time come?, Anal Bioanal Chem. 406 (2014) 2289–2301. https://doi.org/10.1007/s00216-013-7542-5.
[71] J. Bazile, B. Picard, C. Chambon, A. Valais, M. Bonnet, Pathways and biomarkers of marbling and carcass fat deposition in bovine revealed by a combination of gel-based and gel-free proteomic analyses, Meat Science. 156 (2019) 146–155. https://doi.org/10.1016/j.meatsci.2019.05.018.
[72] M. Gagaoua, V. Monteils, S. Couvreur, B. Picard, Identification of Biomarkers Associated with the Rearing Practices, Carcass Characteristics, and Beef Quality: An Integrative Approach, J. Agric. Food Chem. 65 (2017) 8264–8278. https://doi.org/10.1021/acs.jafc.7b03239.

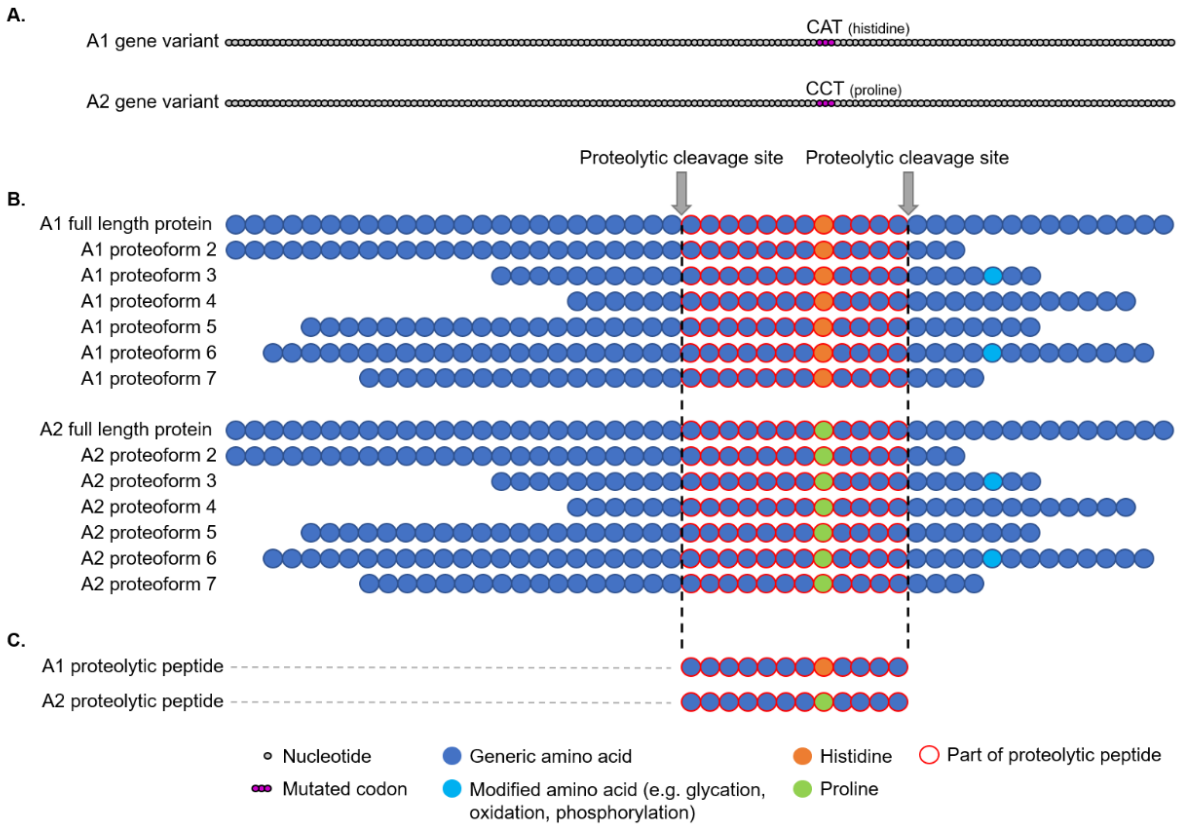

Figure 1. The complexity of A1 and A2 β-CN gene variants and proteoforms. A: A1 and A2 β-CN gene sequences have a single nucleotide difference that results in a proline or histidine during translation. B: Eight example proteoforms of A1 β-CN protein are shown, demonstrating the impact of sequence truncation and modifications (such as glycation, oxidation, phosphorylation, etc.). The A1 β-CN protein may be expected to present as a similar set of eight proteoforms. In an ideal scenario, the region of the protein containing the amino acid change can be proteolytically isolated using enzymes. C: The complex set of A1 and A2 β-CN proteoforms is reduced to two peptides that can be quantified to provide absolute concentration measurements when combined with synthetic peptides and isotopically labelled standards.