Journal of Proteomics 147 (2016) 156–168

Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot

Defining the wheat gluten peptide fingerprint via a discovery and targeted proteomics approach

María José Martínez-Esteso, Jørgen Nørgaard, Marcel Brohée, Reka Haraszi [1], Alain Maquet, Gavin O'Connor *

European Commission, Joint Research Centre — Institute of Reference Materials and Measurements, Retieseweg 111, Geel 2440, Belgium

ARTICLE INFO

Article history:
Received 1 September 2015
Received in revised form 2 March 2016
Accepted 8 March 2016
Available online 16 March 2016

Keywords:
Peptide markers
Coeliac disease
Mass spectrometry
Proteomics
Wheat gluten fractionation
Food labelling legislation
Food allergen

ABSTRACT

Accurate, reliable and sensitive detection methods for gluten are required to support current EU regulations. The enforcement of legislative levels requires that measurement results are comparable over time and between methods. This is not a trivial task for gluten which comprises a large number of protein targets. This paper describes a strategy for defining a set of specific analytical targets for wheat gluten.

A comprehensive proteomic approach was applied by fractionating wheat gluten using RP-HPLC (reversed phase high performance liquid chromatography) followed by a multi-enzymatic digestion (LysC, trypsin and chymotrypsin) with subsequent mass spectrometric analysis. This approach identified 434 peptide sequences from gluten. Peptides were grouped based on two criteria: unique to a single gluten protein sequence; contained known immunogenic and toxic sequences in the context of coeliac disease.

An LC-MS/MS method based on selected reaction monitoring (SRM) was developed on a triple quadrupole mass spectrometer for the specific detection of the target peptides. The SRM based screening approach was applied to gluten containing cereals (wheat, rye, barley and oats) and non-gluten containing flours (corn, soy and rice).

A unique set of wheat gluten marker peptides were identified and are proposed as wheat specific markers.

Significance: The measurement of gluten in processed food products in support of regulatory limits is performed routinely. Mass spectrometry is emerging as a viable alternative to ELISA based methods. Here we outline a set of peptide markers that are representative of gluten and consider the end user's needs in protecting those with coeliac disease. The approach taken has been applied to wheat but can be easily extended to include other species potentially enabling the MS quantification of different gluten containing species from the identified markers.

1. Introduction

For regulatory purposes gluten is defined as 'the protein fraction from wheat, barley, rye, oats or their crossbred varieties and derivatives thereof, to which some persons are intolerant and that is insoluble in water and 0.5 M NaCl' [1]. However, there are a diverse range of definitions for gluten in the literature that have been constructed from different perspectives. Cereal grain proteins, including gluten, have been classically defined according to their solubility. The resulting categories, or Osborne fractions, in wheat are defined as: water soluble albumins; salt soluble globulins; alcohol soluble gliadin (prolamins) and insoluble glutenin (glutelins) [2]. Wheat gluten is a collection of seed storage proteins found in the wheat grain that can be divided into two main groups, gliadins and glutenins. Gliadins are subdivided into three main groups according to their electrophoretic mobilities. The first group contains α- and β-gliadins, the second contains γ-gliadins and the third consists of Ω-gliadins. Glutenins are divided into two groups according to their apparent molecular weight on a SDS-PAGE, high molecular weight glutenin subunits (HMW-GS) and low molecular weight glutenin subunits (LMW-GS). Glutenins and gliadins are widely studied due to their contribution to the quality of the end-product of bakery and pasta goods, including the rheological characteristics of dough made from wheat flour. Besides the importance of gluten proteins in food quality, they have a direct impact on the human health by triggering wheat related food disorders such as coeliac disease (CD) [3]. Gluten is not an individual protein but rather the generic name given to a complex protein fraction from a multigenic family of seed storage proteins. The homolog gluten genes are found in wheat, rye, barley, oats and their translated proteins can elicit CD by recognition of specific sequences in susceptible individuals. However, a better definition of gluten for the analytical community in terms of a specific set of analytes to be measured is still lacking [4].

Coeliac disease, the most common chronic inflammatory condition in developed countries, is thought to affect 1% of the population [5]. CD is an immune-mediated enteropathy caused by abnormal immune response to gluten in genetically susceptible people carrying the HLA-DQ2 or DQ8 haplotypes. In affected individuals, gluten intake

* Corresponding author.
E-mail address: gavin.o'connor@ec.europa.eu (G. O'Connor).
[1] Campden BRI, Station road, Chipping Campden, GL55 6LD, UK.

http://dx.doi.org/10.1016/j.jprot.2016.03.015

in the diet results in a harmful substance since the proteases in the human digestive system cannot completely digest the gluten proteins. These highly resistant peptides can flow through the gastrointestinal tract where they can initiate the inflammatory cascade. These gluten peptides are involved in the disease process differently, either being toxic or immunogenic [6]. Despite the recent progresses in understanding the mechanism of the disease, no cure currently exists for this condition and coeliac patients have to adhere to a strict life-long gluten-free diet.

Cereals represent an important source of nutrients for humans. In particular, wheat (*Triticum* sp.) is one of the most important cereal grain crops, which serves as a staple food for 30% of the human population. Wheat as a raw material for the production of starch is of enormous relevance to the food industry and is a cheap source of functional agents used in food formulations. These include baked products and many other products that would not normally be recognized as containing wheat, such as fruit juices, beverages, dairy and meat-derived food products [7]. This makes it very difficult for sensitive individuals to avoid gluten in their diet.

To protect affected consumers, worldwide regulations on gluten-free labelling have been implemented; most of these are based on the recommendations by Codex Stan 118–1979 [1]. Commission Regulation (EC) No 41/2009, and more recently Regulation (EU) No 828/2014, which will be enforced from July 2016, addresses the labelling of gluten-free and low-gluten food products in Europe. Products containing ingredients derived from wheat, rye, barley and oats, which have been processed to eliminate gluten or that can be potentially contaminated with these sources may be labelled with a specific claim as to their gluten content. These products can be labelled as 'very low-gluten', if the gluten content does not exceed of 100 mg/kg and as 'gluten-free' if the gluten content is equal to or lower than 20 mg/kg. The threshold established in this regulation is based on clinical and serological evidence that chronic exposure to traces of gluten are not harmful to sensitive individuals based on a normal dietary intake of cereal derived foods [8].

In support of current legislation and an increasing gluten-free market there is a critical need to develop analytical methods to quantify the gluten content in 'very low-gluten' and 'gluten-free' food products. The detection and quantification of gluten is important not only to detect gluten in relatively high amounts but also at trace levels, since long exposure to low levels of gluten has been reported as more detrimental than an acute intoxication [7].

Currently ELISA (enzyme linked immunosorbent assay) based on the R5 antibody, is the only method recommended by Codex for the quantification of gluten [1]. However, ELISA measurements based on commercially available kits have proven inconsistent and the measurement results have not been comparable. Therefore, there is a lack of standardization in gluten measurements. The main issues identified for this lack of comparability are: the lack of a certified reference material; cross-reactivity of antibodies; unique kit procedures (different buffers, concentration of antibodies, extraction protocols, etc.); and the measurement of different protein targets. Moreover, current ELISA methods do not differentiate the cereal source of the gluten (wheat, rye or barley). Many ELISA detection methods have focused on reporting the gluten content in reference to wheat, whereby assay components (calibrants and conversion factors etc.) and antibodies have been developed to recognize gliadins. The results are converted into an amount of gluten by applying a factor of 2. This assumes that the ratio of gliadin/glutenins remains stable regardless of different conditions (cultivars, extraction procedures, antibodies cross-reactivity, etc.) which introduce a significant degree of measurement uncertainty [9,4,10,11,12]. Therefore, species specific variations in the detected proteins are not taken into account when expressing the final gluten content of a sample. An agreement on the specific analyte(s)/target(s) is critical for the analytical community to improve the current situation and make gluten measurements comparable enabling the current legislation to be enforced.

New methodologies that can address these issues are demanded [11]. Mass spectrometry (MS), in particular, is a promising alternative to ELISA for the detection and quantification of proteins in contaminated food, as it can target multiple and very specific analytes with good sensitivity and accuracy. MS-based methods have recently been developed to quantify protein allergens like milk, egg, nuts [13] using selected reaction monitoring (SRM) as a targeted approach for specific protein derived peptides. SRM on triple quadrupole (QqQ) instruments allows the precise quantification of target proteins in complex samples with a broad dynamic range [14]. In a SRM targeted approach, peptide sequences that are uniquely specific to a particular protein, termed as signature peptides, are detected and quantified [15].

One solution to facilitate and improve the comparability of gluten measurements is the production of certified reference materials (CRMs). The first step in enabling the production of CRMs and the development of new confirmatory methods, is an agreement on a set of markers that can be used to infer the presence of gluten and that enable the quantity of gluten present to be determined. Currently these specific gluten targets have yet to be defined.

The strategy applied in our work identified a set of gluten markers from a single source, the main source of gluten (*Triticum sp.*), using both discovery and targeted based proteomic approaches. A combined strategy, based on the first dimension fractionation of GluVital™ wheat gluten using RP-HPLC followed by a multi-enzymatic based digestion of the resulting fractions and high resolution MS or MS/MS measurements was employed. The utility of these peptides, to act as wheat gluten markers, was investigated via a number of theoretical and practical approaches. A set of target peptides unique to *Triticum* sp. were selected. The basis for selection was either they were unique to a single wheat protein in the current protein databank (NCBInr) or were identified containing sequences known to cause intolerance in CD. The complete set of selected peptides belong to the protein fractions currently detected by commercially available ELISAs and were not detected in other gluten-containing cereals such as rye, barley or oats as assessed by an SRM screening approach. The detection of specific markers is the first step in enabling the quantification of gluten from specific sources providing both gluten content and identity of the cereal species in one method. The current EU legislation (Regulation (EC) No 41/2009 and Regulation (EU) No 828/2014) requires the quantification of total gluten from a mixture of cereals. Therefore the discovery of the proposed markers identified addresses the first step toward gluten quantification in food.

2. Materials and methods

2.1. Materials

HPLC solvents including ULC grade acetonitrile (ACN), ULC grade water and 99% formic acid (FA) were purchased from Biosolve (Valkenswaard, NL). Wheat gluten, GluVital™, was obtained from Cargill, Bergen op Zoom, NL. The following food grade flours were obtained from their respective producers: Wheat flour, *Triticum aestivum* (Albona, Mills United, Germany); rye flour, *Secale cereale* (Natudis B.V., Harderwijk, NL); barley flour, *Hordeum vulgare* ground from barley grain (Lima nv., Aalter, Belgium); oat flour, *Avena sativa* ground from oat grain (Lima nv., Maldegem, Belgium); soy flour, *Glycine max* (Biofresh, Gavere, Belgium); rice flour, *Oryza sativa* (Ekoland, Natudis B.V., Harderwijk, NL) which contained the warning "may contain gluten traces"; corn flour, *Zea mays* ground from polenta (Dutch Organic Int. Trade, Barneveld, NL).

Reagents used for sample extraction/preparation including dithiotreitol (DTT), iodoacetamide, ammonium bicarbonate (BCA), phosphate buffered saline (PBS), urea as well as the peptide standards

MRFA, bradykinin, angiotensin I, angiotensin II, glu-fibrinopeptide B and Insulin chain B were purchased from Sigma Aldrich (St. Luis, MO, USA). Trifluoroacetic acid and mercaptoethanol were purchased from Merck (Darmstadt, Germany). Trypsin, (Trypsin Gold, Mass spectrometry grade), rLys-C (Mass Spec grade) and chymotrypsin (sequencing grade) were purchased from Promega (Madison, USA). Nylon 0.45 μm centrifuged filters were obtained from VWR (Leuven, Belgium).

2.2. Protein extraction and fractionation

GluVital™ wheat gluten was resuspended in 1/10 (w/v) 6 M urea by placing the vial in a mixing rotator for 1 h at room temperature (rt) (21 ± 3 °C). The dissolved gluten proteins were reduced by adding 50 μl of 0.2 M DTT in 1 ml of extract and the solution was incubated at 37 °C for 1 h. The proteins were then alkylated by adding 200 μl of 0.2 M iodoacetamide per ml and the solution was kept at rt. for 1 h in the dark. 100 μl of the reduced and alkylated gluten proteins were diluted 5-fold in solvent A (0.1% FA in water) and were filtered through a 0.45 μm filter. The proteins in the filtrate were separated by HPLC (Proteome Lab PF 2D, Beckman Coulter) using a Zorbax 300 Å SB C18, 4.6 mm (ID) × 250 mm (5 μm) column (Agilent Technologies) with a solvent flow rate of 0.8 ml/min. The sample constituents were separated using a binary solvent consisting of solvent A, 0.1% FA in water, and solvent B, 0.1% FA in 95% ACN:5% water. Samples were loaded, using a 50 μl partial loop injection onto the column that had been equilibrated with the starting mobile phase consisting of 90:10 mobile phase A:B. Proteins were eluted using a 40 min linear gradient from 10 to 70% solvent B. The gluten proteins were fractionated into twenty-nine separate wells of a microtitre plate with each fraction consisting of 0.5-min elution time (0.4 ml). The fractions were dried under vacuum and resuspended in either 20 μl of 6 M or 8 M urea depending on the selected enzyme for digestion for further analysis.

Flour materials from rye, barley, oat, wheat, rice, corn and soy and GluVital™ were resuspended in 1/100 (w/v) of extraction buffer (0.5% (w/v) SDS, 2% (w/v) mercaptoethanol in PBS, pH 7.4) and placed in a mixing rotator overnight at rt. The resulting homogenate was centrifuged at 3893 ×*g* for 15 min. A 200 μl aliquot of the supernatant was diluted 10-fold in cold acetone and the proteins were left to precipitate overnight at −20 °C. Proteins were collected by centrifugation for 10 min at 12,000 ×*g* followed by the removal of the supernatant. The precipitated proteins were washed by adding 1 ml acetone and repeating the centrifugation step. Dried precipitated proteins were resuspended in either 50 μl of 6 M or 8 M urea depending on the selected enzyme for digestion for further analysis.

The total protein content was estimated by using the RCDC™ (reducing agent and detergent compatible) method (Bio-Rad, Hercules, CA, USA) using bovine serum albumin as the calibration standard. The RCDC assay is based on the Lowry protocol.

2.3. Protein digestion of entire gluten extracts

The unfractionated gluten extracts were digested using two single step digestions, one with trypsin, and one with chymotrypsin. A combined two-step digestion with LysC and trypsin was also performed. For the single step digestions, a 60 μl aliquot of the 6 M urea gluten extract, containing approx. 100 μg of protein, was reduced by adding 10 μl of 0.2 M DTT in 25 mM BCA and was incubated for 1 h at 37 °C. Cysteine residues were then blocked by the addition of 40 μl of 0.2 M iodoacetamide in 25 mM BCA and the resulting solution was incubated at rt. for 1 h in the dark. 100 μl of the reduced and alkylated protein solution were diluted by adding 240 μl of 0.1 M BCA containing 1 μg of chymotrypsin. The solutions were placed in an incubator at 25 °C overnight. For trypsin digestion, 1 μg of trypsin (1:100 ratio) was added and placed in an incubator at 37 °C overnight. Digestions were stopped by the addition of 1 μl of conc. FA to each vial.

For a two-step digestion, 50 μl of the gluten extract containing approx. 100 μg of protein in 8 M urea were reduced with 10 μl of 0.2 M DTT in 50 mM Tris-HCl pH 8 and incubated for 1 h at 37 °C. Cysteine residues were blocked by adding 40 μl of 0.2 M iodoacetamide in 50 mM Tris-HCl pH 8 and incubated at rt. for 1 h in the dark. 1 μg of LysC (1:100 ratio) was added and samples were incubated at 37 °C for 6 h. After the first digestion step, samples were diluted by adding 312 μl of 50 mM Tris-HCl (pH 8). Then, 1 μg of trypsin (1:100 ratio) was added to each sample and digestion was allowed to proceed overnight at 37 °C. Digestions were stopped by the addition of 1 μl of FA to each vial. Gluten proteins extracted from the flour materials were digested following the same protocols.

Each digest was concentrated by solvent evaporation using a vacuum centrifuge. The residue was resuspended in 100 μl of 0.1% TFA in water and the resulting solution was desalted using Bond Elut OMIX pipette tips (Agilent Technologies), by following the manufacturer's recommendations. Desalted peptides were resuspended in 100 μl of LC-MS grade water prior to LC-MS analysis.

2.4. Protein digestion of gluten fractions

The HPLC fractionated gluten proteins were digested with chymotrypsin and the two-step digestion protocol using LysC and trypsin with some small modifications. The protein content of the individual fractions, collected from the same chromatographic retention time, from three replicate sample injections were pooled. The protein fractions in 6 M urea were reduced by the addition of 2 μl 0.2 M DTT in 25 mM BCA and the fractions were incubated for 1 h at 37 °C. Cysteine residues were alkylated by the addition of 6 μl of 0.2 M iodoacetamide in 25 mM BCA and the fractions were further incubated at rt. for 1 h in the dark. Each fraction was diluted by adding 150 μl of 0.1 M BCA. 0.15 μg of chymotrypsin was then added to each pooled fraction and digestion was allowed to proceed overnight at 25 °C.

For the two-step combined digestion protocol, using LysC and Trypsin, proteins contained in the pooled fractions were solubilised in 30 μl of 8 M urea in 0.1 M Tris-HCl pH 8. Then 2 μl of 0.2 M DTT in 50 mM Tris-HCl pH 8 were added and the fractions were incubated for 30 min at 37 °C. Cysteine residues were blocked by the addition of 6 μl of 0.2 M iodoacetamide in 50 mM Tris-HCl pH 8 and the fractions were incubated at rt. for 30 min in dark. LysC (0.15 μg) was added to each pooled fraction and digestion was allowed to proceed for 6 h at 37 °C. After the first digestion step, samples were further diluted by adding 188 μl of 50 mM Tris-HCl pH 8. Trypsin (0.15 μg) was added to each pooled fraction and digestion was allowed to proceed overnight at 37 °C. Digestions were stopped by adding 1 μl of FA to each vial.

Each digest was concentrated by solvent evaporation using a vacuum centrifuge and the peptides were resuspended in 50 μl of 0.5% TFA (v/v). The peptides were isolated and desalted with PepClean™ C18 Spin Columns (Thermo Fisher Scientific, Rockford, IL, USA) by following the manufacturer's recommendations. The peptides eluted in 80 μl of 70% ACN/0.1% FA were dried in a vacuum centrifuge and the residue was resuspended in 15 μl of LC-MS grade water.

2.5. Shotgun analysis by liquid chromatography electrospray ionization tandem mass spectrometry

Peptide separation was performed using a nanoAcquity UPLC (Waters, Manchester, UK). Samples were eluted using a binary gradient consisting of: solvent A, 0.1% FA in water; solvent B, 0.1% FA in ACN. 2 μl injections of the digested samples were pre-concentrated on a Symmetry C18, 180 μm (ID) × 20 mm 5 μm, trap column for 1 min with a solvent flow of 8 μl/min consisting of 1% solvent B. Peptides were separated using an Atlantis, 75 μm (ID) × 100 mm, 3 μm particle size (Waters, Manchester, UK), analytical column using a 98 min linear gradient consisting of 0–1 min 1% B, 1 to 78 min 1 to 25% solvent B, 78–88 min

25%-40% B and 88–98 min 40%-95% solvent B. A constant flow rate of 400 nl/min was used throughout. The column was coupled with a picotip emitter (Waters, Manchester, UK; 360 μm (OD) × 20 μm (ID); 10 μm tip; 6.35 cm) and sprayed directly into the nano-ESI Z-spray source of a Synapt G2 HDMS (Waters, Manchester, UK). Data dependant acquisition (DDA) was controlled using the MassLynx v4.1 operating software optimized for product ion spectral acquisition over a 3.5-second cycle, which was repeated throughout the chromatographic run. The MS-TOF survey scan lasted 0.5 s and data were recorded in the 300–1800 *m/z* range, targeting precursor ions which exceeded a threshold of 800 counts per second. Preselected precursor ions within a 100 mDa window were excluded from reselection for the next 15 s. Following each survey spectrum the top three most intense precursors were selected for subsequent isolation and generation of product ion spectra. Product ions were generated by CID and data were recorded between 50 and 2000 *m/z* and summed for 1 s. A trap collision energy (CE) linear ramp was used from 12 V to 30 V for low *m/z* ions and from 30 V to 55 V for high *m/z* ions for each product ion spectrum. The instrument was operated in positive ion resolution mode resulting in an MS resolution of approximately 16,000 FWHM (full width at half maximum). A lock mass reference standard consisting of 100 fmol/μl glu-fibrinopeptide B was continuously sprayed through a second nano-spray lock mass taper tip emitter and the lock mass was monitored every 50-s of the chromatographic run for a duration of 0.7 s.

2.6. Targeted mass spectrometric analysis using a selected reaction monitoring (SRM) method

A Xevo TQ-S triple quadrupole mass spectrometer (Waters, Manchester, UK) equipped with a trizaic ion source was used to acquire SRM data for this study. Peptide mixtures were separated on a 150 μm × 100 mm Peptide BEH C18 130 Å, 1.7 μm UPLC column with an integral ESI emitter which formed part of the ionKey based separation system (Waters, Manchester, UK). 2 μl of the peptide standards and extracts were separated using a binary linear gradient from 10% to 40% solvent B over 15.5 min followed by a gradient from 40%-99% solvent B over 1 min using a flow rate of 1.4 μl/min. Solvent A was 0.1% FA in water and solvent B was 0.1% FA in ACN. The gradient was followed by a 3 min wash step using 99% ACN and 5 min reconditioning using 10% solvent B resulting in a total run time of 25 min. The CE applied whilst performing the individual SRM experiments was calculated using the linear relationship of $CE = 0.034 * m/z + 3.314$ (values were calculated using Skyline v2.5 software [15]), In the tune page the capillary voltage was set to 3.2 kV, the collision gas to 0.1 ml/min, and the cone voltage to 25 V. SRM transitions were first monitored using the automated dwell time function of Masslynx v4.1 by setting a minimum of 12 points per chromatographic peak. The quadrupole parameters were set to unit mass resolution for Q1 and Q3. All of the sample extracts were assayed in triplicate and a blank injection was made between each analytical replicate to check for sample carry over. The autosampler needle was rinsed with strong wash solvent (2% DMSO in 95% ACN:0.1% FA in water) followed by a weak wash solvent (1% ACN:0.1% FA in water) after each injection to minimize carry over effects. A control injection of a mixture of five standard peptides (50 ng/ml of MRFA, bradykinin, angiotensin I and II, glu-fibrinopeptide B and 200 ng/ml insulin chain B) was made every 15 injections to check retention time stability and equipment performance.

SRM assays were developed by first selecting five specific transitions from each peptide precursor ion, which corresponded to the five most intense product ions from the b- and y-type ion series. The chosen products preferably had an *m/z* greater than that of the precursor. For the final scheduled SRM method three transitions were acquired for each peptide over a retention time window of ±1 min from the initially observed retention time.

3. Data analysis of shotgun analysis

3.1. Protein identification

Raw data files were processed using Protein Lynx Global Server v2.5 (PLGS, Waters, Manchester, UK) to perform lock mass correction, noise reduction, de-isotoping and conversion from continuum to centroid data. The following parameters were used: lock spray calibration was performed using the lock spray exact mass of 785.8426 *m/z* and 684.3469 *m/z* in MS survey and MS/MS, respectively. Background subtraction was performed in PLGS using an order 5 polynomial with a threshold of 35%, data smoothing was performed using a 3 point Savitzky-Golay method with peak centroid calculation being performed at 80% peak width based on a minimum peak width of 4 channels at half height. The resulting .mzML files were submitted to a Mascot server search engine (Matrix Science, London, UK.) in order to perform database searching for protein identification. Database searching was performed against a custom database for *Triticum* sp. downloaded from NCBInr (containing 76,428 sequences, 11/06/2014) and using the following parameters: chymotrypsin or trypsin used were selected as the digestion enzyme, carbamidomethyl Cys as a fixed modification, Met oxidation and Asn and Gln deamidation as variable modifications and allowing up to 2 missed cleavage sites per peptide. Data were matched with a parent ion tolerance of ±20 ppm and a fragment ion tolerance of ±0.1 Da. Data were also searched against a random automatic decoy database. All peptides with scores less than the identity threshold ($p = 0.05$) were automatically discarded.

3.2. Selection of unique peptides

Exported .MzIdentML files from Mascot obtained individually for each replicate and fraction were uploaded into PAnalyzer [16] for protein inference identification or assigning peptides to proteins. A maximum false discovery rate of 1% was used to filter proteins reported by Mascot. The spectra from the unique peptides identified were manually revised and validated by BLASTp searching in the NCBInr database to further confirm their provenance.

3.3. Data analysis of targeted analysis

Raw data were imported into Skyline software v.2.6, to integrate the acquired SRM peak areas. The choice of precursor and product transition monitored SRM were selected from the spectral libraries built using the MS1 filtering analysis option in Skyline.

4. Results and discussion

4.1. Fractionation of gluten proteins

GluVital™ is an industrially extracted gluten product from wheat (*Triticum sp.*). According to the manufacturer, this product provides consistency in flour by improving the mechanical properties of dough and extending shelf life of finished products. A comprehensive proteomic approach was undertaken to characterize and catalogue the gluten proteins and peptides in wheat from this common gluten source. A multi-dimensional separation of the gluten proteins was achieved by RP-HPLC followed by enzymatic digestion and separation of the peptides using reverse phase chromatography coupled to high resolution tandem mass spectrometry.

Gluten proteins, extracted in 6 M urea, were reduced and alkylated by the use of DTT and iodoacetamide. This dramatically increased the solubility of the gluten proteins as observed by SDS-PAGE (data not shown). The presence of conserved cysteine residues in the protein sequences of gluten allows disulphide bond formation which makes it difficult to solubilise. The glutenin fraction in particular, which comprises large polymer aggregations via disulphide bonds, is known to be

Table 1
Number of the identified peptides when GluVital was digested using either chymotrypsin (QM) or a combined digestion with LysC and trypsin (LT). Identified peptides are shown as number of total wheat peptides and gluten peptides and are given when gluten was digested unfractionated (total) or peptides derived from digestion of the fractions of gluten after the separation by RP-HPLC. Protein searches were performed against a custom database for *Triticum* sp. and against a coeliac protein database (FARRP CD, Allergen Online).

Enzyme	LysC + Trypsin (LT) (wheat/gluten peptides)		Chymotrypsin (QM) (wheat/gluten peptides)	
Database	Total	Fractionated	Total	Fractionated
NCBInr	345/208	344/207	223/172	418/361
FARRP CD	37	67	105	163

Table 2
Number of identified gluten peptides when GluVital was digested using chymotrypsin (QM) or a combined digestion with LysC and trypsin (LT). Identified peptides are classified according to the sub-fraction of gluten protein and are shown as number of peptides and as a percentage of the total gluten peptides identified. Protein searches were performed against a custom database for *Triticum* sp. and against a coeliac protein database (FARRP CD, Allergen Online).

Enzyme	LysC + Trypsin (LT) (number peptides/ % total peptides)		Chymotrypsin (QM) (number peptides/ % total peptides)	
Gluten proteins	NCBIn	FARRP CD	NCBIn	FARRP CD
α/β gliadin	45/21.7	9/13.4	109/30.2	47/28.8
γ gliadin	38/18.4	16/23.9	69/19.1	40/24.5
Ω gliadin	–	–	25/6.9	6/3.7
HMW glutenin	59/28.5	33/49.3	48/13.3	38/23.3
LMW glutenin	41/19.8	9/13.4	91/25.2	32/19.6
Avenin-like	24/11.6	–	19/5.3	–
TOTAL	207	67	361	163

difficult to solubilise [17,18]. The gluten fraction from cereals is known by its extreme heterogeneity, in number and type of proteins. This complexity derives from: the high sequence homology; high proline (Pro) and glutamine (Gln) content and high content of hydrophobic amino acids. Being one of the more complex protein fractions in nature makes its analysis by conventional proteomics techniques challenging. The reduced gluten proteins were fractionated by RP-HPLC into 29 fractions to decrease sample complexity and separate the different types of gluten proteins prior to digestion. RP-HPLC chromatography has been classically applied for gliadin and glutenin separation and characterization of their protein components [7,19–22].

4.2. Digestion of gluten proteins

In conventional proteomics studies enzymatic digestion follows the initial protein separation step. Three endoproteases were considered to digest gluten proteins in this study: trypsin, chymotrypsin and LysC. Trypsin is often the endoprotease of choice for protein identification and quantification. However, gluten proteins contain a low number of lysine and arginine amino acids required for tryptic digestion, therefore alternative proteases are usually applied for gluten digestion. Traditionally, pepsin and chymotrypsin [23,24] have been used to digest gluten and the use of thermolysin has also been reported [25,26,27,28]. The aim of this work is to discover wheat gluten markers suitable for the development of a mass spectrometry based quantification method. Thus enzymes capable of producing highly reproducible and specific peptides, such as trypsin and chymotrypsin [29], were used in this study. Digestion of the whole gluten fraction with trypsin yielded 84 peptides (data not shown) versus 172 peptides with chymotrypsin (Table 1). Previous works have reported similar numbers of tryptic peptides. Fiedler et al., [30] obtained 66 peptides and Vensel et al. [27], 37 peptides. Subsequently, a two-step combined digestion was performed by using LysC in the first step followed by trypsin. LysC is an endoprotease used to digest proteolytically resistant and tightly folded proteins [32] which cleaves at the *C*-terminal side of lysine (Lys) thus not adding to the resulting peptide mixture complexity when combined with trypsin. LysC has been suggested to enhance the number of identified peptides, improve reproducibility and enable accurate quantification [33]. In this study, 208 gluten peptides (Table 1) where identified using the two-step digestion procedure LysC/Trypsin (LT) which dramatically

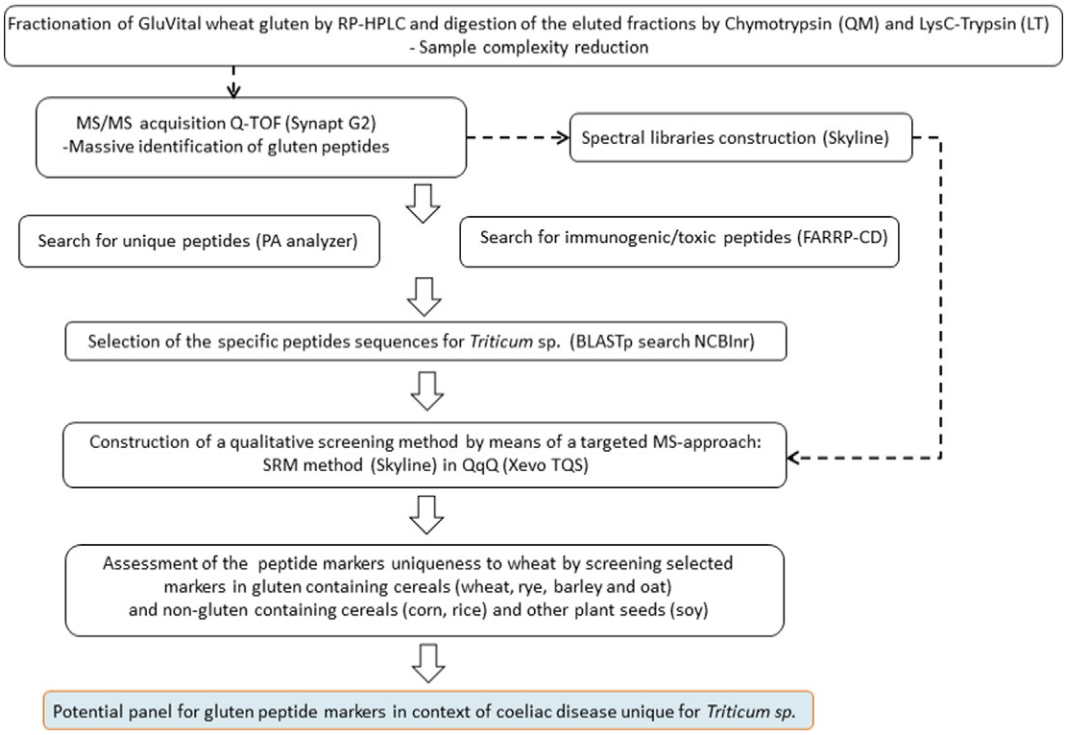

Fig. 1. Strategy followed for the identification of unique wheat gluten markers in the context of protecting consumers suffering CD in support of current EU legislation.

improved the number of identified peptides. LysC enables the use of high denaturing conditions (6 M urea) when performing the first digestion step, which seems to improve the solubility of gluten proteins. Consequently the partially digested gluten sequences are more accessible to trypsin during their secondary digestion. Thus, the digestion in two steps results in a more efficient strategy to produce tryptic peptides from gluten. The use of chymotrypsin (QM) generates peptides which when fragmented by conventional collision-induced dissociation (CID) results in product ion spectra dominated by a b-type fragment ion series. However, in this study complementary sequence information was obtained for both b- and y-type ion series as previously reported [27]. Chymotrypsin cleavage produces higher protein coverage, when compared with trypsin where the C-terminal part of the protein sequence is preferentially hydrolysed. This releases complimentary fragments from different domains in the protein sequence being studied. This observation was more acute for α/β-gliadin- and LMW/HMW-GS. The use of the two digestions methods resulted in a complementary set of peptides increasing protein coverage.

4.3. Comprehensive proteomic analysis of wheat gluten

The three replicates from each of twenty-nine fractions collected, representing each individual time point of the RP-HPLC, were pooled and digested with LT and QM separately. 344 and 418 peptides were identified from the gluten fractionated digests using LT and QM (Table 1) respectively. These comprised 60.2% and 86.4%, of the total proteins identified. The remainder of the wheat proteins identified were related to other seed storage proteins which have different roles in grain development. The 207 and 361 peptides for the gluten proteins identified from the digested gluten with LT and QM present a missed cleavage rate of 9% and 70%, respectively. The QM digestion of wheat gluten generated proteolytic resistant peptides which were still suitable for MS/MS identification. This was thought to be due to the high abundance of Pro in their sequence which can explain the high percentage of missed cleavages when using chymotrypsin. The lack of smaller fragments of these peptides seems to suggest that these are not random missed cleavages but stable proteolytic fragments that may still be

Table 3a

Peptides unique to a single gluten protein from *Triticum* sp. observed after chymotrypsin digestion (QM). Selection of the peptide set was aided by the PAnalyzer tool (Prieto et al., 2013). Specificity to *Triticum* sp. was assessed by Blastp search in the NCBInr database and via practical observations using an SRM screening method in gluten and non-gluten-containing flours. The precursors with the lowest charge state are shown (see in Suppl. material S2 for other charge states). [a]Modified amino acids in the sequence are indicated as follows: Carbamidomethylation of Cys [+57.0 Da], Met oxidation [+16.0 Da], Deamidation of Gln [+1.0 Da]; Peptides with minimum modified amino acids were considered. Known epitopes for the mAbs tested by ELISA are shown in italic and colored as follows: blue (α20 mAb), orange (R5 mAb). *Peptides contained in Table 4a. [b]Monoisotopic mass (MH^+) calculated by ProteinProspector v5.14.1 (http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct).[c]*m/z* for precursors and product ions calculated by Skyline, CE energies are given in Suppl. Material S2).

Protein	NCBInr Accession number	Sequence[a]	Retention time (min)	Missed cleavages	Monoisotopic mass(MH^+)[b]	Charge state	Transitions[c]
LMW glutenin	GI:208605344	GSSLTSIDGQ	10.6	1	964.4582	2	482.7→620.2 (y6) 482.7→519.2 (y5) 482.7→319.1 (y3)
LMW-glutenin		GVGTRVGAY	9.2	0	879.4683	2	440.2→723.3 (y7) 440.2→666.3 (y6) 440.2→214.1 (b3)
LMW-glutenin	GI:338817620	LQQC[+57.0]SQTAY	8.6	1	1098.4884	2	549.7→745.3 (b6) 549.7→846.3 (b7) 549.7→917.4 (b8)
HMW-glutenin	GI:632044	QGQ[+1.0]GPQGKQGY	10.2	0	1148.5331	2	574.7→834.4 (y8) 574.7→597.2 (b6) 574.7→782.3 (b8)
HMW-glutenin		TTSLQQSGQGQQGY	11.7	1	1482.6819	2	741.8→824.3 (y8) 741.8→531.2 (b5) 741.8→494.7 (b10)
ω-5-gliadin		QQYPQQQPSGSDVISISGL	15.7	0	2031.9982	2	1016.5→1131.5 (y12) 1016.5→901.4 (b7) 1016.5→1556.7 (b14)
ω-5-gliadin		HQQQLP*QQQFPQQQFPQQQFP*QQQQFPQQQQL	14.0	0	4016.9539	3	1339.6→741.3 (y6) 1339.6→635.3 (b5) 1339.6→1263.6 (b10)
ω-gliadin, partial	GI:63252971	SPHQPQQPFPQPQRPTPL	11.9	0	2080.0723	3	694.0→900.4 (b8) 694.0→1047.5 (b9) 694.0→808.4 (y7)
		HQQPEQIISQQPQQPF	11.8	0	1932.9563	3	644.9→748.3 (b6) 644.9→861.4 (b7) 644.9→616.3 (y5)
ω-gliadin		ILQPQQPLPQQPQQPF	14.2	1	1887.0123	2	944.0→616.3 (y5) 944.0→708.4 (b6) 944.0→918.5 (b8)
γ-gliadin	GI:205363280	ALRTLPTM[+16.0]C[+57.0]NVY	13.4	1	1454.7130	2	727.8→341.2 (b3) 727.8→442.2 (b4) 727.8→900.4 (b8)
γ-gliadin		LPLSQQQQVGQGSLVQGQGIIQPQQPAQL*	15.3	2	3068.6487	3	1023.5→667.3 (b6) 1023.5→795.4 (b7) 1023.5→923.4 (b8)
α-gliadin		ALQTLPAM[+16.0]C[+57.0]NVY	12.9	1	1396.6599	2	698.8→702.2 (y5) 698.8→414.2 (b4) 698.8→695.4 (b7)
α/ß gliadin		*RPQQPYP*QPQPQYSQPQHPISQQQQQQQQQQQQE QQIL*	11.6	1	4649.2628	4	1163.0→510.2 (b4) 1163.0→770.3 (b6) 1163.0→995.5 (b8)
Gliadin/avenin-like	GI:281335538	TTISPSSDVTTDM[+16.0]GGY	11.1	0	1647.7054	2	824.3→760.2 (y7) 824.3→659.2 (y6) 824.3→558.1 (y5)

suitable for quantification. The existence of large proteolytic resistant peptides has been previously reported, such as the 33 amino acid sequence (LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF), when performing digestions with QM. This was thought to be due to the repetitive motif of PQX which is only weakly susceptible to chymotrypsin cleavage [34]. The combined digestion with LT reduced the number of missed cleavages observed when searching for tryptic peptides.

Highly resistant peptides are also released upon the *in-vivo* digestion of gluten by major gastric (pepsin) and pancreatic proteases (trypsin, chymotrypsin, elastase and carboxypeptidases A and B) and to intestinal brush border membrane proteolysis under physiologically conditions. It is important to understand the link between the molecular fate of gluten under physiological conditions and the *in silico* and *in vitro* digestion of the gluten proteins. Most of the immunogenic or toxic peptides are proteolytic resistant and therefore will present missed cleavage sites in their sequences after chymotryptic digestion [35,36].

Fractionation improved the number of peptides identified, especially for QM, when compared with the digestion of the whole gluten (208 vs. 207 for LT and 172 vs. 361 for QM digestions) increasing the protein coverage of the gluten proteome. The lack of improvement in the LT digestion can be attributed to the low number of cleavage sites for trypsin in the gluten proteins.

When gluten was digested with QM, 203 gliadin peptides accounted for 56% of the total gluten peptides identified. While for tryptic digestion, 83 gliadin peptides accounted for 40% of total gluten peptides observed. Gliadin/glutenin ratios in wheat have been reported to range from 1.5–3.1, [10]. When gluten was digested with QM, 109, 69 and 25 of the resulting peptides were identified from the gliadin subfractions α-, γ- and Ω-gliadin respectively. When the LT digestion protocol was used, the number of peptides identified was 45 and 38 for α- and γ-gliadin respectively. No tryptic peptides were observed or identified for Ω-gliadin. The Ω-gliadins mainly consist of a single repetitive domain where 80% of the sequence is composed by glutamine (Gln), glutamic acid (Glu), proline (Pro) and phenylalanine (Phe) residues [37]. This may explain the lack of tryptic peptides observed. The remaining gluten proteins identified correspond mainly to glutenins, HMW-GS and LMW-GS which accounted for 13% and 25% respectively, when digested with QM, and 29% and 10% respectively, when the LT protocol was used. In addition to the main sub-group of gliadin and glutenins, other seed storage proteins were identified and these comprised 5% and 12% of the gluten proteome when digestion was performed with QM and LT respectively. These proteins have homologous sequences to gluten proteins and are called avenin-like proteins due to their similarity to oat avenins.

LT and QM digestions generated a pool of peptides representative of all the sub-groups of gluten proteins. The number of proteins identified from the different fractions was consistent with the content of the specific gluten proteins in different wheat cultivars [17,19] thus, giving a good representation of the wheat gluten proteome. The consistency in the number of peptides and coverage observed suggests the abundance for each gluten sub-group, of glutenin and gliadin, could be indicative of the ratio of these proteins present. The observed ratio of 1:1.2–1.4 differs from the current convention of 1:1 when converting gliadin content

Table 3b
Peptides unique to a single gluten protein from *Triticum* sp. observed after LysC and trypsin digestion (LT). Selection of the peptide set was aided by the PAnalyzer tool (Prieto et al., 2013). Specificity to *Triticum* sp. was assessed by Blastp search in the NCBInr database and via practical observations using an SRM screening method in gluten and non-gluten-containing flours. The precursors with the lowest charge state are shown (see in Suppl. Material S3 for other charge states). [a]Modified amino acids in the sequence are indicated as follows: Carbamidomethylation of Cys [+57.0 Da], Met oxidation [+16.0 Da], Deamidation of Gln [+1.0 Da]; Peptides with minimum modified amino acids were considered. [†]Peptides whose SRMs signals were detected also in rye. [b]Monoisotopic mass (MH^+) calculated in ProteinProspector v5.14.1 (http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct).[c]m/z for precursors and product ions calculated in Skyline, CE energies are given in Suppl. Material S3).

Protein	NCBInr Accession number	Sequence[a]	Retention time (min)	Missed cleavages	Monoisotopic mass(MH^+)[b]	Charge state	Transitions[c]
HMW glutenin	GI:262,205,115	VAKNQQLAAQLPAMC[+57.0]R	13.9	1	1798.9415	3	600.3 → 634.2 (y5) 600.3 → 853.4 (b8) 600.3 → 1165.6 (b11)
	GI:37,575,363	QQPGQGQQPEQGQQPGQGQQGYYPTFPQQPGQGK	11.8	0	3681.7065	3	1227.9 → 1184.6 (y11) 1227.9 → 839.4 (y8) 1227.9 → 1082.0 (y20)
LMW glutenin	GI:513,130,001	SQ[+1.0]MLQQSIC[+57.0]HVMQQ[+1.0]Q[+1.0]C[+57.0]RQ[+1.0]Q[+1.0]LR	13.4	1	2692.1946	3	898.0 → 1120.5 (y8) 898.0 → 1248.5 (y9) 898.0 → 1379.6 (y10)
	GI:56,126,405	VFLQQQC[+57.0]SPVAMLQ[+1.0]SLAR[†]	12.9	0	2077.0569	3	693.0 → 446.2 (y4) 693.0 → 991.4 (b8) 693.0 → 1187.5 (b10)
γ Gliadin		LQC[+57.0]QAIHNVVHAIILHQQQK	13.4	0	2378.2874	4	595.3 → 668.3 (y5) 595.3 → 781.4 (y6) 595.3 → 1215.6 (y10)
	GI:209,971,811	RPLFQLIQGQ[+1.0]GIIRPQQPAQLEVIR	18.1	0	2899.6629	3	967.2 → 1053.6 (y9) 967.2 → 1278.7 (y11) 967.2 → 1621.9 (b14)
	GI:513,129,949	NYLLQQC[+57.0]DPVSLVSSLVSM[+16.0]ILPR	22.8	0	2648.3786	3	883.4 → 765.8 (b13) 883.4 → 831.4 (y7) 883.4 → 1118.6 (y10)
α/β-gliadin		VSQQSYQLLQQLC[+57.0]C[+57.0]QQLWQTPEQSR	16.5	0	3136.4939	3	1046.1 → 1720.7 (y13) 1046.1 → 1961.8 (y15) 1046.1 → 1992.9 (b16)
	GI:383,210,739	VSQQSYQLLQQLC[+57.0]C[+57.0]LQLWQTPEQSR	18.4	0	3121.5194	3	1041.1 → 1705.7 (y13) 1041.1 → 1946.9 (y15) 1041.1 → 1977.9 (b16)
		NLALQ[+1.0]TLPAM[+16.0]C[+57.0]NVYIPPYC[+57.0]TIVPFGIFGTN	19.1	0	3401.6619	3	1134.5 → 541.2 (b5) 1134.5 → 642.3 (b6) 1134.5 → 755.4 (b7)
Avenin-like	GI:474,329,936	TAWEPHHPSSPEQQPTPQPQEQPVPHQK	9.3	0	3222.5351	5	645.3 → 705.4 (y6) 645.3 → 1187.6 (y10) 645.3 → 1356.5 (b12)
	GI:281,335,538	SAWEPQHPSSPEHQPTPQPQEHPVPHQK	8.9	0	3217.5198	5	644.3 → 509.2 (y4) 644.3 → 1196.6 (y10) 644.3 → 1421.7 (y12)

Table 4a

Peptides containing toxic/immunogenic sequences from *Triticum* sp. observed after chymotrypsin digestion (QM). Selection of the peptide set was aided by the exact BLAST celiac disease (CD) novel protein risk assessment tool provided on the AllergenOnline website (http://www.allergenonline.org/celiacfasta.shtml). Specificity to *Triticum* sp. was assessed by Blastp search in the NCBInr database and via practical observations using an SRM screening method in gluten and non-gluten containing flours. The precursors with the lowest charge state are shown (see in Suppl. Material S2 for other charge states). [a]Toxic/Immunogenic sequences are underlined; Known epitopes for the mAbs tested by ELISA are shown in italic and colored as follows: green (G12 mAb), blue (α20 mAb), orange (R5 mAb); [#]33-mer peptide target of G12 mAb. Peptides with minimum modified amino acids were considered. *Unique peptides contained in Table 3a. [†]Peptides whose SRMs signals were detected also in rye. [b]Monoisotopic mass (MH^+) calculated by ProteinProspector v5.14.1 (http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct). [c]m/z for precursors and product ions calculated by Skyline, CE energies are given in Suppl. Material S2).

Protein	Sequence[a]	Retention time (min)	Missed cleavages	Monoisotopic mass (MH^+)[b]	Charge state	Transitions (m/z)[c]
HMW-glutenin	QQSGQGQHGYYPTSPQLSGQGQRPGQW[†]	11.9	2	2957.3674	3	986.4→1438.71 (y13) 986.4→1213.60 (y11) 986.4→719.86 (y13++)
	YPGQASPQRPGQGQQPGQGQQSGQGQQGYYPTSPQQPGQW	12.2	1	4268.9517	4	1067.9→1125.5 (y10) 1067.9→840.3 (y7) 1067.9→1580.7 (b15)
	QQGYYPTSPQQLGQGQQPRQW	12.7	2	2475.1800	3	825.7→1835.9 (y16) 825.7→1084.5 (y9) 825.7→899.4 (y7)
	QRPGQGQPGYYPTSPQQPGQEQQSGQAQQSGQW	11.2	1	3613.6439	3	1205.2→1232.5 (b11) 1205.2→1517.7 (b14) 1205.2→1870.8 (b17)
LMW-glutenin	SQQQQPPFSQQQPPFSQQQQQPL[†]	13.7	2	2709.3016	2	1355.1→941.4 (b8) 1355.1→1297.6 (y11) 1355.1→1412.6 (b12)
ω-gliadin	SQQLEQTISQQP*QQPFP*QQPHQPQQPYPQQQPY	12.8	1	3956.8951	4	989.9→760.3 (y6) 989.9→279.1 (y2) 989.9→928.4 (b8)
	SQQQISQQPQQLPQQQQIPQQPQQF	12.5	0	2988.4922	3	996.8→1707.8 (y14) 996.8→1594.7 (y13) 996.8→1394.7 (b12)
	QQHQIPQQPQQFPQQQQF	11.8	0	2263.1003	3	755.0→775.3 (y6) 755.0→988.4 (b8) 755.0→1488.7 (b12)
γ-gliadin	LQPQQPQQSFPQQQQPL	12.3	1	2020.0247	2	1010.5→838.4 (y7) 1010.5→1182.5 (b10) 1010.5→1791.8 (b15)
	VQGQGIIQPQQPAQL[†]	13.1	0	1604.8755	2	802.9→781.4 (y7) 802.9→696.4 (y7) 802.9→1177.6 (b11)
	VPLSQQQQVGQGTLVQGQGIIQPQQPAQL	15.3	2	3068.6487	3	1023.5→653.3 (b6) 1023.5→781.4 (b7) 1023.5→909.4 (b8)
	LPLSQQQQVGQGSLVQGQGIIQPQQPAQL*	15.3	2	3068.6487	3	1023.5→667.3 (b6) 1023.5→795.4 (b7) 1023.5→923.4 (b8)
α/β-gliadin	LQ*LQPFP*QPQLPYPQPHLPYPQPQPF[†]	18.3	2	3095.6142	3	1032.5→1320.6 (y11) 1032.5→1775.9 (b15) 1032.5→1873.0 (b16)
	LQ*LQPFP*QPQLPYPHPQLPYPQPQPF	18.3	2	3095.6142	3	1032.5→1214.6 (y10) 1032.5→1784.9 (b15) 1032.5→1882.0 (b16)
	MQ*LQPFP*QPQLPYP*QPQLPY*PQPQPF	18.3	2	3104.5703	3	1035.5→973.5 (y8) 1035.5→713.3 (y6) 1035.5→970.4 (b8)
	LQLQPFP*QPQLPY*P*QPQLPY*P*QPQLPY*PQPQPF[#†]	19.5	2	3910.0367	3	1304.0→973.4 (y8) 1304.0→713.3 (y6) 1304.0→824.4 (b7)
	*RPQQPYP*QPQPQYSQPQHPISQQQQQQQQQQEQQIL*	11.6	1	4649.2628		1163.0→510.2 (b4) 1163.0→770.3 (b6) 1163.0→995.5 (b8)

to gluten content. This may partially explain the differences observed between ELISA kits, from the different suppliers, that aim to measure total gluten but often target different gluten proteins [4].

The data suggests that each enzyme has a preference for different substrates and can be an effective cleavage agent for specific gluten proteins. A higher proportion of α-gliadin peptides were identified when digestion was performed with QM in comparison with LT. While, γ-gliadin peptides observed were of similar abundance when either LT or QM was used for digestion. However, no peptides from the Ω-gliadin were identified when using the LT digestion. This suggests that QM is a suitable enzyme for the digestion of gliadin-type proteins. In the case of the glutenin fraction, including other minor types of gluten-like proteins, LT digestion is a good option and seems a more effective endoprotease to digest the HMW glutenin sub-group. This confirms the observations of Mucilli et al. [38], who compared HMW-GS gluten protein content between different durum wheat cultivars using trypsin digestion. In contrast, the use of LT was not advantageous for LMW-GS digestion, which is explained by their similarity in amino acid composition and protein structure to gliadin type proteins [17].

There was an overlap of the sequences identified between both sets of peptides observed when digesting gluten with QM and LT. A common set of peptides were observed whereby, 75 peptides out of 361 for QM and 59 peptides out of 207 for LT, had partially overlapping sequences. Thus, the combined information from both sets of peptides which do not share overlapping amino acid sequence results in a complete dataset of 434 different peptide sequences covering gluten. The combination of two digestions enabled the detection of 286 and 148 different peptide sequences for QM and LT digestions, which emphasized the complementary nature of the dual digestion procedure. The importance of the use of a multi-enzymatic approach at different levels either in the digestion [24,36] or data analysis steps [26] has been reported in the literature for the comprehensive identification of gluten proteins. However, the number of identifications is still limited by: the current information in protein public databases; the fact that protein sequence information for many cereal species is incomplete as their genomes are not yet fully sequenced. Efforts to sequence the 17 gigabase genome of *Triticum aestivum* and its comparison with eight related wheat genomes has recently been published [39].

4.4. Defining the wheat gluten fingerprint: Potential set of specific peptide markers for wheat gluten detection

Peptide identification was facilitated using two complementary digestions, QM and LT, which allowed the detection of specific peptides from wheat gluten following the strategy described in Fig. 1. A set of peptide markers specific for *Triticum* sp. were selected which included, peptides unique in protein sequence and peptides containing known immunogenic/toxic sequences in the context of CD. A crucial issue for gluten quantification is the selection of the gluten peptide markers. The legislation concerning the labelling of gluten-free food products does not specify what is to be measured. As reported, the best hazard management and risk assessment processes require methodology that measures the hazard itself [40]. Thus, the proposed set of specific peptides markers to wheat includes those peptides identified as containing known clinically relevant sequences triggering CD. Also, an effort has been made to include peptides markers unique to a single protein sequence to act as indicators of gluten quantity.

In order to find peptides from gluten which are unique to a single protein, the Mascot search results from all twenty-nine digested fractions from QM and LT were combined using the software tool PAanalyzer [16]. The issue of protein inference is a well-known problem in shotgun proteomics studies [41] and proteomic strategies based on proteome digestion followed by MS/MS-based peptide sequencing [42,43]. One such issue resides in the fact that the same set of peptides can be shared by multiple different proteins. The shared peptides can lead to ambiguities in reported protein identities from samples where connectivity between peptides and proteins is lost at the digestion stage. The difficulty of relating individual peptides back to the protein level is increased in the case of gluten where the homology between sequences is extremely high. To address this issue, PAnalyzer was used to search for unique peptide sequences from the set of 568 peptides identified, 361 for QM and 207 for LT, and correlate these peptides to protein sequences in GluVital™ wheat gluten (Table 2). A unique set of 83 peptides were found in both digestions, 48 and 35 from the QM and LT digestions respectively (Table 3a, 3b).

In order to meet our second criteria, to select peptides that inform on the toxic nature of the sample, the sequences from the 568 peptides identified were searched against the FARRP allergen database (http://www.allergenonline.org/celiachome.shtml). The filtered peptides sequences from both digestions are shown in the Supplementary material (S1). The AllergenOnline database comprised 68 protein sequences and 1016 naturally occurring, mutated or deamidated (Gln converted to Glu by tissue transglutaminase) peptides from wheat and wheat relatives (barley, rye and two proteins from oats) that have been demonstrated to elicit coeliac disease or activate MHC Class II restricted T cells of coeliac patients. A database containing these sequences was compiled and used in the MS1 filtering approach (described in Materials and Methods

Table 4b
Peptides containing toxic/immunogenic sequences from *Triticum* sp. observed after Lysc and trypsin digestion (LT). Selection of the peptide set was aided by the exact BLAST celiac disease (CD) novel protein risk assessment tool provided on the AllergenOnline website (http://www.allergenonline.org/celiacfasta.shtml). Specificity to *Triticum* sp. was assessed by Blastp search in the NCBInr database and via practical observations using an SRM screening method in gluten and non-gluten containing flours. The precursors with the lowest charge state are shown (see in Suppl. Material S3 for other charge states) and the peptides with minimum modified amino acids were considered. [a]Toxic/Immunogenic sequences are underlined. [b]Monoisotopic mass (MH^+) calculated by ProteinProspector v5.14.1 (http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct).[c]m/z for precursors and product ions calculated by Skyline, CE energies are given in Suppl. Material S3).

Protein	Sequence*	Retention time (min)	Missed cleavages	Monoisotopic mass (MH^+)[b]	Charge state	Transitions[c]
γ-Gliadin	QGVQILVPLSQQQQVGQGTLIQGQGIIQPQQPAQLEVIR	17.5	0	4218.3357	4	1055.3 → 925.5 (y8) 1055.3 → 1278.7 (y11) 1055.3 → 1406.7 (b27)
	QPFPQQPQQPYPQQPQQPFPQTQQPQQPFPQSK	13.1	0	3927.9201	4	982.7 → 1056.5 (y9) 982.7 → 1079.5 (b9) 982.7 → 1339.6 (b11)
HMW glutenin	QGYYPTSLQQPGQGQQIGQGQQGYYPTSPQHTGQR	12.0	0	3863.8121	3	1288.6 → 554.7 (y10++) 1288.6 → 1271.6 (y10) 1288.6 → 1989.9 (y18)
	GQQGYYPTSLQQPGQGQQGYYPTSLQHTGQR	12.4	0	3453.6207	3	1151.8 → 534.2 (b5) 1151.8 → 1124.5 (y10) 1151.8 → 1287.6 (y11)
	QQPVQGQQPEQGQQPGQWQQGYYPTSPQQLGQGQQPR	12.3	0	4189.9823	4	1048.2 → 1236.6 (y11) 1048.2 → 1323.6 (y12) 1048.2 → 1521.7 (y14)

section) as a background proteome. This highlighted 163 and 67 peptides, for QM and LT digestions respectively, as belonging to the proteins contained in the database, but not necessarily that the peptide was either toxic or immunogenic (Table 2). A third Blast search was performed on this reduced set of peptides using the exact BLAST Celiac Disease (CD) Novel Protein Risk Assessment Tool provided on the AllergenOnline website (http://www.allergenonline.org/celiacfasta.shtml). 54 and 12 peptides for QM and LT were found to have recognized toxic and immunogenic sequences. All sub-groups of the gluten proteins were represented in both digestions. QM released peptides containing sequences that previously have been demonstrated to be involved in triggering coeliac disease, while only 12 such peptides could be identified with LT. Proven peptides susceptible to elicit an immunogenic repose in terms of CD are therefore good candidates for the systematic screening of foods.

In order to meet our third criteria for selecting peptides, a total of 102 peptides for QM and 47 for LT, were BLASTp search in the NCBInr

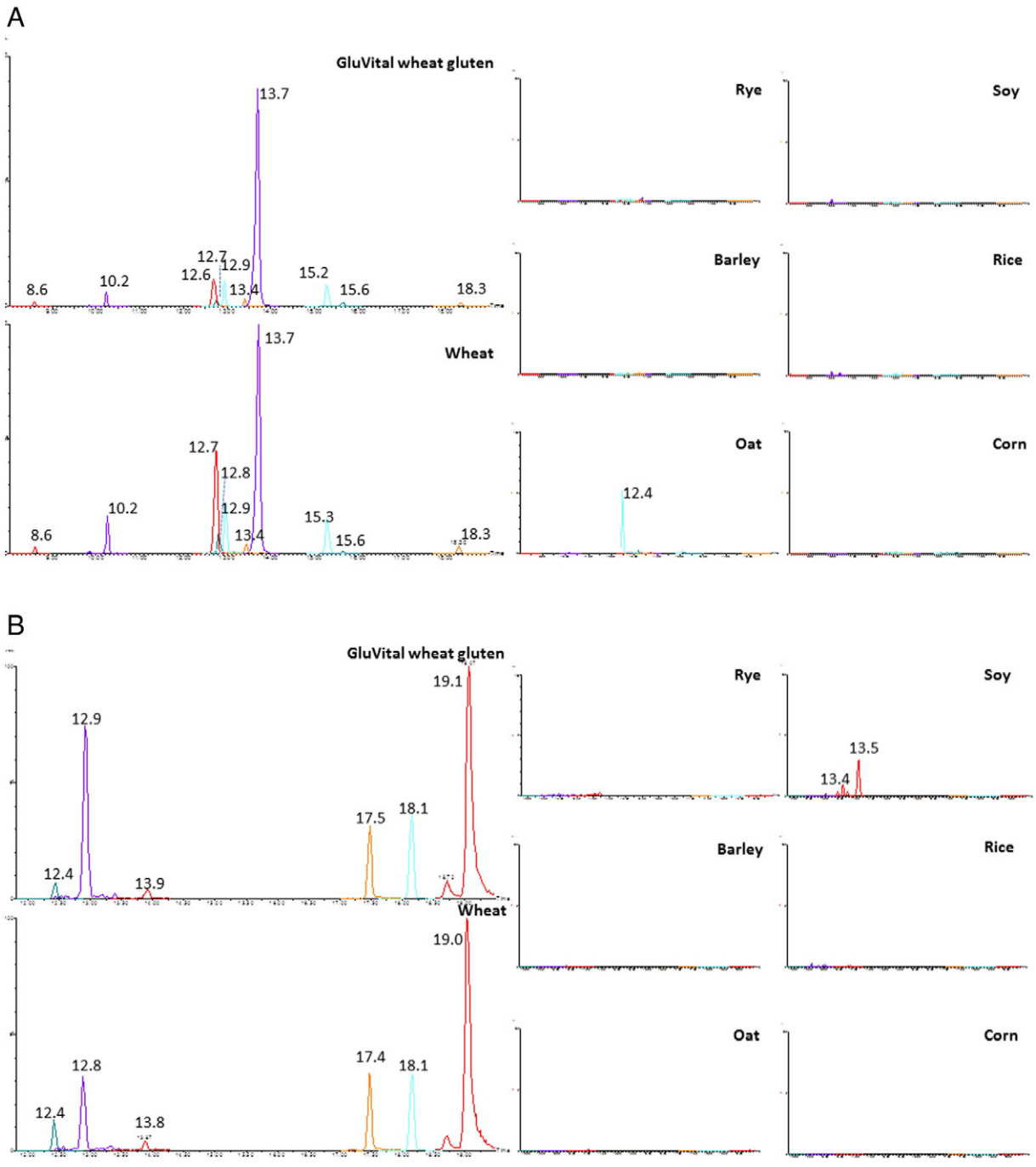

Fig. 2. Selection of a set of peptides from Tables 3a, 3b and 4a, 4b as an example of the SRM signals after screening with the final SRM method of the GluVital™ wheat gluten, the gluten-containing flours (wheat, rye, barley and oats), non-gluten containing flours (corn and rice) and the plant seed flour, soya. The retention time (RT) for each peptide is annotated at the top of the chromatographic peak. A) Peptides derived from gluten digested with chymotrypsin (QM), unique (Table 3a): LQQCSQTAY (8.6 min, LMW-GS), QGQGPQGKQGY (10.2 min, HMW-GS), QQYPQQQPSGSDVISISGL (15.6 min, omega-gliadin), ALRTLPTMCNVY (13.4 min, gamma-gliadin), ALQTLPAMCNVY (12.9 min, alpha gliadin); immunogenic/toxic peptides (Table 4a): QQGYYPTSPQQLGQGQQPRQW (12.7 min, HMW-GS), SQQLEQTISQQPQQPFPQQPHQPQQPYPQQQPY (12.8 min, Ω-gliadin), VPLSQQQQVGQGTLVQGQGIIQPQQPAQL (15.3 min, γ-gliadin), LQLQPFPQPQLPYPHPQLPYPQPQPF (18.3 min, α-gliadin). B) Peptides derived from GluVital™ gluten digested with LysC and trypsin (LT), unique (Table 3b): VFLQQQCSPVAMLQSLAR (12.9 min, LMW-GS), VAKNQQLAAQLPAMCR (13.9 min, HMW-GS), RPLFQLIQGQGIIRPQQPAQLEVIR (18.1 min, γ-gliadin), NLALQTLPAMCNVYIPPYCTIVPFGIFGTN (19.1 min, α-gliadin); immunogenic/toxic peptides (Table 4b): GQQGYYPTSLQQPGQGQQGYYPTSLQHTGQR (12.4 min, HMW-GS), QGVQILVPLSQQQQVGQGTLIQGQGIIQPQQPAQLEVIR (17.5 min, γ-gliadin).

database to select peptides that are specific to *Triticum* sp. A fourth filter was applied whereby only peptides exhibiting a good product ion sequence coverage, enabling confident identification, were selected. Identification of proteins by MS/MS is often based on the best hit of the spectral data to databases of protein sequences. For complex and highly homologous sequences, such as gluten proteins, where the difference can be a few amino acids in a protein sequence or a single amino acid switching positions, the careful interpretation of product ion data is essential for correct sequence assignment. Heterogeneity of gluten proteins is due to numerous gene duplications and subsequent divergences of multigene families encoding a highly polymorphic set of proteins [44].

After applying the four criteria above, 77 and 31 peptides for QM and LT digestions were identified as markers and selected to build an SRM method to specifically target these peptide sequences using a sensitive triple quadrupole mass spectrometer. For each peptide, 5 specific transitions (from the y and b ion series) were selected based on the spectral libraries compiled in the MS1 filtering. A maximum of 10 precursors were monitored sequentially in any one MS method resulting in 50 transitions each with a dwell time of 36 ms over the 25 min run. After performing this initial SRM method for each peptide and determining the appropriate retention time, a scheduled SRM method was developed. The SRM screening method consisted of the three most abundant transitions from the selected peptide signals that were detected from the digestion of an unfractionated (total) gluten sample and that were free from interferences from similar peptides. This resulted in a final number of peptides from the QM and LT digestions of 45 and 26 peptides for the total gluten digests respectively.

The SRM methods were used to screen digested gluten-containing cereals of rye, barley, wheat and oat flours and non-gluten containing flours from corn and rice and soya. The screening of these materials enabled the selection of specific peptide markers for wheat. Protein extracts from flours and GluVital™ wheat gluten as described in the materials and methods were digested in triplicates by QM and LT and were screened for the presence of the selected peptides. Positive identifications were based on: chromatographic retention time; the observed ratios for the multiple transitions from the same precursor; co-eluting well-defined SRM chromatograms for all nine analytical and technical replicates.

A final SRM method containing specific wheat peptide markers, detected only in wheat flour and GluVital™ wheat gluten was constructed. It contained 14 unique (Table 3a) and 15 immunogenic/toxic peptides (Table 4a) from QM digestion and 12 unique (Table 3b) and 5 immunogenic/toxic (Table 4b) from LT digestion. Two unique peptides from QM digestion were found also to containing immunogenic/toxic sequences (Tables 3a and 4a).The Skyline template for the final SRM method for each digestion is provided in the Supplementary materials (S2 and S3) with information about the individual precursor and product *m/z*, RT and peptide modifications. As an example, the SRM signals for specific peptides from the gluten protein subgroups, containing unique and immunogenic/toxic sequences from each digestion (QM and LT), are shown in Fig. 2A and B.

Besides the hydrophobic nature and long amino acid sequence for most of the proposed marker peptides, the good chromatographic reproducibility of ± 10 s, quality of SRMs (as described in materials and methods section) and their reproducible detection in all the extraction and digestion replicates (data in Supplementary material S4), demonstrate the suitability of these markers for developing an MS-method for gluten detection and quantification.

Several peptides were identified as not being unique to wheat, as the same SRMs were also observed in the rye flour (Table 5). The SRMs for

Table 5
Peptides not specific to *Triticum* sp. SRMs were found in other gluten-containing cereal source as indicated in the table. The precursors with the lowest charge state are shown (see in Suppl. Materials S2 and S3 for other charge states). [a]Toxic/Immunogenic sequences are underlined; [*]Peptides common to all the gluten containing cereals. Peptides with minimum modified amino acids were considered. [b]Monoisotopic mass (MH^+) calculated by ProteinProspector v5.14.1 (http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct). [c]m/z for precursors and product ions calculated by Skyline, CE energies are given in Suppl. Material S2 and S3).

Protein	Sequence[a]	Cereal	Retention time (min)	Missed cleavages	Monoisotopic mass(MH^+)[b]	Charge state	Transitions[c]
γ-Gliadin	SQQQQVGQGSLVQGQGIIQPQQPAQL	Rye, Wheat	14.3	1	2745.4278	3	915.8 → 699.3 (b6) 915.8 → 781.4 (y7) 915.8 → 909.4 (y8)
	SQQPQQTFPQPQQTFPHQPQQQVPQPQQPQQPF	Rye, Wheat	13.6	0	3907.8899	3	1303.3 → 945.4 (b8) 1303.3 → 969.4 (y8) 1303.3 → 1194.5 (y10)
	YQQQQVGQGTLVQGQGIIQPQQPAQL	Rye, Wheat	14.5	2	2835.4748	3	945.8 → 428.2 (y4) 945.8 → 775.3 (b6) 945.8 → 781.4 (y7)
	SQQQQVGQGILVQGQGIIQPQQPAQL	Rye, Wheat	15.5	1	2771.4799	3	924.4 → 781.4 (y7) 924.4 → 1054.5 (b10) 924.4 → 1266.6 (b12)
	FQLVQGQGIIQPQQPAQL*	Rye, Barley, Wheat, Oat	15.9	2	1993.0865	2	997.0 → 428.2 (y4) 997.0 → 781.4 (y7) 997.0 → 971.5 (b9)
	RPLFQLVQGQGIIQPQQPAQLEVIR*	Rye, Barley, Wheat, Oat	16.5	0	2856.6207	3	952.8 → 1278.7 (y11) 952.8 → 925.5 (y8) 952.8 → 1578.9 (b14)
γ-Gliadin	QPQQPFPQPQQPQQSFPQQQPSLIQQSLQQQLNPC[+57.0]K	Rye, Barley, Wheat	15.7	0	4197.0934	4	1064.2 → 404.1 (y4) 1064.2 → 1304.6 (b11) 1064.2 → 951.4 (b8)
α/β-gliadin	VRVPVPQLQPQNPSQQQPQEQVPL*	Rye, Barley, Wheat, Oat	14.8	1	2734.4635	2	1367.7 → 1017.6 (b9) 1367.7 → 1356.7 (b12) 1367.7 → 1925.0 (b17)
	RPSQQNPQAQGSVQPQQLPQFEEIRNL*	Rye, Barley, Wheat, Oat	15.0	1	3117.5824	3	1039.8 → 1506.7 (b14) 1039.8 → 1135.5 (b10) 1039.8 → 1378.6 (b13)
LMW glutenin	SHIPGLERPSQQQPLPPQQTL*	Rye, Barley, Wheat, Oat	12.5	1	2351.2466	3	784.4 → 338.1 (b3) 784.4 → 683.3 (y6) 784.4 → 1458.7 (b13)
HMW-glutenin	NQQLAAQLPAMC[+57.0]R	Rye, Barley, Wheat	14.9	0	1443.7195	2	750.8 → 634.2 (y5) 750.8 → 946.4 (y8) 750.8 → 1017.4 (y9)

two peptides were found in rye and barley flours (Table 5). Other peptides exhibited measurable signals with an intensity much lower than that observed in wheat flour (peptides marked with the sign of a cross (†) in Tables 3a, 3b and 4a), which were included in the final method as the intensities of the SRMs were found to be between 50 and 600-fold lower in rye flour when compared to wheat flour. Due to the low abundance of these peptides in the other flours these peptides may still provide useful indicators for the quantity of wheat present. However, a much broader survey on different wheat and flour samples would be required to prove this.

A set of peptides common to the gluten-containing cereals, wheat, rye, barley and oat, which were not observed in the non-gluten cereals tested were identified (Table 5). These peptides have the potential to enable the quantification of gluten irrespective of the cereal source, enabling the detection of gluten in compliance with current legislation for the labelling of gluten-free food products (Regulation (EC) No 41/209 and Regulation (EU) No 828/2014). Five SRMs, originally predicted to be unique to wheat according with current public databases, were found to be common to gluten containing-cereals. However, the current study can't distinguish if the same peptide sequences are present in the other cereals as very similar sequences may share the same SRM transitions. An alternative explanation is that these peptides may exist in these flours due to the close sequence homology between the cereal species. The presence of these sequences was not predicted by the current databases. However, a complete and comprehensive set of protein sequences regarding all the cereals studied is not currently available.

Grain-specific marker peptides for wheat, rye and barley have previously been identified by MS/MS [30,45] and by an SRM targeted approach [46,47]. In the former study candidate markers from flours, were identified based on a DDA method and precursor ion alignment in Skyline. In the later studies, an SRM approach was investigated for the detection of unique peptides enabling the speciation of gluten-cereals. However, in both cases the identified targets were not specific for gluten peptides in the context of CD. A set of pathogenic peptides have been identified, but their specificity to the cereal source was not studied [48]. Two further studies have addressed the issue of gluten quantification by developing an MS-based method. An SRM quantitative method to target six peptides [24] and an MS^E approach to quantify gluten proteins [49]. The number of recent investigations about the detection of gluten peptides markers show the relevance of the issue and the critical need to identify and agree on a set of robust gluten markers suitable for future use in the MS quantification of gluten. However, a complete set of gluten markers indicative of the species specific source to target all the sub-fractions of the gluten proteins for the cereals included in current EU regulations have not yet been defined. It is generally agreed that detection methods for gluten used to support a gluten-free statement, must target toxic sequences, should refer to gluten and not target individual gluten sub fractions [11,40].

The set of gluten markers proposed here, meet different characteristics by: containing toxic/immunogenic fractions in the context of coeliac disease; containing unique diagnostic peptides for wheat; have potential for the quantification of total gluten after further MS method development. The proposed wheat gluten fingerprint comprises coeliac toxic and unique peptides for *Triticum* sp. that completely represent all gluten protein fractions, as each single gluten protein sub-fraction is represented by a peptide detected with the SRM method. Diagnostic unique peptides and peptides containing toxic/immunogenic sequences for each type of gluten protein were identified: α/β-gliadin with 6QM/3LT peptides; γ-gliadin with 5QM/5LT peptides; Ω-gliadin 8QM peptides; HMW glutenin 6QM/4LT and LMW-GS 4QM/2LT. Eight marker peptides derived from QM digestion include in their sequences known epitopes for the G12 mAb, QPQLPY [50], α20 mAb, RPQQPYP [51] and R5 mAb, QQPFP [52] (Tables 3a and 4a). The 33mer peptide (LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF), the target for raising the G12 mAb, is also included in the set of markers. Using peptide markers that contain ELISA epitopes may help in establishing a link between both MS and ELISA platforms and may enable comparison of results. Although existing studies show issues in the correlation of the relative content of gluten peptides determined by MS and ELISA gluten content [53,54].

This is the first study that compiles a comprehensive set of gluten peptide markers for MS-method development. The targets contain species-specific sequences as well as common peptides for all gluten cereals, are representative of all sub-groups of gluten proteins, contain toxic sequences for CD and peptides containing epitopes for the different mAbs used in commercially available ELISA test kits for gluten.

5. Conclusion

Combined multiple enzymatic digestion and MS was a powerful tool that enabled the detection of specific peptide sequences from all the sub-groups of gluten proteins. The final targeted screening approach based on SRM enabled the detection of specific peptides representative of gluten from wheat whilst also indicating the presence of more generic gluten peptides. Many of the species specific gluten peptides initially identified as markers, using the conventional proteomics approach of product ion spectra combined with database searching, gave rise to measurable non-specific signals in other cereal flours when using SRM. Therefore, over reliance on discovery based approaches alone is not sufficient for peptide marker identification in foods.

The target peptides are ideal candidates for the quantification of gluten from *Triticum* sp. in processed products and this is the first step in providing a comprehensive set of markers to define gluten quantity in the context of protecting those with coeliac disease and supporting the enforcement of EU legislation. The strategy applied could be easily extended to other cereal sources, rye, barley and oats, yielding species specific and total gluten markers that could be used in the quantification of gluten.

An appropriate set of analytical targets may enable the harmonization of various detection and quantification techniques and the design of appropriate certified reference materials enabling standardised methods in the future.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jprot.2016.03.015.

Transparency Document

The Transparency document associated with this article can be found, in the online version.

References

[1] Codex Alimentarious Commission, Foods for Special Dietary Use for Persons Intolerant to Gluten., Rome, Italy, 2008.
[2] T.B. Osborne, The Vegetable Proteins, 2nd Edition, London, 1924.
[3] G. Mamone, G. Picariello, F. Addeo, P. Ferranti, Proteomic analysis in allergy and intolerance to wheat products, Expert Rev. Proteomics 8 (1) (2011) 95–115.
[4] C. Diaz-Amigo, B. Popping, Accuracy of ELISA detection methods for gluten and reference materials: a realistic assessment, J. Agric. Food Chem. 61 (24) (2013) 5681–5688.
[5] Husby S, Olsson C, Ivarsson A. Celiac disease and risk management of gluten. In: Charlotte Madsen C, Crevel R, Mills C, Taylor S, editors. Risk Management for Food Allergy, Chapter 7; 2014, p. 129–152.
[6] R. Ciccocioppo, A. Di Sabatino, G.R. Corazza, The immune recognition of gluten in coeliac disease, Clin. Exp. Immunol. 140 (3) (2005) 408–416.
[7] P. Ferranti, G. Mamone, G. Picariello, F. Addeo, Mass spectrometry analysis of gliadins in celiac disease, J. Mass Spectrom. 42 (12) (2007) 1531–1548.
[8] C. Catassi, A. Fasano, Celiac disease, Curr. Opin. Gastroenterol. 24 (6) (2008) 687–691.
[9] I.D. Bruins Slot, M.G.E.G. Bremer, I. van der Fels-Klerx, R.J. Hamer, Evaluating the performance of gluten ELISA test kits: the numbers do not tell the tale, Cereal Chem. 92 (5) (2015) 513–521.
[10] H. Wieser, P. Koehler, Is the calculation of the gluten content by multiplying the prolamin content by a factor of 2 valid? Eur. Food Res. Technol. 229 (1) (2009) 9–13.
[11] C. Diaz-Amigo, B. Popping, Gluten and gluten-free: issues and considerations of labeling regulations, detection methods, and assay validation, J. AOAC Int. 95 (2) (2012) 337–348.

[12] T.B. Koerner, M. Abbott, S.B. Godefroy, B. Popping, J.M. Yeung, C. Diaz-Amigo, J. Roberts, S.L. Taylor, J.L. Baumert, F. Ulberth, P. Wehling, P. Koehler, Validation procedures for quantitative gluten ELISA methods: AOAC allergen community guidance and best practices, J. AOAC Int. 96 (5) (2013) 1033–1040.
[13] J. Heick, M. Fischer, S. Kerbach, U. Tamm, B. Popping, Application of a liquid chromatography tandem mass spectrometry method for the simultaneous detection of seven allergenic foods in flour and bread and comparison of the method with commercially available ELISA test kits, J. AOAC Int. 94 (4) (2011) 1060–1068.
[14] N.R. Kitteringham, R.E. Jenkins, C.S. Lane, V.L. Elliott, B.K. Park, Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 877 (13) (2009) 1229–1239.
[15] M. Koeberl, D. Clarke, A.L. Lopata, Next generation of food allergen quantification using mass spectrometric systems, J. Proteome Res. 13 (8) (2014) 3499–3509.
[16] G. Prieto, K. Aloria, N. Osinalde, A. Fullaondo, J.M. Arizmendi, R. Matthiesen, PAnalyzer: a software tool for protein inference in shotgun proteomics, BMC Bioinformatics 13 (2012) 288.
[17] H. Wieser, Chemistry of gluten proteins, Food Microbiol. 24 (2) (2007) 115–119.
[18] R. Haraszi, H. Chassaigne, A. Maquet, F. Ulberth, Analytical methods for detection of gluten in food—method developments in support of food labeling legislation, J. AOAC Int. 94 (4) (2011) 1006–1025.
[19] H. Wieser, Comparative investigations of gluten proteins from different wheat species I. Qualitative and quantitative composition of gluten protein types, Eur. Food Res. Technol. 211 (4) (2000) 262–268.
[20] W. Seilmeier, I. Valdez, E. Mendez, H. Wieser, Comparative investigations of gluten proteins from different wheat species II. Characterization of v-gliadins, Eur. Food Res. Technol. 212 (2001) 355–363.
[21] F.M. DuPont, R. Chan, R. Lopez, W.H. Vensel, Sequential extraction and quantitative recovery of gliadins, glutenins, and other proteins from small samples of wheat flour, J. Agric. Food Chem. 53 (5) (2005) 1575–1584.
[22] J.H. Mejías, X. Lu, C. Osorio, J.L. Ullman, D. von Wettstein, S. Rustgi, Analysis of wheat prolamins, the causative agents of celiac sprue, using reversed phase high performance liquid chromatography (RP-HPLC) and matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS), Nutrients 6 (4) (2014) 1578–1597.
[23] J. Salplachta, M. Marchetti, J. Chmelík, G. Allmaier, A new approach in proteomics of wheat gluten: combining chymotrypsin cleavage and matrix-assisted laser desorption/ionization quadrupole ion trap reflectron tandem mass spectrometry, Rapid Commun. Mass Spectrom. 19 (18) (2005) 2725–2728.
[24] J.A. Sealey-Voyksner, C. Khosla, R.D. Voyksner, J.W. Jorgenson, Novel aspects of quantitation of immunogenic wheat gluten peptides by liquid chromatography-mass spectrometry/mass spectrometry, J. Chromatogr. A 1217 (25) (2010) 4167–4183.
[25] S.B. Altenbach, W.H. Vensel, F.M. Dupont, Analysis of expressed sequence tags from a single wheat cultivar facilitates interpretation of tandem mass spectrometry data and discrimination of gamma gliadin proteins that may play different functional roles in flour, BMC Plant Biol. 10 (2010) 7.
[26] F.M. Dupont, William H. Vensel, Charlene K. Tanaka, William J. Hurkman, Susan B. Altenbach, Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry, Proteome Sci. 9 (2011) 10.
[27] W.H. Vensel, F.M. Dupont, S. Sloane, S.B. Altenbach, Effect of cleavage enzyme, search algorithm and decoy database on mass spectrometric identification of wheat gluten proteins, Phytochemistry 72 (10) (2011) 1154–1161.
[28] Rombouts, B. Lagrain, M. Brunnbauer, J.A. Delcour, P. Koehler, Improved identification of wheat gluten proteins through alkylation of cysteine residues and peptide-based mass spectrometry, Sci. Rep. 3 (2013) 2279.
[29] R. Sturm, G. Sheynkman, C. Booth, L.M. Smith, J.A. Pedersen, L. Li, Absolute quantification of prion protein (90–231) using stable isotope-labeled chymotryptic peptide standards in a LC-MRM AQUA workflow, J. Am. Soc. Mass Spectrom. 23 (9) (2012) 1522–1533.
[30] K.L. Fiedler, S.C. McGrath, J.H. Callahan, M.M. Ross, Characterization of grain-specific peptide markers for the detection of gluten by mass spectrometry, J. Agric. Food Chem. 62 (25) (2014) 5835–5844.
[32] W.H. McDonald, et al., Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single-dimension LC- MS/MS, 2-phase MudPIT and 3-phase MudPIT, Int. J. Mass Spectrom. 219 (2002) 245–251.
[33] T. Glatter, C. Ludwig, E. Ahrné, R. Aebersold, A.J. Heck, A. Schmidt, Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion, J. Proteome Res. 11 (11) (2012) 5145–5156.
[34] L. Shan, S.W. Qiao, H. Arentz-Hansen, Ø. Molberg, G.M. Gray, L.M. Sollid, C. Khosla, Identification and analysis of multivalent proteolytically resistant peptides from gluten: implications for celiac sprue, J. Proteome Res. 4 (5) (2005) 1732–1741.
[35] L. Shan, Ø. Molberg, I. Parrot, F. Hausch, F. Filiz, G.M. Gray, L.M. Sollid, C. Khosla, Structural basis for gluten intolerance in celiac sprue, Science 297 (5590) (2002) 2275–2279.
[36] B. Prandi, A. Faccini, T. Tedeschi, A. Cammerata, D. Sgrulletta, M.G. D'Egidio, G. Galaverna, S. Sforza, Qualitative and quantitative determination of peptides related to celiac disease in mixtures derived from different methods of simulated gastrointestinal digestion of wheat products, Anal. Bioanal. Chem. 406 (19) (2014) 4765–4775.
[37] C.C. Hsia, O.D. Anderson, Isolation and characterization of wheat Ω-gliadin genes, Theor. Appl. Genet. 103 (1) (2001) 37–44.
[38] V. Muccilli, M. Lo Bianco, V. Cunsolo, R. Saletti, G. Gallo, S. Foti, High molecular weight glutenin subunits in some durum wheat cultivars investigated by means of mass spectrometric techniques, J. Agric. Food Chem. 59 (22) (2011) 12226–12237.
[39] International Wheat Genome Sequencing Consortium (IWGSC), A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome, Science 345 (6194) (2014) 1251788.
[40] P.E. Johnson, S. Baumgartner, T. Aldick, C. Bessant, V. Giosafatto, J. Heick, G. Mamone, G. O'Connor, R. Poms, B. Popping, A. Reuter, F. Ulberth, A. Watson, L. Monaci, E.N. Mills, Current perspectives and recommendations for the development of mass spectrometry methods for the determination of allergens in foods, J. AOAC Int. 94 (4) (2011) 1026–1033.
[41] A.I. Nesvizhskii, R. Aebersold, Interpretation of shotgun proteomic data: the protein inference problem, Mol. Cell. Proteomics 4 (10) (2005) 1419–1440.
[42] R. Aebersold, M. Mann, Mass spectrometry-based proteomics, Nature 422 (6928) (2003) 198–207.
[43] M.P. Washburn, D. Wolters, J.R. Yates 3rd., Large-scale analysis of the yeast proteome by multidimensional protein identification technology, Nat. Biotechnol. 19 (3) (2001) 242–247.
[44] M. Ribeiro, J.D. Nunes-Miranda, G. Branlard, J.M. Carrillo, M. Rodriguez-Quijano, G. Igrejas, One hundred years of grain omics: identifying the glutens that feed the world, J. Proteome Res. 12 (11) (2013) 4702–4716.
[45] M. Manfredi, M. Mattarozzi, M. Giannetto, M. Careri, Multiplex liquid chromatography-tandem mass spectrometry for the detection of wheat, oat, barley and rye prolamins towards the assessment of gluten-free product safety, Anal. Chim. Acta 895 (2015) 62–70.
[46] S. Lock, Gluten detection and speciation by liquid chromatography mass spectrometry (LC–MS/MS), Foods 3 (1) (2014) 13–29.
[47] M.L. Colgrave, H. Goswami, K. Byrne, M. Blundell, C.A. Howitt, G.J. Tanner, Proteomic profiling of 16 cereal grains and the application of targeted proteomics to detect wheat contamination, J. Proteome Res. 14 (2015) 2659–2668.
[48] H.C. Van de Broeck, J.H.G. Cordewener, M.A. Nessen, et al., Label free targeted detection and quantification of celiac disease immunogenic epitopes by mass spectrometry, J. Chromatol. A 1391 (2015) 60–71.
[49] L. Uvackova, L. Skultety, S. Bekesova, S. McClain, M. Hajduch, The MS(E)-proteomic analysis of gliadins and glutenins in wheat grain identifies and quantifies proteins associated with celiac disease and baker's asthma, J. Proteome 93 (2013) 65–73.
[50] B. Morón, M.T. Bethune, I. Comino, H. Manyani, M. Ferragud, M.C. López, A. Cebolla, C. Khosla, C. Sousa, Toward the assessment of food toxicity for celiac patients: characterization of monoclonal antibodies to a main immunogenic gluten peptide, PLoS One 3 (5) (2008), e2294.
[51] C. Mitea, Y. Kooy-Winkelaar, P. van Veelen, A. de Ru, J.W. Drijfhout, F. Koning, L. Dekking, Fine specificity of monoclonal antibodies against celiac disease-inducing peptides in the gluteome, Am. J. Clin. Nutr. 88 (4) (2008) 1057–1066.
[52] I. Valdés, E. García, M. Llorente, E. Méndez, Innovative approach to low-level gluten determination in foods using a novel sandwich enzyme-linked immunosorbent assay protocol, Eur. J. Gastroenterol. Hepatol. 15 (5) (2003) 465–474.
[53] G.J. Tanner, M.L. Colgrave, M.J. Blundell, H.P. Goswami, C.A. Howitt, Measuring hordein (gluten) in beer- a comparison of ELISA and mass spectrometry, PLoS One 8 (2) (2013), e56452.
[54] P. Cressey, P. Grounds, S. Jones, E. Ashmore, D. Saunders, Gluten residues in gluten-free foods in Christchurch, New Zealand: comparison of LC-MS and ELISA methods, Qual. Assur. Saf. Crop. 5 (3) (2013) 207–2013.