

Assessment of Recovery of Milk Protein Allergens from Processed Food for Mass Spectrometry Quantification

KATE GROVES, ADAM CRYAR, MICHAEL WALKER, and MILENA QUAGLIA¹

LGC Limited, Queen's Rd, Teddington, Middlesex TW11 0LY, United Kingdom

Assessing the recovery of food allergens from solid processed matrixes is one of the most difficult steps that needs to be overcome to enable the accurate quantification of protein allergens by immunoassay and MS. A feasibility study is described herein applying International System of Units (SI)-traceably quantified milk protein solutions to assess recovery by an improved extraction method. Untargeted MS analysis suggests that this novel extraction method can be further developed to provide high recoveries for a broad range of food allergens. A solution of α -casein was traceably quantified to the SI for the content of α -S1 casein. Cookie dough was prepared by spiking a known amount of the SI-traceable quantified solution into a mixture of flour, sugar, and soya spread, followed by baking. A novel method for the extraction of protein food allergens from solid matrixes based on proteolytic digestion was developed, and its performance was compared with the performance of methods reported in the literature.

Food allergy is an increasingly important public health problem, with implications for the general health status, economy, and legislation of a country. Currently, European rules for prepacked and certain non-prepacked foods establish a list of 14 food allergens, which must be disclosed by referencing the allergen source whenever they, or ingredients made from them, are formulated for use in a food, including when a list of ingredients need not be given (e.g., alcoholic drinks; Regulation No. 1169/2011, Annex II). The list consists of cereals containing gluten, crustaceans, molluscs, eggs, fish, peanuts, certain nuts, soya, milk, celery, mustard, sesame, lupin, and sulfur dioxide. The majority of these food allergens are proteins with a defined structure in their native form that can induce allergic reactions in sensitized consumers. The two most common food allergies in childhood are to milk and egg proteins. Currently, there is no generally accepted cure for food allergy, and avoidance of the triggering protein(s) is the main clinical management regime. Risk assessment and risk management in aid of avoidance depend heavily on the detection

and quantification of food allergens at clinically relevant concentrations. However, all current analytical approaches suffer from drawbacks (1). Herein, we report the preparation of a solution of milk protein fractions of α -S casein and the metrological quantification of α -S1 casein by exact-matching isotope dilution MS (EM-IDMS) and

- its application to assess extraction recoveries, including improved recovery from solid matrixes by an independently developed direct tryptic proteolysis, which shows promise as a generic extraction method for MS analysis; and
- the characterization of the extracted peptides and their potential modifications.

Allergen Measurement

Methods for detecting food allergens include immunoassays, particularly ELISA, MS, and PCR methods, among which commercial ELISAs are currently the most commonly used. In ELISA, the specificity of the assay antibodies and, therefore, their ability to detect allergenic proteins, primarily depends on how the antibodies have been raised and what epitopes they react to. These are, therefore, likely to be major causes of measurement variability. Furthermore, from an analytical point of view, modifications of the tertiary structure of the allergen protein induced by food processing may affect the antibody recognition (2–7).

LC-MS-based methods are available for the measurement of allergen proteins (8–18), with many based on detecting and quantifying peptides following proteolytic digestion and the use of isotopically labeled internal standards (ISs) corresponding to a portion of the sequence of the protein to be detected/quantified (19, 20). LC-MS-based methods offer a number of advantages over ELISA and PCR as they can be made specific to the protein to be quantified (11), provide protein identity information, permit a wider linear dynamic range, are less prone to be affected by food processing, and, if appropriately applied, can be used as reference methods or for the production of certified reference materials. However, LC-MS methods remain as vulnerable as other approaches in the following crucial respects (1):

- Poor extraction efficiency (stemming from thermal- and processing-induced protein insolubility and protein reaction with matrix compounds, such as tannins and polyphenols).
- Difficulty in determining recovery owing to a lack of reference methods and reference materials.
- Lack of metrological traceability to secure the standardization and harmonization of results.

Milk Allergens

Milk allergy is an adverse immunological reaction to milk proteins of different mammalian species, and it is one of the most common and widespread food allergies, particularly in

¹ Corresponding author's e-mail: milena.quaglia@lgcgroup.com
Guest edited as a special report on "A Global Reflection on Food Allergen Regulations, Management, and Analysis" by Carmen Diaz-Amigo and Bert Popping.

Supplemental Information is available on the *J. AOAC Int.* Web site, <http://aoac.publisher.intentaconnect.com/content/aoac/jaoac>

The research leading to the results presented herein was funded by the Government Chemist Programme, which is a work funding program of the U.K. National Measurement Office.

DOI: <https://doi.org/10.5740/jaoacint.17-0214>

children. Cow's milk contains several proteins that could each, in principle, elicit an allergic reaction in a sensitized individual. The proteins present in milk are subdivided into two major groups: caseins (about 80%) and whey proteins (20%; 21, 23). The World Allergy Organization defines α -casein as the major milk allergen; α -S1 casein represents 80% of α -caseins (22). Immunoassay platforms are currently routinely applied to the analysis of milk allergens. However, a number of MS-based methods have been recently developed for the characterization and quantification of milk protein allergens in solutions and food matrixes (21, 23–39).

In order to facilitate the calculation of the recovery of milk protein allergens from solid matrixes and to assess the suitability of the methods currently available to underpin legislation and support risk management using allergen thresholds, we describe herein the preparation of a prototype reference material containing a known amount of α -S1 casein traceably quantified to the International System of Units (SI) by EM-IDMS. The material was used for the assessment of the efficiency of extraction protocols reported in the literature and aided the development of an extraction method based on the tryptic digestion of solid matrixes and heating cycles to denature proteins and facilitate solubilization. This method could also be used to check for efficiency in extraction and could be developed further as a generic extraction procedure for MS experiments. With further work, the final goal will be the addition of known amounts of SI-traceable quantified milk proteins in a food matrix to mimic the presence of milk (19, 40, 41).

Experimental

Materials

Tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), urea, ammonium bicarbonate (NH_4HCO_3), lyophilized α -S casein from bovine milk (250 mg, $\geq 70\%$ α -S casein basis), and tosyl phenylalanyl chloromethyl ketone-treated lyophilized trypsin from bovine pancreas were purchased from Sigma-Aldrich (Gillingham, United Kingdom). Sodium dodecyl sulfate (SDS) 20% was from MP Biomedicals (Loughborough, Fisher Scientific, United Kingdom). LC-MS grade acetonitrile and methanol were from LGC Standards (Teddington, United Kingdom). Sodium hydroxide, glacial acetic acid, and formic acid (98/100%) were all of analytical reagent grade and purchased from Fisher Scientific (Loughborough, United Kingdom). The water was ultra-pure Elga ($18.0 \text{ M}\Omega \text{ cm}^{-1}$). The peptides used for quantitation of α -S1 casein were T3 (HQGLPQEV-LNENLLR) and the isotopically labelled T3 (T3*) [HQGLPQEV-(Leu($^{13}\text{C}_6$, ^{15}N))-NEN-(Leu($^{13}\text{C}_6$, ^{15}N))-LR], T4 (FFVAPFPEVFGK) and the isotopically labelled T4 (T4*) [FFVAPF-(P($^{13}\text{C}_5$, ^{15}N))EV-(F($^{13}\text{C}_9$, ^{15}N))-GK], and T11 (YLGYLEQLLR) and the isotopically labelled T11 (T11*) [YLGYLE-(L($^{13}\text{C}_6$, ^{15}N))-EQ-(L($^{13}\text{C}_6$, ^{15}N))-LR] and were custom synthesized by Biopolymers (Thermo Fisher Scientific, Ulm, Germany). Peptides were delivered in vials containing peptides at a nominal concentration of $50 \pm 2.5 \text{ nmol/g}$ in water–acetonitrile (95 + 5, v/v). These solutions were aliquoted into individual 50 μL aliquots and stored at -20°C for single use. The concentration of the peptides was verified by amino acid analysis and EM-IDMS (18).

Cookie prototype materials were prepared by mixing Tesco Everyday Value plain flour, Tate & Lyle granulated cane sugar, and Pure dairy free soya spread. A preliminary selection of the peptides to be used for quantification was performed by using the

following commercially available cookies: Co-Operative All Butter Shortbread Biscuits and M^cVitie's Original Digestives (both containing milk) and Lyon's Rich Tea Biscuits (indicating "may contain milk"). Additionally, another commercially available biscuit, Co-Operative White Chocolate and Blueberry Biscuits, and a Tesco instant dried skim milk powder were also used to evaluate extraction recoveries.

The expanded uncertainties, corresponding to volume pipets of 2, 20, 100, 200, and 1000 μL , were used throughout experimentation: ± 0.02 , ± 0.12 , ± 0.15 , ± 1.8 , and $\pm 2.9 \mu\text{L}$, respectively. The appropriate pipets were used in order to minimize the uncertainty of each volume being pipetted. The expanded uncertainty associated with the balance used for the weighing of materials was $\pm 0.000284 \text{ g}$.

Apparatus

For in-house cookie preparation, a food processor (FF586 Multipro; Kenwood) and a fan oven (Heraeus; Thermo Fisher Scientific) were used. For sample preparation, a vacuum centrifuge (AVC 2-22; Christ), an incubator shaker (Camlab, Microtherm), and a 24-place microliter centrifuge (Biofuge Pico; Heraeus) were used. For LC-tandem MS (MS/MS) analysis, an Agilent Technologies Q-TOF 6530 coupled with Agilent nano-HPLC chip technology (Agilent Technologies, Santa Clara, CA) was used and for semiquantitative LC-MS^E analysis tandem mass spectrometry acquisition using alternating low-energy-induced dissociation and high energy induced dissociation, a Xevo G2-XS Q-TOF coupled to an ACQUITY UPLC H-Class system (Waters, Milford, MA) was employed. For targeted LC-MS/MS quantification, an Agilent 6490 triple quadrupole coupled with the Agilent nano-HPLC chip technology (Agilent Technologies) was used.

LC-MS Methods

An Agilent Technologies Q-TOF 6530 coupled with Agilent nano-HPLC chip technology (Agilent Technologies) was used for the characterization of the customized peptides, the optimization of the α -S1 casein proteolytic digestion, and the analysis of the proteolytic extracts from the commercially available cookies. The separation of the peptides was carried out on an Agilent ProtID-Chip-150 (II), an analytical column (150 mm $\times$ 75 μm), and an enrichment column. Mobile phases were 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The following gradient was applied: 0–2 min, 5–10% B; 2–30 min, 10–50% B; 30–31 min, 50–98% B; 31–35 min, 98% B; and 35–36 min, 98–5% B. The flow rate of the nanopump was set to 600 nL/min and the capillary pump was set to 4 $\mu\text{L}/\text{min}$. The injection volume was 1 μL . Data were acquired in positive-ion mode and analyses were performed both in full-scan (m/z 100–1200) and in MS/MS mode for the identification of potential impurities and to monitor modifications throughout digestion/extraction. The capillary voltage was 1950 V. The drying gas was set to 325°C at 4 L/min. LC-MS/MS experiments were carried out only on precursor ions with +3, +2 and +1 charges between m/z 300 and 1200, using a fragmentor voltage of 170 V. The collision energy (CE) applied was dependent on the m/z of the precursor ion, and the following formula was applied: $\text{CE} = [\text{slope} \times (m/z)]/[100 + \text{offset}]$, with the slope and offset set to 2.5 and -1 , respectively. Only product ion spectra within the range of m/z 50–2500 were acquired.

Quantification experiments were performed in positive-ion mode on the Agilent 6490 triple quadrupole coupled with Agilent

nano-HPLC chip technology. LC conditions and columns were as described above. The source gas was set to 350°C and a flow rate of 11 L/min was applied. Nitrogen was used for all gases. Selected-reaction monitoring experiments were optimized to monitor the MS ratio of unlabeled to labeled peptides. Final transitions and optimum conditions for each transition are reported in Table 1. The selection of the optimal transitions and CE was facilitated by using Skyline Version 1.4 software (MacCoss Laboratory, University of Washington; 42). The conditions were further verified and optimized experimentally. The dwell time was set to 40 ms, the cell acceleration potential set at 5 V, and the resolution for MS1 and MS2 scans were set to wide (1.2 amu) and unit (0.7 amu), respectively.

The Xevo G2-XS Q-TOF coupled to an ACQUITY UPLC H-Class system was used for the qualitative and semiquantitative analysis of the proteins extracted from the cookies incurred with SI-traceable quantified casein. Peptide solutions were chromatographically separated on a CORTECS C18+ column (3.0 × 150 mm, 2.7 µm; Waters). Peptides were separated across a 30 min linear binary gradient, starting at 3% acetonitrile–0.1% formic, increasing up to 40% acetonitrile–0.1% formic. The flow rate was set to 0.25 mL/min throughout the run, with a column temperature of 35°C. The mass spectrometer was operated in MS^E mode, which has been described in detail previously (43). Mass measurements were made from *m/z* 50 to 2000 while the quadrupole was operated such that the transmission of ions between *m/z* 400 and 2000 was optimal. Low-energy measurements were made with a CE of 4 eV, whereas high-energy measurements were made with a CE ramp of 15–45 eV. The scan time was set to 0.6 s, with an interscan delay of 0.05 s, meaning that a complete low- and high-energy cycle was completed every 1.3 s (44). For lock mass correction, a 100 fmol/µL [Glu1]fibrinopeptide B solution was infused through the reference fluidics system and sampled every 60 s.

Preparation of the Stock Solution of α-S1 Casein

A stock solution of α-S1 casein was prepared gravimetrically by diluting 100 mg lyophilized α-S casein in 10 mL water. The stock solution was mixed on a vortex mixer thoroughly and left to equilibrate at room temperature for 40 min. Aliquots of 550, 120, and 70 µL were prepared using the stock solution and stored at

–20°C, to be thawed for single usage. Aliquots of the thawed frozen solution were gravimetrically diluted 1:1000 and SI-traceably quantified as described below.

SI-Traceable Quantification of an Aqueous Solution of α-S1 Casein

SI-traceable quantification of the aqueous solution of α-S1 casein was performed by EM-IDMS, using isotopically labeled peptides as ISs (19, 20). Five sample blends (SBs), containing 500 ng protein standard and equimolar amounts of labeled peptides, were gravimetrically prepared plus five calibration blends (CBs), containing equimolar amounts of unlabeled and labeled peptides. Labeled peptides added to the blends were prepared so as to have equivalent mass fractions in all blends and to achieve an observed MS area ratio of 1:1 (unlabeled + labeled peptide). After gravimetric preparation, the blends were frozen at –80°C for 30 min and dried in a vacuum centrifuge at –54°C and 0.7 mbar (AVC 2-22) for 2.5 h.

Tryptic digestion was carried out by dissolving SBs and CBs in 100 µL Tris-HCl buffer (500 mM, pH 7.5). Samples were mixed on a vortex mixer and incubated at 60°C under shaking conditions for 30 min. Finally, samples were removed from the incubator and allowed to cool to room temperature for 15 min. Repeat 3 µL additions of 1 µg/µL trypsin, solubilized in 1% (v/v) aqueous acetic acid, were made at time points of 0, 15, 30, 60, 120, and 150 min while incubating at 37°C. For EM-IDMS quantification, SBs were injected five times each and bracketed by CBs. Quantification results plus standard and combined uncertainties were determined by application of the EM-IDMS equation (19, 20).

Preparation of Prototype Biscuit Reference Materials

Two α-S1 casein cookies and one blank cookie (cookie spiked with water instead of aqueous casein solution) were prepared separately, starting with the blanks to avoid contamination. Each cookie was prepared by gravimetrically weighing 10 g granulated sugar, 10 g soya spread, and 15 g plain flour and adding volumetrically 500 µL water to the mixture (or 250 µL casein solution and 250 µL water). The mixture was transferred to the Kenwood FF586 Multipro food processor and the unbaked material patted into single-form cookie tins in an attempt to make

Table 1. MRM transitions and collision energies optimized for the quantification of unlabeled and labeled α-S1 casein peptides for both the unlabeled and labeled forms of the peptide standards used for the quantification of α-S1 casein

Peptide	Amino acid sequence	MW, Da	Q1, <i>m/z</i> ^a	<i>z</i> , + ^b	Q3, <i>m/z</i> ^c	Fragment	CE, V ^d
T3	HQGLPQEVLNENLLR	1761.1	587.3	3	758.4	y6	16.3
T3*	HQGLPQEVL*NENL*LR ^e	1774.7	592.0	3	765.4	y6	16.3
T4	FFVAPFPEVFGK	1385.6	692.9	2	920.5	y8	20.1
T4*	FFVAPFP**EVF***GK ^{f,g}	1401.1	701.4	2	936.5	y8	20.1
T11	YLGYLEQLLR	1268.2	634.4	2	991.6	y6	18
T11*	YLGYL*EQL*LR	1282.2	641.4	2	1005.6	y6	18

^a Q1 = Precursor ion.

^b *z* = Charge.

^c Q3 = Product ion.

^d CE = Collision energy.

^e * = Isotopically labeled amino acid Leu(¹³C₆, ¹⁵N).

^f ** = Isotopically labeled amino acid Pro(¹³C₅, ¹⁵N).

^g *** = Isotopically labeled amino acid Phe(³C₉, ¹⁵N).

the material as uniform as possible across the cookie. The cookies were then baked at 180°C for 18 min using the Heraeus fan oven. Following baking, the cookies were allowed to cool and stored separately in sealed bags at room temperature. Cookies were weighed before and after baking, with average weights of 34.9 ± 0.1 and 28.3 ± 0.3 g, respectively. The same procedure was followed for the preparation of the α -S1 casein cookies, with the only variation being the addition of 250 μ L casein stock solution and 250 μ L water instead of 500 μ L water to obtain cookies containing approximately 60 μ g/g α -S1 casein. The volume of aqueous solution to be added was optimized to minimize absorption of casein on the walls of the food processor during mixing.

Extraction of α -S1 Casein from Prototype Cookies

α -S1 casein was extracted from all cookies by using both an extraction procedure reported in the literature by Ansari et al. (36), in which extraction occurs at the intact protein level, and a method developed in-house based on tryptic digestion and consecutive heating cycles to facilitate protein denaturation and extraction from the solid matrix. All cookies were separately blended in the Kenwood FPP225 food processor and four 0.18 g aliquots per each prespiked cookie (total of eight aliquots) and four aliquots of blank cookies were weighed.

For the Ansari et al. extraction method, two aliquots from the blank blended cookies were spiked with a solution of α -S casein to achieve a concentration of approximately 60 μ g casein/g cookie. The four prespiked cookie aliquots were spiked with an equivalent volume of water; 1.8 mL of 50 mM NH_4HCO_3 (pH 9.6) were added to each 0.18 g material and the samples heated at 60°C for 30 min. After heating, samples were centrifuged at $4000 \times g$ for 20 min at room temperature using the Biofuge Pico 24-place microliter centrifuge. Following this, 250 μ L supernatant was pipetted into a low-binding polypropylene tube (Eppendorf). Samples were dried using a vacuum centrifuge, further solubilized in 106 μ L of 10 mM NH_4HCO_3 (pH 7.6), heated at 60°C for 30 min in the presence of urea, and tryptic digested overnight after addition of 10 μ g trypsin. Finally, samples were diluted 1:2 with water and analyzed by LC-MS/MS as reported above.

For the proteolytic digestion extraction procedure developed in-house, two 0.18 g blended biscuits were placed in Eppendorf tubes and 1.8 mL extraction buffer (50 mM NH_4HCO_3 , 1 M urea, and 0.002% SDS in water, pH 8) were added. Twenty microliters of a 1:10 dilution of a SI-traceable solution of α -S1 casein was gravimetrically added to two blank blended cookie aliquots with extraction buffer to produce a 0.1 g/g solution (biscuit/extraction buffer) and 20 μ L water added to the prespiked samples. The extraction of α -S1 casein was carried out through five heating-cooling cycles, each followed by addition of trypsin. Samples were heated at 60°C for 30 min and allowed to cool to room temperature for 15 min before adding 3 μ L of a solution of 1 μ g/ μ L trypsin in 1% acetic acid solution, followed by 30 min incubation at 37°C. After five cycles, samples were centrifuged for 20 min and the supernatant removed and further centrifuged for 10 min. Two hundred fifty microliters of the supernatant were frozen overnight at -80°C and subsequently dried in a vacuum centrifuge at -54°C at 0.7 mbar. The samples were resolubilized in 112 μ L buffer (50 mM NH_4HCO_3 and 1 M urea, pH 8) and heated at 60°C for 30 min. Samples were allowed to cool to room

temperature before final addition of 3 μ g trypsin and further incubation at 37°C for 30 min. Finally, the samples were diluted 1:2 with water and analyzed by LC-MS/MS (Agilent 6490).

Each sample was injected three times and the average results of all prespiked and all postspiked data used to calculate the recovery for each peptide (T3, T4, and T11). Errors were calculated as the square root of the sum of average RSD squares obtained by multiple injections of the same sample.

LC-MS/MS Data Processing

LC-MS/MS data acquired on the Agilent Technologies Q-TOF 6530 were exported in .mgf format from within the MassHunter workstation software. These files were then searched against a database using a Mascot search engine. The Mascot search parameters were as follows: SwissProt database, peptide tolerance of 10 ppm, peptide fragment mass tolerance of 0.4 Da, 2 missed cleavages, variable modification oxidation, and deamidation.

Data acquired on the Xevo G2-XS Q-TOF were processed on a ProteinLynx Global Server 3.2 (Waters). Apex3D peak picking was carried out with the following parameters: automatic chromatographic peak width, automatic MS time-of-flight resolution, lock mass for charge +2 of 785.8426 Da/e, low-energy threshold of 135 cts, elevated energy threshold of 30 cts, automatic retention time window, and an intensity threshold of the precursor/fragment ion cluster of 750 cts. Resultant peak lists were searched against a fasta file consisting of four concatenated UniProt protein databases: *Bos taurus* (24 148 entries), *Gallus gallus* (29 485 entries), *Triticum aestivum* (136 884 entries), and *Glycine max* (66 205 entries). For all searches, a minimum of three fragment ions per peptide, five fragments per protein, and one peptide per protein were required for identification to be retained. Trypsin was set as the primary digest reagent and a maximum of two missed cleavages allowed. No fixed modifications were set and oxidation and carbamidomethylation were set as variable modification. A maximum false-discovery rate of 4% was maintained by searching against a reverse database. Semiquantitative data were extracted from exported .csv ion accounts, and Top3 protein intensity was used as a proxy for protein abundance.

Results and Discussion

Selection of Internal Standard Peptides for Quantification of α -S1 Casein

The SI-traceable quantification of proteins can be performed by applying EM-IDMS and using proteins or peptides as ISs (18, 19, 39, 41). In this work, appropriate peptides were selected for the SI-traceable quantification of α -S1 casein, and SI traceability was maintained through amino acid quantification of the peptides, as reported by Cryar et al. (19). An in silico tryptic digest of α -S1 casein was performed to assess the suitability of trypsin to generate peptides to be used for quantification. The sequences of the most appropriate peptides were searched against the Basic Local Alignment Search Tool (Swissprot Database; blastp algorithm) for their unicity. The following peptides fulfilled the above criteria: T3 (HQGLPQEVLNENLLR), T4 (FFVAPFPEVFGK), T10 (EDVPSER), T11 (YLGYLEQLLR), T15 (VPQLEIVPNSAEER), and T17 (EGIHAAQK). To experimentally identify the most

appropriate peptides to be used as ISs, milk proteins were extracted from milk, skim milk, buttermilk cookies, cookies containing milk, and “may contain milk” cookies. The extracts were tryptic digested and analyzed by LC-MS/MS. In Table 2, the peptides that were identified in the individual samples by processing the data acquired by LC-MS/MS by Mascot are presented. Three peptides (T3, T4, and T11) were detected in all the samples, excluding the extracts from the “may contain milk” cookies, and were selected as ISs for the SI-traceable quantification of α -S1 casein.

Optimization of Tryptic Digestion for SI-Traceable Quantification of α -S1 Casein

A requirement for SI-traceable quantification of proteins by using proteolytic digestion and isotopically labeled peptides as ISs is the complete release of the peptides used for quantification from the protein to be quantified (19, 20). Three sets of proteolytic digestion time course experiments were performed to identify the ideal conditions of release of the peptides of interest from the protein. For each experiment, nine aliquots of α -S1 casein containing equimolar amounts of isotopically labeled peptides and three aliquots containing equimolar amounts of natural T3, T4, and T11 peptides and their isotopically labeled equivalents were prepared, frozen at -80°C for 30 min, and dried by using a vacuum centrifuge. Samples were resolubilized in Tris-HCl buffer and incubated at 37°C under shaking conditions for 6 h. Repeat additions of $3\text{ }\mu\text{g}$ trypsin [solubilized in 1% (v/v) acetic acid] were performed hourly in the first set of experiments and every 30 min in the second set of experiments. In the third set of experiments, samples were heated at 60°C for 30 min before digestion, and trypsin additions were performed at 0, 15, 30, 60, 120, 150, and 180 min during incubation at 37°C . Before each trypsin addition, one aliquot was removed, diluted 1:10 in water, and analyzed by LC-MS/MS to check the release of the peptide from the protein. An increase of the unlabeled-labeled MS area ratio showed an increased release of the peptide from the protein. A constant ratio over the time and for the three peptides was indicative of the complete release of the peptides from the protein. A decrease in the area of the labeled peptide combined with an increased unlabeled-labeled ratio suggested degradation of the peptide throughout digestion. In Figure 1, the results from the third set of experiments are reported. The complete release of the three peptides monitored was considered achieved after 60 min.

SI-Traceable Quantification of α -S1 Casein

Two of the three peptides selected were used to provide quantitative data for the SI-traceable quantification of α -S1 casein. The T4 peptide appeared unstable over time when diluted to a low concentration and was used to monitor only the consistency of the digestion.

Five SB-containing α -S1 casein and equimolar amounts of isotopically labeled peptides standards and three CB-containing natural and isotopically labeled peptides were prepared, tryptic digested, and analyzed as previously described. Figure 2 shows the quantification values for T3 and T11 peptides obtained by EM-IDMS. These were averaged to give the final SI-traceable quantification value of the aqueous solution of α -S1 casein. By considering the α -S1 casein MW equal to $22\,974.588 \pm 0.479\text{ g/mol}$,

Table 2. α -S1 tryptic casein peptides identified in two aliquots of cookies containing milk (Bs1 and Bs2) and cookies containing milk spiked with casein (BsSP1 and BsSP2), two aliquots of “may contain milk” cookies (BsN1 and BsN2), and “may contain milk” cookies spiked with casein (BsNSP1 and BsNSP2), two aliquots of buttermilk cookies (BBs1 and BBs2) and buttermilk cookies spiked with casein (BBsSP1 and BBsSP2), and two aliquots of milk and skim milk

α -S1 casein tryptic peptides	Amino acid sequence	Bs1	Bs2	BsSP1	BsSP2	BsN1	BsN2	BsNSP1	BsNSP2	BBs1	BBs2	BBsSP1	BBsSP2	Milk 1	Milk 2	Skim milk 1	Skim milk 2
T3	HQGLPQEVLENLLR	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
T4	FFVAPFPEVFGK	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
T6	VNELSK			X										X	X	X	X
T5-T6	EKVNELSK													X	X	X	X
T7	DIGSESTEDQAMEDIK													X	X	X	X
T10	EDVPSEK													X	X	X	X
T9-T10	HIQK EDVPSEK													X	X	X	X
T11	YLGYLEQLLR	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
T14-T15	YKVPQLEIVPNSAEER													X	X	X	X
T16	LHSMK													X	X	X	X
T17	EGIAQK													X	X	X	X
T18	ERMIGNQELAYFPFLFR													X	X	X	X
T19	QFYQLDAYPSGAWYYVPLGTQYT DAPSFSDIPNIGSENSEK													X	X	X	X

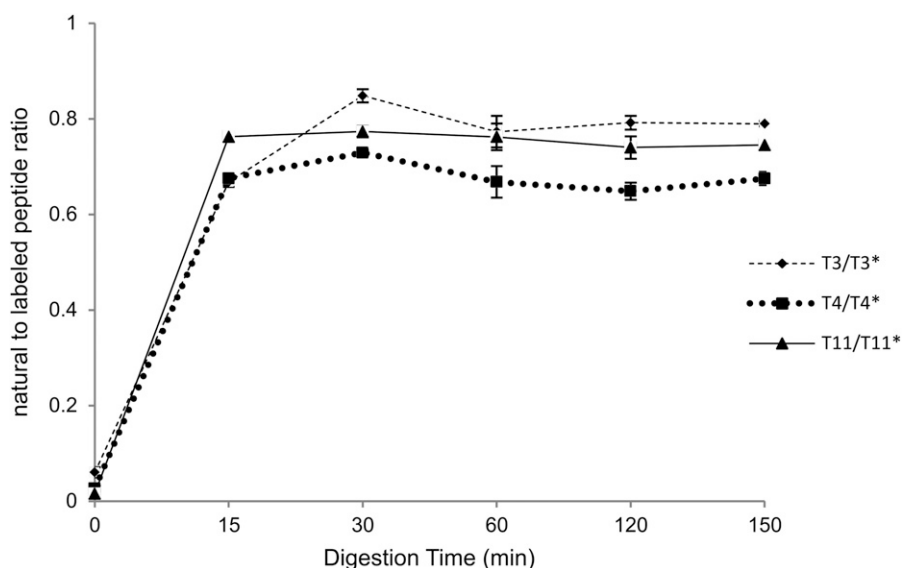


Figure 1. Time course digestion experiments showing the measured MS area ratios of the unlabeled peptide (Tx) released from the protein during digestion and labeled peptide IS (T*) of the three peptides (T3, T4, and T11). The values were normalized for gravimetric preparation. Following an initial stage of heated denaturation at 60°C for 30 min, repeat additions of trypsin were made to a 500 ng solution of α -S1 casein spiked with equimolar isotopically labeled peptide. Prior to each addition of enzyme, the ratio of digested protein was measured against the labeled IS. Error bars represent the SDs of repeat measurements ($n = 3$).

based on International Union of Pure and Applied Chemistry reference masses for elemental weights (45), the 1:1000 aqueous dilution of the stock solution of α -casein was SI-traceably quantified for α -S1 to 6.989 ± 0.793 $\mu\text{g/g}$, corresponding to an expanded uncertainty of 11.34%. The expanded uncertainty was calculated as reported by Cryar et al. (19).

Extraction of α -S1 Casein and Other Food Allergens

The extraction of α -S1 casein from the in-house and the commercially available cookies was performed by applying the Ansari et al. extraction method and the procedure developed in-house previously described. The recovery of

α -S1 casein was calculated by comparing the amount of casein extracted from the cookies containing a known amount of α -S1 casein with the amount of casein spiked in the blank cookies that had been subjected to the same extraction procedure. The recoveries of α -S1 casein from the in-house baked cookies by using the Ansari et al. method for T3, T4, and T11 were 11 ± 3 , 7 ± 1 , and $10 \pm 1.4\%$. The recoveries were from four aliquots from two cookies and, therefore, the SDs of the extracted amount of casein were representative of the homogeneity of the dough used to prepare the cookies and the variability between cookies. No IS were used to account for the variability of the instrument.

To increase the recovery of α -S1 casein and to better understand the data obtained, preliminary experiments were

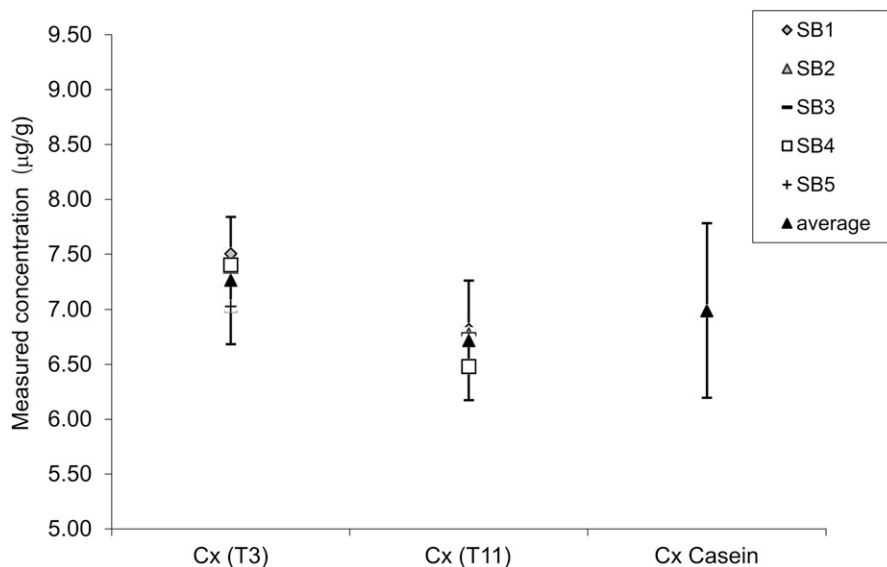


Figure 2. Results of the EM-IDMS quantification of an aqueous solution of the milk allergen, α -S1 casein. For five repeat sample blends, individual concentrations were calculated for T3 and T11 peptides. The final quantification value for α -S1 casein was taken as an average across blends and peptides. Error bars represent the expanded measurement uncertainty of each individual concentration calculation ($k = 2$).

performed in which the ground matrix samples were subjected to multiple cycles of heated extraction, followed by tryptic digestion. This workflow is presented in Figure 3. Using this approach, it was hoped that the amount of protein/peptide extracted from the solid material would be maximized, hence increasing protein recovery. The postspiking of the SI-traceably quantified solution to the blank cookie prior to the heating cycles applied for extraction allows for any losses due to absorption of the protein onto any surface or due to protein instability. The recoveries for T3, T4, and T11 calculated by using the proteolytic digestion extraction procedure developed in-house were 26 ± 4 , 26 ± 4 , and $50 \pm 5\%$. Overall increased recovery was observed for the proteolytic digest method, as previously reported in the literature (46). The recovery calculated for the T11 peptide was higher than the recovery calculated for the T3 and T4 peptides, probably due to difficulties in extracting the T3–T4 region of the protein from the solid matrix. Obviously, this was not observed when the protein was extracted as intact by using the Ansari et al.

method. The same extraction procedures were also applied to the commercially available cookies and reconstituted milk powder before analysis by LC-MS/MS. The peak areas for the T3, T4, and T11 peptides extracted from the milk powder were comparable with the commercially available extracted cookies. It was possible to observe that also in this case, by using the proteolytic extraction procedure developed in-house, T11 peptide recovery from the commercially available cookies was higher than the recovery for the T3 and T4 peptides. Although losses of the SI-traceable quantified casein solution could have occurred during cookie preparation and, therefore, negatively bias the calculation of the recoveries, the improved values obtained when proteolytic digestion was applied directly to the matrix show a large margin of potential improvement.

In order to check that modifications did not bias the calculation of the recovery for the peptides monitored, the extracts were analyzed by LC-MS^E on the quadrupole time-of-flight apparatus.

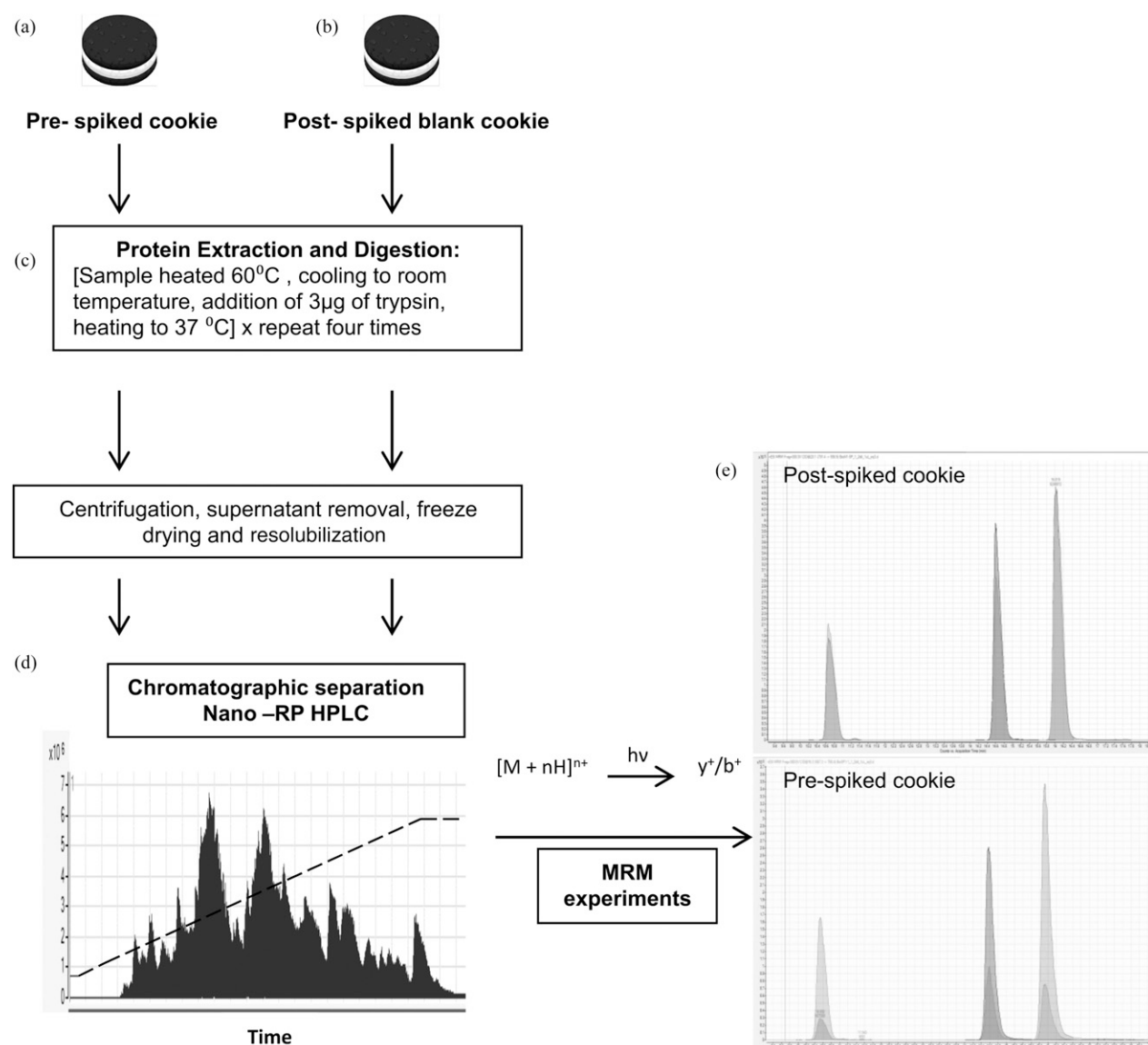


Figure 3. Optimized workflow for the evaluation of the extraction of the milk allergen, α -S1 casein, from an in-house-prepared solid biscuit matrix. Nano-RP HPLC is nano reverse phase high performance liquid chromatography.

Table 3. Overview of results from the LC-MS^E analysis of cookie extractions^a

Parameter	Ansari et al. method		In-house method	
	Sample 1	Sample 2	Sample 1	Sample 2
No. of proteins	10	11	100	99
No. of peptides	56	67	1413	1501
α-S1 casein sequence coverage, %	32.2	32.2	76.6	97

^a Compared are the number of protein and peptides successfully identified using a previously published protocol (Ansari et al.; 46) that had been optimized for casein protein from cows milk and an in-house developed method that uses multiple heat denaturation step. Also included is a comparison of the sequence coverage of α-S1 casein (most abundant cow's milk casein) obtained using the two extraction methods.

All extracts from the in-house-developed protocol showed a significant degree of carbamylation. However, for the three peptides monitored, carbamylation was consistently 5–6% of the total amount of peptide. Deamidation and oxidation were also observed across the sequence, but not on the peptides of interest.

LC-MS^E data were searched against a protein database containing sequences of wheat, soya, chicken, and cow proteins to evaluate the potential of the digest protocol to be used as generic method for the extraction of food allergens for MS analysis. Analysis of protein and peptide identities clearly demonstrate that although the Ansari et al. method is able to efficiently extract casein proteins, this is not the case for other proteins contained within the cookie tested. On average, only 11 proteins were successfully identified with the Ansari et al. extraction method (Table 3). In contrast, the in-house method successfully identified 1 order of magnitude more protein, with an average identification of 100 proteins per sample run. Importantly, proteins identified included proteins derived from the well-established allergens of egg, milk (other than caseins), and wheat (gluten). The details for all protein identification can be found in the Supplemental Information. When peptide numbers are considered, the differences between the two methods are even more pronounced, with the Ansari et al. and

in-house methods identifying an average of 61 and 1457 peptides, respectively. Interestingly, despite the Ansari et al. method extraction process being visibly more biased toward the extraction of caseins, the sequence coverage obtained for these proteins was far lower than that achieved by the in-house method (Table 3). By using well-established Top3 protein intensity as an estimate for protein abundance, it can be seen from the weighted scatterplot in Figure 4 that even for proteins with a higher sample relative abundance in the Ansari et al. method, the in-house results provided consistently higher protein sequence coverage. This can be attributed to the absence of multiple denaturation steps within the Ansari et al. method in which efficient protein denaturation has been previously shown to be a key component in protein extraction and digestion (47, 48). The in-house method conversely uses sequential heat denaturation steps, followed by tryptic digestion, to ensure the extraction of the incurred proteins. The work described here aims to evaluate the extraction efficiency of current procedures for α-S1 casein, a protein which does not contain sulfur bridges and thus no reduction and alkylation step was performed. It is reasonable to believe that by introducing this step, further improvement in the extraction of food allergen proteins from solid matrixes can be achieved. However, this was considered outside the scope of the present work, which focused on α-S1 casein.

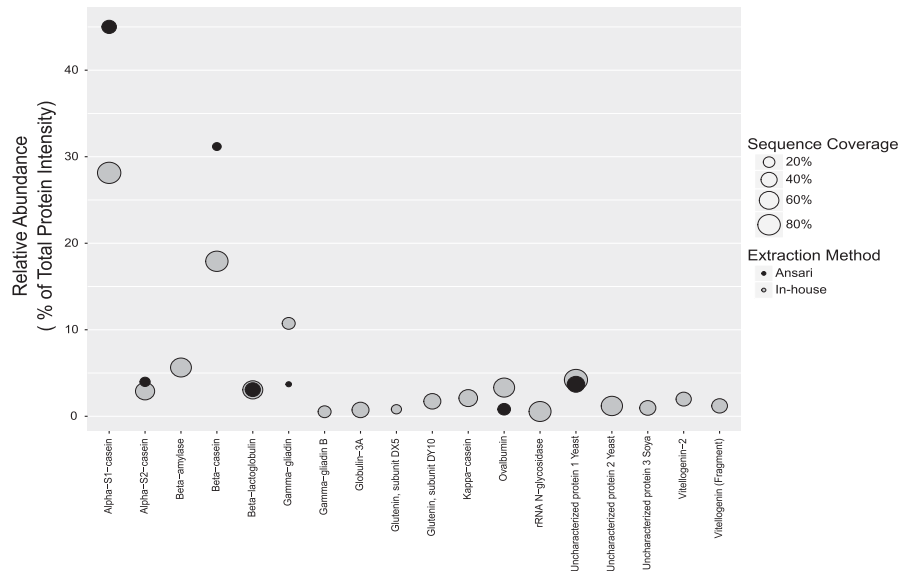


Figure 4. Weight scatterplot of the semiquantitative results from the LC-MS^E analysis of store-bought cookie extracted with the Ansari et al. extraction method and our in-house method. The relative abundance (y-axis) and sequence coverage (plot size) obtained by the two extraction methods are compared with the 18 most abundant proteins (as determined from the in-house method). Results from the Ansari et al. and in-house methods are represented by black and gray plots, respectively.

Conclusions

The work here described highlights the problems associated with the detection and quantification of protein allergens in solid food matrixes by MS and provides solutions through the development of a method for improved recovery that could potentially be used to monitor extraction efficiency and be further developed as a method that is applicable to a broader range of matrixes and food allergens. Here, we also described a feasibility study in which we spiked SI-traceable quantified proteins into food matrixes of baked cookies containing known amounts of protein. These materials may not only be used for method development of food allergen protein extraction procedures, but could also be used to investigate the changes in protein allergens that were induced by food processing and the consequences on quantification results.

Authors Contributions

Kate Groves performed the majority of the LC-MS experiments and prepared and analyzed the cookies; Adam Cryar performed the LC-MS^E experiments and analyzed the cookies; Michael Walker jointly conceived and was the coordinator of the project and involved in stakeholder consultation; and Milena Quaglia jointly conceived the work and was responsible for the experimental design, performed some of the LC-MS experiments, and was heavily involved in stakeholder consultation. All authors critically commented on drafts of the manuscript.

References

- Walker, M.J., Thorburn Burns, D., Elliott, C.T., Hazel Gowland, M., & Mills, E.C. (2016) *Analyst* **141**, 24–35. doi:10.1039/C5AN01457C
- Official Methods of Analysis* (2012) AOAC INTERNATIONAL, Gaithersburg, MD Appendix M: Validation Procedures for Quantitative Food Allergen ELISA Methods: Community Guidance and Best Practices
- Hei, W., Li, Z., Ma, X., & He, P. (2012) *Anal. Chim. Acta* **734**, 62–68. doi:10.1016/j.aca.2012.05.009
- Vieths, S., Barber, D., Chapman, M., Costanzo, A., Daas, A., Fiebig, H., Hanschmann, K.M., Hrabina, M., Kaul, S., Ledesma, A., Moingeon, P., Reese, G., Schömer, C., van Ree, R., Weber, B., & Buchheit, K.H. (2012) *Pharmeuropa Bio. Sci. Notes* **8**, 118–134
- Iqbal, A., & Ateeq, N. (2013) *Food Chem.* **141**, 1651–1654. doi:10.1016/j.foodchem.2013.04.102
- Sakai, S., Adachi, R., Akiyama, H., & Teshima, R. (2013) *J. Agric. Food Chem.* **61**, 5675–5680. doi:10.1021/jf3033396
- Johnson, P.E., Rigby, N.M., Dainty, J.R., Mackie, A.R., Immer, U.U., Rogers, A., Titchener, P., Shoji, M., Ryan, A., Mata, L., Brown, H., Holzhauser, T., Dumont, V., Wykes, J.A., Walker, M., Griffin, J., White, J., Taylor, G., Popping, B., Crevel, R., Miguel, S., Lutter, P., Gaskin, F., Koerner, T.B., Clarke, D., Sherlock, R., Flanagan, A., Chan, C.-H., & Mills, E.N.C. (2014) *Food Chem.* **148**, 30–36. doi:10.1016/j.foodchem.2013.09.115
- Popping, B., & Godefroy, S.B. (2011) *J. AOAC Int.* **94**, 1005
- Heick, J., Fischer, M., Kerbach, S., Tamm, U., & Popping, B. (2011) *J. AOAC Int.* **94**, 1060–1068
- Heick, J., Fischer, M., & Popping, B. (2011) *J. Chromatogr. A* **1218**, 938–943. doi:10.1016/j.chroma.2010.12.067
- Johnson, P.E., Baumgartner, S., Aldick, T., Bessant, C., Giosafatto, V., Heick, J., Mamone, G., O'Connor, G., Poms, R., Popping, B., Reuter, A., Ulberth, F., Watson, A., Monaci, L., & Mills, E.N.C. (2011) *J. AOAC Int.* **94**, 1026–1033
- Carrera, M., Cañas, B., & Gallardo, J.M. (2012) *J. Proteomics* **75**, 3211–3220. doi:10.1016/j.jprot.2012.03.030
- Cucu, T., De Meulenaer, B., & Devreese, B. (2012) *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess.* **29**, 1821–1831
- Di Stefano, V., Avellone, G., Bongiorno, D., Cunsolo, V., Muccilli, V., Sforza, S., Dossena, A., Drahos, L., & Vékey, K. (2012) *J. Chromatogr. A* **1259**, 74–85. doi:10.1016/j.chroma.2012.04.023
- Abdel Rahman, A.M., Kamath, S.D., Gagné, S., Lopata, A.L., & Helleur, R. (2013) *J. Proteome Res.* **12**, 647–656. doi:10.1021/pr300755p
- Monaci, L., Losito, I., De Angelis, E., Pilolli, R., & Visconti, A. (2013) *Rapid Commun. Mass Spectrom.* **27**, 2009–2018. doi:10.1002/rcm.6662
- Uvackova, L., Skultety, L., Bekesova, S., McClain, S., & Hajduch, M. (2013) *J. Proteome Res.* **12**, 4862–4869. doi:10.1021/pr400336f
- Posada-Ayala, M., Alvarez-Llamas, G., Maroto, A.S., Maes, X., Muñoz-Garcia, E., Villalba, M., Rodríguez, R., Perez-Gordo, M., Vivanco, F., Pastor-Vargas, C., & Cuesta-Herranz, J. (2015) *Food Chem.* **183**, 58–63. doi:10.1016/j.foodchem.2015.02.139
- Cryar, A., Pritchard, C., Burkitt, W., Walker, M., O'Connor, G., Burns, D.T., & Quaglia, M. (2013) *J. AOAC Int.* **96**, 1350–1361. doi:10.5740/jaoacint.12-438
- Quaglia, M., & O'Connor, G. (2010) *Chromatogr. Today* **3**, 11–15
- Weber, D., Raymond, P., Ben-Rejeb, S., & Lau, B. (2006) *J. Agric. Food Chem.* **54**, 1604–1610. doi:10.1021/jf052464s
- Fiocchi, A., Brozek, J., Schunemann, H., Bahna, S.L., von Berg, A., Beyer, K., Bozzola, M., Bradsher, J., Compalati, E., Ebisawa, M., Guzman, M.A., Li, H., Heine, R.G., Keith, P., Lack, G., Landi, M., Martelli, A., Rance, F., Sampson, H., Stein, A., Terracciano, L., & Vieths, S. (2010) *World Allergy Organ. J.* **3**, 57–161. doi:10.1097/WOX.0b013e3181defeb9
- Sabbadin, S., Seraglia, R., Allegri, G., Bertazzo, A., & Traldi, P. (1999) *Rapid Commun. Mass Spectrom.* **13**, 1438–1443. doi:10.1002/(SICI)1097-0231(19990730)13:14<1438::AID-RCM663>3.0.CO;2-Z
- Chen, R.-K., Chang, L.-W., Chung, Y.-Y., Lee, M.-H., & Ling, Y.-C. (2004) *Rapid Commun. Mass Spectrom.* **18**, 1167–1171. doi:10.1002/rcm.1460
- Natale, M., Bisson, C., Monti, G., Peltran, A., Garoffo, L.P., Valentini, S., Fabris, C., Bertino, E., Coscia, A., & Conti, A. (2004) *Mol. Nutr. Food Res.* **48**, 363–369. doi:10.1002/mnfr.200400011
- Akashi, S., Suzuki, K., Arai, A., Yamada, N., Suzuki, E.-I., Hirayama, K., Nakamura, S., & Nishimura, Y. (2006) *Rapid Commun. Mass Spectrom.* **20**, 1932–1938. doi:10.1002/rcm.2541
- Muller, L., Barták, P., Bednár, P., Frysová, I., Sevcik, J., & Lemr, K. (2008) *Electrophoresis* **29**, 2088–2093. doi:10.1002/elps.200700794
- Cereda, A., Kravchuk, A.V., D'Amato, A., Bachi, A., & Righetti, P.G. (2010) *J. Proteomics* **73**, 1732–1739. doi:10.1016/j.jprot.2010.05.010
- D'Amato, A., Bachi, A., Fasoli, E., Boschetti, E., Peltre, G., Sénéchal, H., & Righetti, P.G. (2009) *J. Proteome Res.* **8**, 3925–3936. doi:10.1021/pr900221x
- Cunsolo, V., Muccilli, V., Saletti, R., & Foti, S. (2011) *Eur. J. Mass Spectrom. (Chichester, Eng.)* **17**, 305–320. doi:10.1255/ejms.1147
- Fæste, C.K., Rønning, H.T., Christians, U., & Granum, P.E. (2011) *J. Food Prot.* **74**, 316–345. doi:10.4315/0362-028X.JFP-10-336

- (32) Guarino, C., Fuselli, F., La Mantia, A., Longo, L., Faberi, A., & Marianella, R.M. (2010) *Rapid Commun. Mass Spectrom.* **24**, 705–713. doi:10.1002/rcm.4426
- (33) Picariello, G., Mamone, G., Addeo, F., & Ferranti, P. (2011) *J. Chromatogr. A* **1218**, 7386–7398. doi:10.1016/j.chroma.2011.06.033
- (34) Monaci, L., Losito, I., Palmisano, F., & Visconti, A. (2010) *J. Chromatogr. A* **1217**, 4300–4305. doi:10.1016/j.chroma.2010.04.035
- (35) Monaci, L., & van Hengel, A.J. (2008) *J. Chromatogr. A* **1192**, 113–120. doi:10.1016/j.chroma.2008.03.041
- (36) Ansari, P., Stoppacher, N., Rudolf, J., Schuhmacher, R., & Baumgartner, S. (2011) *Anal. Bioanal. Chem.* **399**, 1105–1115. doi:10.1007/s00216-010-4422-0
- (37) Lutter, P., Parisod, V., & Weymuth, H. (2011) *J. AOAC Int.* **94**, 1043–1059
- (38) Czerwenka, C., Maier, I., Pittner, F., & Lindner, W. (2006) *J. Agric. Food Chem.* **54**, 8874–8882. doi:10.1021/jf061646z
- (39) Monaci, L., Piolli, R., De Angelis, E., Godula, M., & Visconti, A. (2014) *J. Chromatogr. A* **1358**, 136–144. doi:10.1016/j.chroma.2014.06.092
- (40) Quaglia, M., Pritchard, C., Hall, Z., & O'Connor, G. (2008) *Anal. Biochem.* **379**, 164–169. doi:10.1016/j.ab.2008.05.005
- (41) Pritchard, C., Quaglia, M., Burkitt, W.I., Mussell, C., O'Connor, G., & Parkes, H. (2009) *Clin. Chem.* **55**, 1984–1990. doi:10.1373/clinchem.2009.124354
- (42) MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., & MacCoss, M.J. (2010) *Bioinformatics* **26**, 966–968. doi:10.1093/bioinformatics/btq054
- (43) Geromanos, S.J., Vissers, J.P.C., Silva, J.C., Dorschel, C.A., Li, G.-Z., Gorenstein, M.V., Bateman, R.H., & Langridge, J.I. (2009) *Proteomics* **9**, 1683–1695. doi:10.1002/pmic.200800562
- (44) Jeffrey, C.S., Gorenstein, M.V., Li, G.-Z., Vissers, J.P.C., & Geromanos, S.J. (2006) *Mol. Cell. Proteomics* **1**, 144–156
- (45) Wieser, M.E., & Coplen, T.B. (2011) *Pure Appl. Chem.* **83**, 359–396
- (46) Chen, Q., Zhang, J., Ke, X., Lai, S., Tao, B., Yang, J., Mo, W., & Ren, Y. (2015) *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess.* **32**, 25–34
- (47) Chen, E.I., Cociorva, D., Norris, J.L., & Yates, J.R. (2007) *J. Proteome Res.* **6**, 2529–2538. doi:10.1021/pr060682a
- (48) Wiśniewski, J.R., Zougman, A., Nagaraj, N., & Mann, M. (2009) *Nat. Methods* **6**, 359–362. doi:10.1038/nmeth.1322