Comparative study of concatemer efficiency as an isotope-labelled internal standard for allergen quantification

Gavage, Maxime; Van Vlierberghe, Kaatje; Van Poucke, C.; De Loose, M.; Gevaert, K.; Dieu, Marc; Renard, Patricia; Arnould, Thierry; Filee, Patrice; Gillard, Nathalie

Published in:
Food Chemistry

DOI:
10.1016/j.foodchem.2020.127413

Publication date:
2020

Document Version
Peer reviewed version

Comparative study of concatemer efficiency as an isotope-labelled internal standard for allergen quantification

Maxime GAVAGE[a,b], Kaatje VAN VLIERBERGHE[c], Christof VAN POUCKE[c], Marc DE LOOSE[c], Kris GEVAERT[d,e], Marc DIEU[b,f], Patsy RENARD[b, f], Thierry ARNOULD[b], Patrice FILEE[a], Nathalie GILLARD[a,*]

a. CER Groupe, Rue du Point du Jour 8, 6900 Marloie, Belgium
b. Laboratory of Biochemistry and Cell Biology (URBC)-Namur Research Institute for Life Sciences (NARILIS), University of Namur, 61, Rue de Bruxelles, 5000 Namur, Belgium
c. Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Brusselsesteenweg 370, 9090 Melle, Belgium
d. VIB Center for Medical Biotechnology, Albert Baertsoenkaai 3, B9000 Ghent, Belgium
e. Department of Biomolecular Medicine, Ghent University, Albert Baertsoenkaai 3, B9000 Ghent, Belgium
f. MaSUN, Mass Spectrometry Facility, University of Namur, 61, Rue de Bruxelles, 5000 Namur, Belgium

* Corresponding author; telephone: +32 (0) 84 31 00 90; e-mail: n.gillard@cergroupe.be
CER Groupe, rue du Point du Jour, 8, 6900 Marloie, Belgium

Abstract

Mass spectrometry-based methods coupled with stable isotope dilution have become effective and widely used methods for the detection and quantification of food allergens. Current methods target signature peptides resulting from proteolytic digestion of proteins of the allergenic ingredient. The choice of appropriate stable isotope-labelled internal standard is crucial, given the diversity of encountered food matrices which can affect sample preparation and analysis. We propose the use of concatemer, an artificial and stable isotope-labelled protein composed of several concatenated signature peptides as internal standard. With a comparative analysis of three matrices contaminated with four allergens (egg, milk, peanut, and hazelnut), the concatemer approach was found to offer advantages associated with the use of labelled proteins, ideal but unaffordable, and circumvent certain limitations of traditionally used synthetic peptides as internal standards. Although used in the proteomic field for more than a decade, concatemer strategy has not yet been applied for food analysis.

Keywords

Food allergen analysis, mass spectrometry, isotope dilution, isotope-labelled internal standard, isotope-labelled concatemer

1. Introduction

Food allergy is defined as an adverse health effect arising from a specific reproducible immune response that occurs on exposure to a given food (Boyce et al., 2011). Several studies indicate an increase in the prevalence of food allergy with nearly 5% of adults and 8% of children being affected (Sicherer & Sampson, 2014). Given the absence of accepted treatment, the current solution for allergic patients relies on allergen avoidance to circumvent allergic reactions. However, this essentially requires correct food labelling and efficient risk management from food business operators to reduce the risk of contamination by allergens to acceptable levels. European legislation (Regulation [EU] No 1169/2011) requires the labelling of 14 allergenic ingredients when they are part of a foodstuff recipe. However, this legislation does not cover the presence of hidden allergens that are due to cross-contamination during food processing. Even if strongly requested by food producers and control laboratories, no harmonized regulatory framework for managing hidden allergens or action thresholds have been enacted in Europe. Some countries have set legal thresholds but with a high disparity among allergens and among countries (Planque et al., 2019). A quantitative risk assessment was also developed by VITAL® (Voluntary Incidental Trace Allergen Labelling) combining reference doses and exposure (Allen et al., 2014). The thresholds for allergenic proteins in food are based on clinical data and are indicators of the action levels, expressed as the total protein amount of the allergenic food (mg), below which only the most sensitive allergic subjects might react (1% of allergic patients or 5% of them for the less common foods). These values are often used by laboratories as a targeted limit of quantification (LOQ) in the absence of legal thresholds.

The development of a quantitative allergen risk assessment requires quantitative allergen analysis. During the last decade, mass spectrometry became the method of choice for allergen analysis (Ahsan, Rao, Gruppuso, Ramratnam, & Salomon, 2016). Allergen analysis by mass spectrometry is predominantly performed by specific analysis of peptides obtained by an enzymatic digestion of the proteins of the sample, including the proteins of the allergenic ingredients. One of the advantages of mass spectrometry-based methods is the possibility to simultaneously detect multiple peptides from multiple allergens, thus enabling time- and money-saving multiplexed analysis. Such a targeted approach, named multiple reaction monitoring (MRM), offers high sensitivity and specificity. Targeted proteomics is often used for absolute peptide quantification in combination with isotope dilution, a technique based on the use of an internal standard corresponding to the stable isotope-labelled version of the analyte (Monaci, Losito, De Angelis, Pilolli, & Visconti, 2013; Nitride et al., 2019; Planque et al.,

2019). The introduction of this isotope-labelled internal standard corrects for variability and various matrix effects during the actual analysis. Notably, ion suppression effects and, depending on the type of internal standard, matrix effects and analyte loss during sample preparation may be corrected by the use of isotope-labelled internal standards.

Peptides specific for allergen proteins are the analytes in mass spectrometry analysis of food allergens; however, the initial analytes are proteins. Stable isotope-labelled internal standards can therefore adopt different forms. In theory, a stable isotope-labelled protein is the ideal internal standard as, when added to the food that needs to be analyzed, it can correct for sample losses during all the steps of the sample preparation procedure (including protein extraction and digestion), as well as for matrix effects during mass spectrometry analysis. Such an approach was proposed by Newsome and Scholl (Newsome & Scholl, 2013) for the quantification of bovine milk α_{S1}-casein in baked goods. The main limitation of this approach, besides technical issues for protein production, is its cost. When one aims at multiplexed analysis, this necessitates the use of multiple isotope-labelled proteins, which is unrealistic for laboratories performing routine analyses (Planque, Arnould, & Gillard, 2017). Therefore, most laboratories rely on stable isotope-labelled synthetic peptides (Boo, Parker, & Jackson, 2018; Henrottin et al., 2019; Planque et al., 2019). However, in food allergen analysis, the initial analytes are proteins. Peptide internal standard and protein analytes can exhibit different behaviors during the extraction, leading to different extraction yield. Moreover, the peptides do not undergo the enzymatic digestion step which is known to be highly affected by the matrix effects (Korte, Oberleitner, & Brockmeyer, 2019).

Here, we implemented an alternative method based on the synthesis of a concatemer used as a stable isotope-labelled internal standard for allergen quantification. This strategy has been well adopted by proteomics researchers, and the concatemers are known as QconCAT (Pratt et al., 2006), but, as far as we know, these molecules have not yet been explored for food analysis. Concatemers are artificial proteins composed of concatenated, proteotypic peptides originating from different proteins of interest. The peptides themselves are typically first identified following mass spectrometry or are predicted from theoretical peptide sequences. Concatemers are typically recombinantly produced in an environment that allows labelling with stable isotopes (e.g.,^{13}C or ^{15}N). In contrast to synthetic peptides, concatemers need to be proteolytically digested to release their peptides, and thus, this peptide release is also affected by the interference caused by the matrix during the digestion step, in a manner similar to the analyte of interest. Another advantage of concatemers is their potential for multiplexing. A single

concatemer can be composed of numerous proteotypic peptides and can therefore be used for multiplexed allergen analysis. The limitation of this approach is fixed by the protein size reachable with recombinant protein expression, which is more than 100 kDa (Chambers, Austen, Fulghum, & Kim, 2004). This approach can be cost-effective when compared with using synthetic peptides for multiplexed analysis. For our study, we developed, produced, and purified a ^{15}N isotopically labelled concatemer composed of 19 proteotypic peptides, allowing for the analysis of 4 allergenic ingredients (egg, milk, peanut, and hazelnut). We evaluated the performance of this concatemer by the analysis of three uncontaminated food matrices spiked with increasing and defined concentrations (2.5 ppm to 50 ppm, where ppm corresponded to mg of total allergen protein per kg of matrix) of the selected allergen extracts. In addition, we compared the use of the concatemer with that of five synthetic peptides corresponding to tryptic peptides from the four considered allergens and with β-lactoglobulin, a bovine milk protein that was ^{15}N isotopically labelled.

2. Material and methods

2.1. Reagent and materials

Gene synthesis and cloning were ordered from GeneCust (Boynes, France). Acetic acid, ammonium bicarbonate, ampicillin sodium salt, chloramphenicol, dimethyl sulfoxide (DMSO), DL-dithiothreitol (DTT), expression plasmid pET17b(+) Novagen, HiLoad® 26/600 Superdex® 200 pg, imidazole hydrochloride, iodoacetamide (IAA), kanamycin monosulfate, Lennox broth (LB), Ni Sepharose® 6 fast flow GE Healthcare, Origami™ B(DE3) pLysS competent cells Novagen, phenylmethanesulfonyl fluoride (PMSF), Q Sepharose® Fast Flow, select agar, sodium chloride, sodium phosphate dibasic, sodium phosphate monobasic, tetracycline hydrochloride, tetraethylammonium bicarbonate (TEAB), trypsin from bovine pancreas, tris(hydroxymethyl)aminomethane (Tris) and urea were obtained from Sigma-Aldrich (Bornem, Belgium). One Shot™ BL21(DE3) chemically competent *Escherichia coli*, isopropyl β-D-thiogalactopyranoside (IPTG), SnakeSkin™ dialysis tubing, 3.5K MWCO, 22 mm were purchased from Thermo Fisher Scientific (Waltham, MA, USA). Bioexpress cell growth media (U-15N, 98%) (10x concentrate) was obtained from Buchem B.V. (Apeldoorn, the Netherlands), Trypsin Gold, Mass Spectrometry Grade from Promega (Madison, WI, USA), 4–20 Mini-PROTEAN® TGX™ precast protein gels from Bio-Rad (Hercules, CA, USA), Sep-Pak C18 6 cc Vac solid-phase extraction (SPE) cartridges from Waters (Milford, MA, USA), and 0.2 µm acrodisc syringe filters with supor membrane from Pall Corporation (Port Washington, NY, USA). Water, acetonitrile (ACN), and formic acid (FA) were obtained from Biosolve (Valkenswaard, the Netherlands). Labelled synthetic peptides ADIYTEQ**V**[$^{13}C_5{}^{15}N$]GR, FFVAPFPEVFG**K**[$^{13}C_6{}^{15}N_2$], GGLEPIN**F**[Ring-D_5]QTAADQAR, LS**F**[Ring-D_5]NPTQLEEQCHI, TANELNLLI**L**[$^{13}C_6{}^{15}N$]R were ordered from Eurogentec (Seraing, Belgium).

Food samples were analyzed by ultra-high performance liquid chromatography–tandem mass-spectrometry (UHPLC MS/MS) using an Acquity liquid chromatograph equipped with a C18 Acquity BEH130 column (2.1 x 150 mm; 1.7 µm) and coupled with a Xevo TQ-S micro triple quadrupole system (Waters, Milford, MA, USA). Characterization of ^{15}N isotopically labelled concatemer and β-lactoglobulin was performed by ultra-high performance liquid chromatography–high resolution mass spectrometry (UHPLC-HRMS) using an Acquity liquid chromatograph equipped with a C18 Acquity BEH130 column (2.1 x 150 mm; 1.7 µm) and coupled to a Xevo G2-XS QTof quadrupole time-of-flight system (Waters, Milford, MA, USA).

2.2. ^{15}N isotopically labelled concatemer production and purification

Design and production of the concatemer were adapted from the method of Pratt (Pratt et al., 2006).The first step focused on concatemer design and the selection of the concatenated peptides. Here, we considered 19 peptides (Table 1) originating from seven proteins of four allergenic ingredients (α_{S1}-casein and β-lactoglobulin from cow milk; ovalbumin, ovotransferrin, and vitellogenin-1 from hen's egg; Cor a 9 allergen from hazelnut; and Ara h 1 allergen from peanut). These 19 peptides were selected from a set of relevant peptide biomarkers identified by an empirical approach based on UHPLC-HRMS analysis of incurred and processed samples. The applied food processing steps, sample preparation, and selection criteria have already been detailed in our previous studies (Gavage et al., 2019, 2020; Van Vlierberghe et al., 2020). The peptides were then *in silico* concatenated, and the resulting polypeptide was flanked with an N-terminus initiator sequence including a methionine start and a C-terminus hexahistidine purification tag (His-tag). Hydrophobicity of each of the 19 peptides was evaluated based on their grand average of hydropathy (GRAVY) parameter. Hydrophobic and hydrophilic peptides were alternated in the concatemer sequence to avoid the formation of high hydrophobic clusters that can interfere with solvent accessibility of concatenated peptides and thus with their subsequent proteolysis during the sample preparation. Translation-associated aspects such as tRNA-mediated codon usage bias and mRNA secondary structure, known to impact the translation process (Gorochowski, Ignatova, Bovenberg, & Roubos, 2015), were also considered. Visual Gene Developer (University of California-Davis, Davis, CA, USA) was used to predict and optimize the mRNA secondary structure. The *in silico* designed DNA construct was finally chemically synthesized and cloned into the pET17b(+) expression vector using *NdeI* and *XhoI* restriction sites to give the pET17b(+)-concat1.

The *E. coli* BL21(DE3)/pET17b(+)-concat1 strain was inoculated in a 30 ml starter culture of ^{15}N labelled media (Bioexpress cell growth media [U-15N, 98%] with 100 µg/mL ampicillin) and grown overnight at 37 °C under 300 rpm orbital shaking. Cells were harvested by centrifugation (4000 x g, 5 min) and the pellet was resuspended in 1 mL of ^{15}N labelled media. Next, a volume of 660 µl of this bacterial suspension was used to inoculate a 1L ^{15}N labelled main culture. This culture was grown at 37 °C under 300 rpm orbital shaking until the optical density at 600 nm reached 0.6–0.8. Concatemer expression was next induced with 1 mM IPTG and cells were cultured overnight at 25 °C under 300 rpm orbital shaking. Cells were harvested by centrifugation (5000 x g, 15 min) and stored at –80 °C until concatemer purification.

The cell pellet of the 1 L culture was resuspended in 40 mL of lysis buffer (50 mM Tris - 10 mM imidazole - pH 8) with 1 mM PMSF. Cells were disrupted using a Vibra-Cell™ (Sonics, Newtown, CN, USA) ultrasonic probe. The cell lysate was centrifuged twice (40000 x g, 20 min) and filtered through 0.2 µm syringe filters before to be submitted to metal affinity chromatography purification. The protein solution was loaded on a 8 ml Ni Sepharose 6 Fast Flow column equilibrated with lysis buffer. An intermediate washing step was performed in the presence of 20 mM imidazole and the His-tag labelled concatemer was finally eluted by using a linear imidazole gradient from 20 mM to 250 mM. The elution fractions were analyzed on SDS-PAGE (Supplementary data 1). The positive fractions were pooled and dialyzed against the storage buffer (50 mM Tris - pH 8) to eliminate imidazole.

Total protein concentration was measured by absorbance at 280nm. A SDS-PAGE/densitometry method based on ImageJ software was used to estimate concatemer purity. A total of 84.5 mg of ^{15}N isotopically labelled concatemer were produced and purified with an estimated purity higher than 90%. Protein sequences, concentration calculations, and purity estimation are detailed in Supplementary data 3.

2.3. ^{15}N isotopically labelled β-lactoglobulin production and purification

The production of β-lactoglobulin, a cow milk protein, was adapted from the work of Loch and collaborators (Loch et al., 2016) who implemented a method leading to the cytoplasmic accumulation of correctly folded disulfide bond-dependent proteins. Briefly, two mutations (L2A/I3S) were introduced in the β-lactoglobulin to facilitate *in vivo* cleavage of the N-terminal methionine allowing for correct protein folding.) Further, the *E. coli* Origami B (DE3) pLysS strain, a glutathione reductase (gor) and thioredoxin reductase (trxB) mutated strain, was used for conducting the cytoplasmic co-expression of the protein of interest with DsbC, an *E.coli* cytoplasmic disulfide bond isomerase. The co-expression was achieved with the same expression vector (pET17b(+)-DsbC-BLg) in which the two genes were transcribed from individual T7 IPTG-inducible promoters.

To achieve the production of ^{15}N labelled β-lactoglobulin, expression (starter culture, main culture, and IPTG induction) conditions were similar as used for concatemer production. The antibiotics that were used were tailored to 200 µg/mL ampicillin, 34 µg/mL chloramphenicol, 15 µg/mL kanamycin, and 12.5 µg/mL tetracycline, and the IPTG concentration was 0.5 mM. Harvested cells were resuspended in 50 mM phosphate buffer, pH 6.5, with 1 mM PMSF and prepared for protein purification using the same procedure as for the concatemer. The purification of 15N labelled β-lactoglobulin was performed

according to the procedure described by Loch and collaborator (Loch et al., 2016). Briefly, this protocol combines anion-exchange chromatography (Q Sepharose® Fast Flow) with a NaCl linear elution gradient (up to 2 M) followed by size-exclusion chromatography (HiLoad® 26/600 Superdex® 200 pg) in initial conditions (50 mM phosphate buffer, pH 6.5). Eluates of these two purification steps were collected in 1 mL fractions and analyzed on SDS-PAGE (Supplementary data 2).

Total protein concentration was measured by absorbance at 280nm. A SDS-PAGE/densitometry method based on ImageJ software was used to estimate protein purity. Using this approach, a total of 2.4 mg of ^{15}N isotopically labelled β-lactoglobulin were produced and purified with an estimated purity higher than 70%. Protein sequences, concentration calculations, and purity estimation are detailed in Supplementary data 3.

2.4. Characterization of produced ^{15}N isotopically labelled proteins

Protein ^{15}N stable isotope enrichment was evaluated by UHPLC-HRMS analysis of its constitutive tryptic peptides. In separated containers, concatemer and β-lactoglobulin were diluted to 0.1 mg/mL with 50 mM TEAB, pH 9.2, to a final volume of 20 µl. Disulfide bridges of β-lactoglobulin were successively reduced and alkylated with DTT (10 mM final concentration, 45 min incubation at 37 °C under 300 rpm orbital agitation) and IAA (40 mM final concentration, 45 min incubation in the dark at 37 °C under 300 rpm orbital agitation). Concatemer and β-lactoglobulin were then proteolytically digested by adding 0.1 µg of trypsin gold (protein:trypsin ratio of 1:20). Digestion was conducted for 1 h at 37 °C under 300 rpm orbital agitation and stopped by the addition of 1% (final concentration) of FA followed by centrifugation (20000 x g, 5 min). Samples were ten-fold diluted with 5% ACN before UHPLC-HRMS analysis. Peptides (5 µl of sample was injected) were first separated by reverse-phase liquid chromatography using a 20 min water/ACN + 0.1% FA linear gradient from 5% to 40% of ACN. Data was acquired in MS^E mode with 0.3 s scan time within the 50 to 2000 m/z mass range. The data were processed using UNIFI software (Waters, Milford, MA, USA) and peptide mapping analysis type with traditional tryptic cleavage rules and setting cysteine carbamidomethylation and ^{15}N isotope labelling as a fixed modifications.

For each identified tryptic peptide, the most intense charge state was considered to define the ^{15}N stable isotopic enrichment. The isotopic enrichment or isotope incorporation rate was evaluated for each peptide by comparing the intensity (in counts) of the peak corresponding to the fully ^{15}N labelled (U-^{15}N) peptide with other peaks corresponding to partially ^{15}N labelled peptides. For practicality, we considered

231 a ^{13}C natural abundance of 1.1% and neglected hydrogen and oxygen isotopic distributions in our
232 calculations. Furthermore, only peaks corresponding to peptide with 1 ^{14}N isotope were considered in
233 our calculation. The proportion of U-^{15}N peptide was then obtained after comparing the intensity of the
234 peak corresponding to the (U-^{15}N & U-^{12}C) peptide with the peak corresponding to the [(U-1)-^{15}N & U-
235 ^{12}C] peptide. Protein isotopic enrichment was evaluated with the exponential trend given by the
236 proportion of the U-^{15}N version of each peptide considering its nitrogen content.

237 **2.5. Food matrices preparation**

238 Three blank food matrices – thus, not contaminated with the considered allergenic ingredients –
239 were prepared to assess the variability due to the food sample used in our study. These blank matrices
240 were baked cookies, chocolate, and freeze-dried cookie dough.

241 Cookie dough was produced in batches of 3 kg by mixing (Kenwood Major Titanium, Stainless Steel
242 Dough Hook, 15 min, max speed) the following ingredients purchased from a local supermarket in the
243 respective weight proportions as follows: wheat flour (Carrefour type 55)/water (Milli-Q)/olive oil (Bertoli
244 Classico)/salt (sodium chloride ACS, ≥ 99%, Thermo Scientific™)/baking powder (Dr. Oetker
245 Baking)/Sugar (Grand Pont Crystal Sugar): 57%/18%/10%/0.2%/0.8%/14%. The dough was
246 subsequently rolled out to a thickness of 0.5 mm, and cookies with a diameter of 8 cm were pressed out
247 of the dough (weight = 25 ± 2 g). Cookies were baked for 25 min with the following program: 1–10 min:
248 180 °C heat from above and 180 °C heat from below; 11–25 min: 180 °C heat from above and 160 °C
249 heat from below. This was done to ensure that the warming of the baking plate would not result in uneven
250 cookie baking. Cookies were left at ambient temperatures to cool down, and subsequently milled and
251 sieved (Retsch® ZM 200 ultra-centrifugal mill [Retsch GmbH, Haan, Germany] with a 0.75 mm pore
252 size sieve, 14000 rpm). Cookie powder was stored at 4 °C in the dark until further use.

253 Cookie dough was produced as described above, rolled out to a thickness of 1 cm, stored at –20
254 °C, and subsequently freeze-dried. Freeze-dried cookie dough was then milled and sieved (Retch® ZM
255 200 ultra-centrifugal mill with a 0.75 mm pore size sieve, 14000 rpm). The freeze-dried cookie dough
256 powder was stored at 4 °C in the dark until further use.

257 Chocolate was made by warming chocolate walsenpowder (90%; Callebaut, Belgium) and cacao
258 butter (10%, Callebaut, Belgium) in a water bath at 40 °C (maximum temperature). The mixture was
259 stirred for 15 min, after which 2% ammonium phosphatide (kindly provided by Palsgaard, Julesminde,
260 Denmark) was added. This mixture was again stirred for 15 min and subsequently poured into chocolate

molds, resulting in chocolate chips of around 5 g each. The chocolate was left to cool down and solidify at 4 °C for 2 h, and the chocolate chips were packed under vacuum and stored at 4 °C in the dark until further use.

2.6. Sample preparation for UHPLC-MS/MS analysis

Two series of samples were prepared and analyzed to be able to cover the three internal standards. Concatemer and β-lactoglobulin were isotopically labelled with the same strategy (^{15}N uniform labelling) and share common tryptic peptides, which cannot be distinguished after enzymatic digestion. Two series of samples were prepared. Labelled peptides and β-lactoglobulin were spiked in the first series (only one shared peptide LSFNPTQLEEQCHI) and labelled concatemer in the second one. For each series, the three blank matrices (baked cookie, chocolate, and lyophilized unbaked cookie dough) were spiked, before extraction, with the appropriate internal standard and with increasing amounts of a standard extract of the four allergens (milk, egg, peanut, and hazelnut). These allergen amounts corresponded to 0, 2.5, 5, 10, 25 and 50 ppm level points expressed in total allergen protein per matrix kg. For each series, each blank matrix and each point of the allergen curve, three biological sample replicates were prepared and analyzed. Stock solutions containing the four allergen standards at 20 mg/mL were prepared using a similar extraction protocol as that used for the samples (extraction, sonication, and centrifugation; see below). These stock solutions were then combined and diluted in appropriate ratios to spike samples at different contamination levels with a 100 µl volume. Combination and dilution were calculated based on theoretical protein content of standards assuming 100% extraction yield. Each internal standard was spiked at the similar molar level (0.25 nmol) with a 100 µl volume. Then,1 mg/mL stock solutions of the five considered labelled peptides (ADIYTEQ**V**[$^{13}C_5{}^{15}N$]GR, FFVAPFPEVFG**K**[$^{13}C_6{}^{15}N_2$], GGLEPIN**F**[Ring-D_5]QTAADQAR, LS**F**[Ring-D_5]NPTQLEEQCHI, and TANELNLLI**L**[$^{13}C_6{}^{15}N$]R) were combined and diluted at the appropriate concentration with 0.1% FA. Concatemer and β-lactoglobulin solutions were also diluted to be spiked at 0.25 nmol level with a 100 µl volume. This level, converted in equivalent allergen ppm, ranged from 10 ppm for abundant proteins, such as α_{S1}-casein, to more than 300 ppm for less abundant proteins, such as vitellogenin-1. This estimate was based on the natural abundance of each considered protein in the corresponding allergenic ingredient. Allergen standards and internal standard were added to blank matrices before extraction.

Samples were prepared as previously described (Planque et al., 2016). Briefly, protein from 2 g samples was extracted in 50 mL conical tubes with 20 mL of 200 mM Tris, pH 9.2, 2 M urea by shaking

at 20 °C for 30 min (Agitelec, J. Toulemonde, Paris, France) prior to ultrasonic treatment at 4 °C for 15 min. After centrifugation (4660 x g, 10 min), 10 mL of supernatant were diluted in digestion buffer (200 mM ammonium bicarbonate, pH 8.2). Protein disulfide bridges were successively reduced and alkylated with 45 min incubation steps at room temperature with the addition of 1 mL of 200 mM DTT and 1 mL of 400 mM IAA (in the dark). Protein was then enzymatically digested with the addition of 1 mL of trypsin solution (trypsin from bovine pancreas, 1 mg/mL in 50 mM acetic acid, pH 2.8) and incubation for 1 h at 37 °C. The digestion reaction was stopped by adding 300 µl of 20% FA to the samples, which were then centrifuged (4660 x g, 5 min). Obtained peptides were then purified and concentrated using C18 SPE cartridges, which were first conditioned with 18 mL of ACN followed by 18 mL of 0.1% FA before loading of 20 mL of the centrifuged sample. The cartridges were washed with 18 mL of 0.1% FA and eluted in 15 mL conical tubes with 6 mL of 80% ACN and 0.1% FA. A volume of 30 µl of DMSO was added to the sample before evaporation (40 °C under nitrogen flow) to avoid dryness. The pellet was finally dissolved in 600 µl of 5% ACN with 0.1% FA and centrifuged twice (4660 x g, 5 min in conical tube and 20 000 x g, 5 min in 1.5 mL microtube, keeping the supernatant) before UHPLC-MS/MS analysis.

2.7. UHPLC-MS/MS analysis and data analysis

The peptides were separated by reverse-phase chromatography on-line connected to a triple quadrupole mass spectrometer. The following 26 min solvent gradient (solvent A, 0.1% FA and solvent B, ACN and 0.1% FA) was applied to the 20 µl injected sample volume: 0–3 min: 92% solvent A; 3–18 min: linear gradient from 92% to 58% solvent A; 18–22.5 min: 15% solvent A; and 22.5–26 min: 92% solvent A, always at constant 0.2 mL/min flow rate. Eluted peptides were ionized using the positive electrospray source and analyzed in MRM mode. The source gas flow was set at 50 L/h and the source voltage at 2.5 kV for the capillary and 30 V for the cone. The source temperature was set at 150 °C and the desolvation temperature at 400 °C with a gas flow at 1200 L/h. Targeted transitions are summarized in Table 1. For each peptide, three transitions were analyzed, as well as the corresponding transitions for the related isotopically labelled internal standard(s) (peptides, concatemer, and β-lactoglobulin). The transitions were selected beforehand using criteria that included the MS signal intensity and the absence of interference for the three considered matrices. The MS/MS acquisition method was generated using the open source Skyline software (MacLean et al., 2010). The most intense transition was used for internal standard comparison calculation and the two others as confirmatory transitions. Internal standards were compared using the peak area ratio (for the most intense transition) between the peptide

from the allergenic ingredient and its corresponding isotopically labelled version from the internal standard.

3. Results and discussion

3.1. Choice of the isotope labelling strategy

Stable isotope internal standard labelling and associated isotopic enrichment are key elements in the design of quantitative mass spectrometry-based methods. The isotopic enrichment and mass shift combination has to be sufficient to avoid any potential risk of false positive introduction. The resolution of quadrupole analyzers is typically around 1 atomic mass unit (Georgiou & Danezis, 2015). Taking into account that most of the peptide ion precursors carry multiple charges and that peptides contain tens of carbons, which lead to widespread isotopic distribution (see Fig. 1), the mass shift introduced by the stable isotopes has to be sufficient to be able to totally distinguish the natural analyte from its internal standard. Considering these aspects, a mass shift of $m/z \geq 3$ is necessary. Furthermore, attention has to be paid to the actual isotope enrichment. Depending on the labelling strategy, an insufficient isotope enrichment may lead to the introduction of the unlabeled form of the internal standard, thus corresponding to the natural analyte itself and contaminating the quantitative analysis.

Several strategies have been developed to produce isotopically labelled proteins, including selective labelling using auxotrophic *E. coli* strains and growth medium supplemented with isotopically labelled amino acids (Mondal, Shet, Prasanna, & Atreya, 2013) or post-translational protein deuteration (Galan et al., 2018). In this study, we decided to use a rich bacterial cell growth medium specifically designed for ^{15}N labeling protein using *E. coli* as a host cell for recombinant protein expression. This original medium is an algal hydrolysate that contains the same level of amino acids as LB medium. This strategy allowed for stable and protein sequence independent labelling (as each amino acid contains at least one nitrogen) with a high isotopic enrichment. As one of the peptide biomarkers selection criteria concerned the actual peptide length (peptides should have at least 8 amino acids), $m/z \geq 3$ mass shift precaution is respected for triply charged precursor. Indeed, selected peptide biomarkers are tryptic peptides, with a lysine or an arginine in C-terminal position, holding two and four nitrogen atoms, respectively.

3.2. Characterization of ^{15}N isotopically labelled proteins

The isotopic enrichment in the concatemer and β-lactoglobulin was evaluated following analysis of their constitutive tryptic peptides by UHPLC-HRMS. The proportion of the fully ^{15}N labelled version of each tryptic peptide was estimated by comparing the intensities of the monoisotopic peak (U-^{15}N & U-^{12}C) and those of its isotope containing one ^{14}N isotope ([U-1]-^{15}N & U-^{12}C). As shown in Fig. 1, the

intensities of the peaks from peptides with more than one ^{14}N isotope were found to be negligible (relative peak intensity <1% compared to the [U-^{15}N & U-^{12}C] peak). Given the resolution of the MS system (40000), carbon and nitrogen isotopes could not be distinguished. As a result, the monoisotopic peak (U-^{15}N & U-^{12}C) was combined with the peak corresponding to the peptide with one ^{14}N and one ^{13}C isotope ([U-1]-^{15}N & $^{13}C_1$). The proportion of fully ^{15}N labelled peptide was evaluated by comparing (U-^{15}N & U-^{12}C) and ([U-1]-^{15}N & U-^{12}C) peak intensities. The part of the peak intensity corresponding to the (U-^{15}N & U-^{12}C) isotope therefore had to first be discriminated from the combined (U-^{15}N & U-^{12}C) and ([U-1]-^{15}N & $^{13}C_1$) isotopes' peak intensity. Since isotopes with more than one ^{14}N were found to be negligible, we assumed that the ([U-1]-^{15}N & U-^{12}C) isotope peak would only correspond to this combination of isotopes. The peak intensity of the ([U-1]-^{15}N & $^{13}C_1$ isotope could therefore be predicted from the ([U-1]-^{15}N & U-^{12}C) isotope peak intensity assuming a 1.1% natural abundance of ^{13}C isotopes and knowing the number of carbon atoms in the peptide. With this prediction, the (U-^{15}N & U-^{12}C) isotope peak intensity could be deduced from the combined isotopes' peak intensity.

The proportion of fully ^{15}N labelled peptide was evaluated for all the 19 concatenated tryptic peptides of the concatemer and for all identified tryptic peptides from β-lactoglobulin. As shown in Fig. 2, the relation between the labelling proportion and the number of nitrogen atoms in the peptides follows an exponential decay. The associated exponential decay constant corresponds to the natural logarithm of the isotopic enrichment. Indeed, for a given isotopic enrichment (φ), the proportion of fully ^{15}N labelled peptide with n nitrogens is given by φ^n, which can be transformed into $e^{\ln(\varphi)*n}$. Isotopic enrichment is deduced from this mathematical transformation by equating $\ln(\varphi)$ to experimentally obtained exponential arguments (–0.00446 for the concatemer and –0.00411 for β-lactoglobulin). These results give an isotopic enrichment of 99.5% for the concatemer and 99.6% for β-lactoglobulin, and are in agreement with the >98% isotopic enrichment of the growth medium.

By using a method for efficient isotopic labelling of recombinant protein, we demonstrated that the purified ^{15}N isotopically labelled concatemer and β-lactoglobulin internal standards fulfilled the required criteria regarding isotopic enrichment and the introduced mass shift. With this ^{15}N uniform labelling strategy, the introduced mass shift was sufficient to distinguish the internal standard from the natural analyte using the quadrupole analyzer. The lowest mass shift corresponded to the double charged FYTVISSLK peptide (from egg white ovotransferrin), one of the 19 concatenated peptides, which contained 10 nitrogen atoms and an associated mass shift of a m/z of 5. Such a mass shift and obtained isotopic enrichment combination prevented the risk of false positive introduction.

3.3. Comparison of isotopically labelled internal standards

Performance of the three types of isotopically labelled internal standards (peptides, concatemer, and protein) were evaluated following analysis of three food matrices (baked cookie [cookie], chocolate, and lyophilized unbaked cookie dough [dough]). In theory, a perfect internal standard would have the same exact behavior as its corresponding analyte during sample preparation and analysis. Hence, any analyte loss or matrix effect (during sample preparation or UHPLC-MS/MS analysis) which affects the analyte should equally affect the internal standard. Consequently, for a given natural analyte concentration and internal standard spike level, the signal ratio between a natural analyte and the internal standard would remain constant, independent of analyte losses and matrix effects. The three internal standards considered in this study were compared based on this correlation.

Similar matrix-matched calibration curves were prepared for the three matrices. These curves included a blank and five allergen concentrations ranging from 2.5 to 50 ppm (expressed in mg total allergen protein per kg of matrix), with each sample prepared in triplicate. For each combination of matrix and allergen contamination level, the appropriate internal standard(s) (isotopically labelled peptides and U-^{15}N β-lactoglobulin for the first sample series, and U-^{15}N concatemer for the second one) was spiked at the same concentration. Results are presented separately for each targeted peptide and its corresponding internal standard (five synthetic peptides, 19 allergenic tryptic peptides from U-^{15}N concatemer digestion, and four tryptic peptides from U-^{15}N β-lactoglobulin digestion). Representative peptides of each internal standard are shown in Fig. 3, and complete results are shown in Supplementary data 4. Performance of the different internal standards were evaluated by comparing the peak area ratio for the most intense transition (highlighted in Table 1) between the analyte and its corresponding internal standard for the three considered matrices. As shown, for a given analyte and internal standard concentration, the signal ratio remained constant when the internal standard was effective. The overlay of the generated linear regression lines was therefore used to evaluate internal standards performance. Overlapping regression lines indicated, for each allergen contamination level, a constant peak area ratio among matrices and thus, an effective internal standard, compensating for matrix effects. In addition to visual evaluation, overlapping regression lines were evaluated using the coefficient of variation (CV) between the slopes of the linear regression lines.

Overall, the best results were obtained for the isotopically labelled protein, U-^{15}N β-lactoglobulin. Assuming that recombinant protein folding was similar to the native protein and that the introduced N-terminal mutations had no significant impact, as previously demonstrated (Loch et al., 2016), this

approach seemed to be the one best suited one for quantifying allergen proteins. Aside from their mass (given the mass shift introduced by isotope labelling), both the analyte protein and the internal standard protein must have had the same properties. This was confirmed by the analysis of four constitutive tryptic peptides from β-lactoglobulin. Regression lines overlapped with all CV values below 15%. This confirmed that the internal standard had efficiently balanced matrix effects during sample preparation and UHPLC-MS/MS analysis, further supported by the fact that the analyte absolute peak area varied by a factor of up to 10 among the three considered matrices, depending on the peptide (data not shown) while the analyte/internal standard peak area ratio remained constant. However, the labelled protein was spiked into the different samples after food processing, which is known to impact peptide detectability and quantification (Korte et al., 2019; Parker et al., 2015). Peptide biomarker selection is therefore a crucial preliminary step in the development of a quantitative method, and selected peptides have to be robust to the food process.

The results obtained with isotopically labelled peptides and the concatemer were less straightforward to interpret. For some targeted peptides, such as LSFNPTQLEEQCHI with labelled peptides, or TNDNAQISPLAGR with the U-^{15}N concatemer, the internal standard efficiently compensated for matrix effects with observed CV values below 15%. However, for some other targeted peptides, such as GGLEPINFQTAADQAR with both U-^{15}N concatemer and labelled peptides, the analyte and internal standard signal ratio was highly matrix-dependent. In these cases, internal standards did not correctly balance matrix effects, potentially leading to biased allergen quantification. These results are consistent with those reported by Planque and co-workers (Planque et al., 2019). No significant difference was observed for the three peptides which were common to synthetic peptides and concatemer used as internal standards.

Isotopically labelled peptides are not subject to one of the crucial steps during sample preparation, this being the proteolytic digestion with trypsin. The composition of the food matrix directly impairs the efficiency of enzymatic digestion at least in two different ways. First, different matrices have different protein concentrations, directly affecting the protein/enzyme ratio. Labelled peptides do not balance for this aspect. Second, some other sample components, such as polyphenols and tannins, may also affect the efficiency of trypsin digestion (Gonçalves, Mateus, Pianet, Laguerre, & De Freitas, 2011), which might help to explain why the chocolate matrix gave lower signals for most of targeted peptides. Contrary to the labelled peptides, the U-^{15}N concatemer needed to be digested by trypsin to yield peptides that could be detected upon UHPLC-MS/MS analysis. Therefore, factors such as the sample protein content

or the presence of tannins should be balanced when using such an internal standard. However, our results indicated that the performance of the concatemer was peptide-dependent. For some peptides, such as TNDNAQISPLAGR from hazelnut Cor a 9 allergen and FFVAPFPEVFGK from milk α_{S1}-casein, matrix effects were efficiently balanced with linear regression lines CV below 15% between the matrices. However, for other peptides, such as NVNFDGEILK from egg vitellogenin-1 and TPEVDDEALEK from milk β-lactoglobulin, the associated CVs were much higher (>30%).

Matrix effects can also affect analytes by other means. Robustness to food processing was one of the criteria for peptide biomarkers selection (Gavage et al., 2019, 2020; Van Vlierberghe et al., 2020) and can therefore be excluded. Variation in protein extraction can also be excluded as, for all selected proteins, multiple peptides were included in the U-^{15}N concatemer, and no general trend of the matrix effect was observed for all the peptides of a given protein. Indeed, if protein extraction of the analyte and/or the internal standard was affected by the matrix, all peptides from a given protein should be equally impacted, which was not observed.

Proteolytic digestion of extracted proteins is a key step in sample preparation and could be a source of the observed variability. Even if the concatemer internal standard needs to be digested to release its constitutive peptides, multiple factors could influence the digestion kinetics. For instance, amino acids surrounding trypsin recognition sites are known to influence the efficiency of peptide bond hydrolysis (Siepen, Keevil, Knight, & Hubbard, 2007). Cleavage sites are described using the nomenclature formulated by Schechter and Berger (Schechter & Berger, 1967), as P4-P3-P2-P1-P1'-P2'-P3', in which cleavage of the peptide bond occurs between P1 and P1'. Arginine, lysine, and proline in position P1' have, for instance, a negative effect on the digestion efficiency. The acidic amino acids aspartate and glutamate also negatively influence digestion when they are present near the cleavage site. These aspects were taken into account during peptide biomarkers selection, and sequences known to negatively affect trypsin digestion were rejected. However, peptide biomarkers were synthetically stitched together in the concatemer. Considering a given peptide in the concatemer, its cleavage site is surrounded at the N-terminal side (P4 to P1) by amino acids from this peptide but also by amino acids from its neighboring peptide at the C-terminal side (P1' to P3'). Consequently, at a local scale, enzymatic digestion of the concatemer only partially reflects digestion of the natural proteins. This difference between natural analytes and concatemers might lead to differences in enzymatic digestion kinetics and could have been a source of the observed variations. A relatively simple solution to overcome this would be the introduction of amino acids between each targeted peptide of the concatemer. Such introduced

amino acids could be the flanking amino acids in the corresponding natural protein sequence, a solution known as a peptide-concatenated standard (PCS) (Kito, Ota, Fujita, & Ito, 2007). However, amino acids surrounding the cleavage site in the three-dimensional structure of the protein might also affect trypsin digestion. Hence, cleavage sites surrounded by acidic amino acids, characterized by a greater average exposed area, are more subject to missed-cleavages.

Besides flanking amino acids, structural parameters also interfere with enzymatic digestion of a protein. According to the work of Hamady and co-workers (Hamady, Cheung, Tufo, & Knight, 2005), secondary protein structures affect trypsin digestion efficiency. Cleavage sites within unstructured domains are more prone to be cleaved incorrectly, whereas cleavage sites in alpha-helices are more favorable. The structures of proteins targeted by the UHPLC-MS/MS method, when available, were analyzed to define whether observed variability among peptides could be linked to findings of Hamady and co-workers or not (Hamady et al., 2005). No general trend emerged from our data, limited to the 19 concatenated peptides. However, three-dimensional and structural aspects could be included in a future peptide biomarker selection, in addition to all other criteria already considered in this study.

4. Conclusions

Mass spectrometry-based detection and quantification of food allergens in processed food products remains challenging. Currently, no threshold values for allergen trace-level contamination have been established in European legislation, but these are highly expected by all stakeholders involved in the food chain, from producers to control laboratories, and will require quantitative analysis methods. Quantitative methods based on stable dilution techniques need isotopically labelled internal standards.

Here, we presented and compared the performances of three different types of isotopically labelled internal standards for allergen analysis in processed food products: synthetic peptides, concatemer, and protein. These internal standards were compared through the analysis of three matrix-matched calibration curves (cookie, chocolate, and unbaked lyophilized cookie dough) for four targeted allergens (egg, milk, peanut, and hazelnut). An effective internal standard needs to behave similar to the natural analyte and is therefore identically impacted by matrix effects during sample preparation and UHPLC-MS/MS analysis. As expected from a theoretical point of view, the isotopically labelled protein that was used as an internal standard gave the best results. A constant signal ratio between the analyte and the internal standard peak areas was observed in all matrices tested for the four tryptic peptides generated from the studied protein. However, we need to emphasize that these results only come from one investigated protein, β-lactoglobulin from milk.

Results from our studies using peptides and the concatemer were more equivocal and seemed to be peptide-dependent. For some synthetic peptides or some tryptic peptides from the concatemer, matrix effects during sample preparation and UHPLC-MS/MS analysis could be efficiently countered by the applied internal standards, whereas for other peptides, significant matrix effects were observed. However, the non-inferiority of the results obtained for the tryptic peptides from the concatemer was established, when compared to synthetic peptides. Moreover, the addition of any synthetic peptide in a method represent an additional cost, limiting therefore the number of targeted peptides for routine laboratories. The concatemer production costs are relatively independent of the number of concatenated tryptic peptides. From a rough estimate of ten peptides, the use of a concatemer as internal standard is financially advantageous and supersedes synthetic peptides.

Even though isotopically labelled synthetic peptides are currently the most commonly used internal standard for allergen analysis, they do not exactly reflect the natural situation as they do not need to be subjected to proteolytic digestion, while part of the variability observed in our study could have come from proteolytic digestion. Concatemers clearly need to be digested to release their constituting

peptides. However, our data seem to indicate that the digestion of the concatemer could be improved to more efficiently represent analyte protein digestion. In this respect, introducing flanking amino acids between each individual peptide (i.e. the PCS strategy) could be a future asset. Moreover, for our concatemer construct, peptide biomarker selection was mainly focused on robustness to food processing and local sequences, but additional criteria, such as protein structure and the local digestion site environment, could be included in the peptide selection process. Such possible future improvements strongly suggest that isotopically labelled concatemers could represent relevant internal standards, as they overcome limitations of the use of synthetic peptides, while combining advantages of the use of labelled proteins and, further, allowing for multiple allergen quantification by mass spectrometry.

Acknowledgments

The research that yielded these results was funded by the Belgian Federal Public Service of Health, Food Chain Safety and Environment through the contract RT 15/10 ALLERSENS. The authors also want to gratefully acknowledge the contributions of the University of Namur Mass Spectrometry platform (MaSUN).

5. Bibliography

Ahsan, N., Rao, R. S. P., Gruppuso, P. A., Ramratnam, B., & Salomon, A. R. (2016). Targeted proteomics: Current status and future perspectives for quantification of food allergens. *Journal of Proteomics*. https://doi.org/10.1016/j.jprot.2016.04.018

Allen, K. J., Remington, B. C., Baumert, J. L., Crevel, R. W. R., Houben, G. F., Brooke-Taylor, S., … Taylor, S. L. (2014). Allergen reference doses for precautionary labeling (VITAL 2.0): Clinical implications. *Journal of Allergy and Clinical Immunology*. https://doi.org/10.1016/j.jaci.2013.06.042

Boo, C. C., Parker, C. H., & Jackson, L. S. (2018). A targeted LC-MS/MS method for the simultaneous detection and quantitation of egg, milk, and peanut allergens in sugar cookies. *Journal of AOAC International*. https://doi.org/10.5740/jaoacint.17-0400

Boyce, J. A., Assa'ad, A., Burks, A. W., Jones, S. M., Sampson, H. A., Wood, R. A., … Schwaninger, J. M. (2011). Guidelines for the Diagnosis and Management of Food Allergy in the United States: Summary of the NIAID-Sponsored Expert Panel Report. *Journal of the American Academy of Dermatology*. https://doi.org/10.1016/j.jaad.2010.11.020

Chambers, S. P., Austen, D. A., Fulghum, J. R., & Kim, W. M. (2004). High-throughput screening for soluble recombinant expressed kinases in Escherichia coli and insect cells. *Protein Expression and Purification*. https://doi.org/10.1016/j.pep.2004.03.003

Galan, S. R. G., Wickens, J. R., Dadova, J., Ng, W. L., Zhang, X., Simion, R. A., … Davis, B. G. (2018). Post-translational site-selective protein backbone α-deuteration. *Nature Chemical Biology*. https://doi.org/10.1038/s41589-018-0128-y

Gavage, M., Van Vlierberghe, K., Van Poucke, C., De Loose, M., Gevaert, K., Dieu, M., … Gillard, N. (2019). Selection of egg peptide biomarkers in processed food products by high resolution mass spectrometry. *Journal of Chromatography A*, *1584*, 115–125. https://doi.org/10.1016/J.CHROMA.2018.11.036

Gavage, M., Van Vlierberghe, K., Van Poucke, C., De Loose, M., Gevaert, K., Dieu, M., … Gillard, N. (2020). High-resolution mass spectrometry-based selection of peanut peptide biomarkers considering food processing and market type variation. *Food Chemistry*. https://doi.org/10.1016/j.foodchem.2019.125428

Georgiou, C. A., & Danezis, G. P. (2015). Elemental and Isotopic Mass Spectrometry. *Comprehensive Analytical Chemistry*, *68*, 131–243. https://doi.org/10.1016/B978-0-444-63340-8.00003-0

569 Gonçalves, R., Mateus, N., Pianet, I., Laguerre, M., & De Freitas, V. (2011). Mechanisms of tannin-
570 induced trypsin inhibition: A molecular approach. *Langmuir*. https://doi.org/10.1021/la202280c
571 Gorochowski, T. E., Ignatova, Z., Bovenberg, R. A. L., & Roubos, J. A. (2015). Trade-offs between
572 tRNA abundance and mRNA secondary structure support smoothing of translation elongation
573 rate. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkv199
574 Hamady, M., Cheung, T. H. T., Tufo, H., & Knight, R. (2005). Does protein structure influence trypsin
575 miscleavage? *IEEE Engineering in Medicine and Biology Magazine*, *24*(3), 58–66.
576 https://doi.org/10.1109/MEMB.2005.1436461
577 Henrottin, J., Planque, M., Huet, A.-C., Marega, R., Lamote, A., & Gillard, N. (2019). Gluten analysis in
578 processed foodstuffs by a multi-allergens and grain-specific UHPLC-MS/MS method: One
579 method to detect them all. *Journal of AOAC International*. https://doi.org/10.5740/jaoacint.19-
580 0057
581 Kito, K., Ota, K., Fujita, T., & Ito, T. (2007). A synthetic protein approach toward accurate mass
582 spectrometric quantification of component stoichiometry of multiprotein complexes. *Journal of*
583 *Proteome Research*. https://doi.org/10.1021/pr060447s
584 Korte, R., Oberleitner, D., & Brockmeyer, J. (2019). Determination of food allergens by LC-MS:
585 Impacts of sample preparation, food matrix, and thermal processing on peptide detectability and
586 quantification. *Journal of Proteomics*, *196*, 131–140.
587 https://doi.org/10.1016/J.JPROT.2018.11.002
588 Loch, J. I., Bonarek, P., Tworzydło, M., Polit, A., Hawro, B., Łach, A., … Lewiński, K. (2016).
589 Engineered β-Lactoglobulin Produced in E. coli: Purification, Biophysical and Structural
590 Characterisation. *Molecular Biotechnology*. https://doi.org/10.1007/s12033-016-9960-z
591 MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., … MacCoss,
592 M. J. (2010). Skyline: An open source document editor for creating and analyzing targeted
593 proteomics experiments. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btq054
594 Monaci, L., Losito, I., De Angelis, E., Pilolli, R., & Visconti, A. (2013). Multi-allergen quantification of
595 fining-related egg and milk proteins in white wines by high-resolution mass spectrometry. *Rapid*
596 *Communications in Mass Spectrometry*, *27*(17), 2009–2018. https://doi.org/10.1002/rcm.6662
597 Mondal, S., Shet, D., Prasanna, C., & Atreya, H. S. (2013). High yield expression of proteins in E. coli
598 for NMR studies. *Advances in Bioscience and Biotechnology*.
599 https://doi.org/10.4236/abb.2013.46099

Newsome, G. A., & Scholl, P. F. (2013). Quantification of allergenic bovine milk αs1-casein in baked goods using an intact 15N-labeled protein internal standard. *Journal of Agricultural and Food Chemistry*. https://doi.org/10.1021/jf3015238

Nitride, C., Nørgaard, J., Omar, J., Emons, H., Esteso, M. J. M., & O'Connor, G. (2019). An assessment of the impact of extraction and digestion protocols on multiplexed targeted protein quantification by mass spectrometry for egg and milk allergens. *Analytical and Bioanalytical Chemistry*. https://doi.org/10.1007/s00216-019-01816-z

Parker, C. H., Khuda, S. E., Pereira, M., Ross, M. M., Fu, T. J., Fan, X., … Jackson, L. S. (2015). Multi-allergen Quantitation and the Impact of Thermal Treatment in Industry-Processed Baked Goods by ELISA and Liquid Chromatography-Tandem Mass Spectrometry. *Journal of Agricultural and Food Chemistry*. https://doi.org/10.1021/acs.jafc.5b04287

Planque, Arnould, Delahaut, Renard, Dieu, & Gillard. (2019). Development of a strategy for the quantification of food allergens in several food products by mass spectrometry in a routine laboratory. *Food Chemistry*. https://doi.org/10.1016/j.foodchem.2018.08.095

Planque, Arnould, & Gillard. (2017). Food Allergen Analysis: Detection, Quantification and Validation by Mass Spectrometry. In *Allergen*. https://doi.org/10.5772/intechopen.69361

Planque, M., Arnould, T., Dieu, M., Delahaut, P., Renard, P., & Gillard, N. (2016). Advances in ultra-high performance liquid chromatography coupled to tandem mass spectrometry for sensitive detection of several food allergens in complex and processed foodstuffs. *Journal of Chromatography A*, *1464*, 115–123. https://doi.org/http://doi.org/10.1016/j.chroma.2016.08.033

Pratt, J. M., Simpson, D. M., Doherty, M. K., Rivers, J., Gaskell, S. J., & Beynon, R. J. (2006). Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nature Protocols*. https://doi.org/10.1038/nprot.2006.129

Schechter, I., & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications*, *27*(2), 157–162. https://doi.org/10.1016/S0006-291X(67)80055-X

Sicherer, S. H., & Sampson, H. A. (2014). Food allergy: Epidemiology, pathogenesis, diagnosis, and treatment. *Journal of Allergy and Clinical Immunology*. https://doi.org/10.1016/j.jaci.2013.11.020

Siepen, J. A., Keevil, E. J., Knight, D., & Hubbard, S. J. (2007). Prediction of missed cleavage sites in tryptic peptides aids protein identification in proteomics. *Journal of Proteome Research*, *6*(1), 399–408. https://doi.org/10.1021/pr060507u

631 Van Vlierberghe, K., Gavage, M., Dieu, M., Renard, P., Arnould, T., Gillard, N., … Van Poucke, C.
632 (2020). Selection of universal peptide biomarkers for the detection of the allergen hazelnut in
633 food trough a comprehensive, high resolution mass spectrometric (HRMS) based approach.
634 *Food Chemistry*, 125679. https://doi.org/10.1016/J.FOODCHEM.2019.125679
635