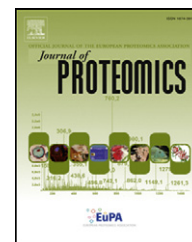


Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/jprot

The MS^E-proteomic analysis of gliadins and glutenins in wheat grain identifies and quantifies proteins associated with celiac disease and baker's asthma[☆]

Lubica Uvackova^a, Ludovit Skultety^{b,c}, Slavka Bekesova^b,
Scott McClain^d, Martin Hajduch^{a,*}

^aDepartment of Reproduction and Developmental Biology, Institute of Plant Genetics and Biotechnology, Slovak Academy of Sciences, Nitra, Slovakia

^bInstitute of Virology, Slovak Academy of Sciences, Bratislava, Slovakia

^cCenter for Molecular Medicine, Slovak Academy of Sciences, Bratislava, Slovakia

^dSyngenta Crop Protection, LLC, Research Triangle Park, NC, USA

ARTICLE INFO

Keywords:

Gliadins
Glutenins
Proteomics
MS^E
Quantification
Wheat grain

ABSTRACT

Precise content of gliadin (Glia) and glutenin (Glu) proteins in wheat grain are largely unknown despite their association with celiac disease, various allergies, and physical processing properties of wheat. Developing methods to quantitatively measure clinically relevant proteins could support advancement in understanding exposure thresholds and clinical study design. The aim of this study was to use a data-independent mass spectrometry (MS^E) approach for quantifying gliadin and glutenin proteins in wheat grain. The biologically replicated analysis yielded concentrations for 34 gliadin and 22 glutenin proteins. The primary focus of this survey was on measuring celiac disease proteins and baker's asthma associated proteins along with the proteins associated with viscoelastic properties of wheat flour and grain texture. The technical coefficients of variation ranged from 0.12 to 1.39 and indicate that MS^E proteomics is a reproducible quantitative method for the determination of gliadin and glutenin content in the highly complex matrix of protein extracts from wheat grain. This article is part of a Special Issue entitled: Translational Plant Proteomics.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Wheat (*Triticum aestivum* L.) is a widely used cereal for human consumption. Wheat grain has some unique properties when processed into wheat flour that depend on protein composition [1]. Based on solubility, wheat grain proteins can be divided into the following: i) water soluble albumin, ii) salt

soluble globulin, and iii) alcohol soluble gluten proteins [2]. Gluten proteins represent up to 80% of wheat proteins [3] and are divided into gliadins (Glia) and glutenins (Glu) based on electrophoretic mobility [4,5]. Glia proteins are mostly monomeric proteins with molecular weight 28–55 kDa and can be divided either based on electrophoretic mobility into α/β -, γ -, and ω -gliadins or based on amino-acid sequence into ω 5-

Abbreviations: Glia, protein fraction containing mostly gliadins; Glu, protein fraction containing mostly glutenins; MSE, data-independent mass spectrometry; emPAI, exponentially modified protein abundance index; GFP, Glu-1-fibrinopeptide B; CV, coefficient of variations; HMW GS, high molecular weight glutenin subunits; LMW GS, low molecular weight glutenin subunits.

[☆] This article is part of a Special Issue entitled: Translational Plant Proteomics.

* Corresponding author. Tel.: +421 37 6943346; fax: +421 37 7336660.

E-mail address: hajduch@savba.sk (M. Hajduch).

1874-3919/\$ – see front matter © 2012 Elsevier B.V. All rights reserved.

<http://dx.doi.org/10.1016/j.jprot.2012.12.011>

Please cite this article as: Uvackova L, et al, The MS^E-proteomic analysis of gliadins and glutenins in wheat grain identifies and quantifies proteins associated with celiac disease..., J Prot (2013), <http://dx.doi.org/10.1016/j.jprot.2012.12.011>

ω 1,2-, α/β - and γ -gliadins [6,7]. Glu proteins can be sorted according to their content of high (HMW 70–90 kDa) and low (LMW 20–45 kDa) molecular weight (MW) glutenin subunits [8,9]. In addition to Glia and Glu, wheat flour contains a number of minor storage proteins, such as “avenin-like” proteins that are similar to oat storage proteins called avenins [10].

In wheat, Glia and Glu proteins are often associated with celiac disease and various allergies [for review see 11–13]. The development of methods for quantitative determination of clinically relevant Glia and Glu proteins is important in order to understand the exposure thresholds and to support clinical allergy study designs. Presently, quantitative methods for the determination of clinically relevant Glia and Glu proteins are mostly based on immunology techniques [14–16]. However, techniques based on mass spectrometry are likely to provide more robust characterization of Glia and Glu proteins [16,3]. Among these, a novel data-independent MS^E acquisition method takes advantage of a data collection approach that focuses on maximizing the amount of data collected for a proteome by rapid switching between low and elevated collision energies to assess multiple target proteins simultaneously [17]. Evaluation of the data output during MS^E analysis of tomato root microsomes showed that MS^E is a reproducible and reliable method for protein identification [18]. In combination with data-dependent MS/MS approaches, MS^E proteomics has also provided reliable identification of cytosolic ribosome proteins isolated from the model plant, *Arabidopsis thaliana* [19]. In this work by Hummel et al. [19], MS^E was used in combination with peak alignment to measure the relative abundance for several ribosomal proteins in a multiplex (proteomic) approach. Previous reports using non-plant systems have shown that MS^E is also capable of providing absolute quantitative analysis by analyzing the signal response of a known quantity of internal standard that has been spiked into the sample [20,21].

The aim of this study was to develop a reliable and robust method based on MS^E that will provide reproducible, quantitative measures of the wheat grain proteins, Glia and Glu, some of which are relevant for wheat food allergy.

2. Material and methods

2.1. Plant material

Bread wheat cultivar “Viginta” was provided by SELEKT Ltd., Bučany. To avoid seed-to-seed variability, 1 kg of wheat grain was divided into 4 equal parts. Each part was then divided into five 50 g portions that were each milled by electric grinder (50 g of grains three times for 5 s to avoid heat). Milled

grain portions were stored at –20 °C until use. Each portion of ground grain was considered a replicate sample.

2.2. Extraction of alcohol soluble proteins from wheat grain

Gliadin (Glia) and glutenin (Glu) proteins were extracted from 200 mg of a milled grain sample according to Broeck [22]. Glia proteins were obtained after extraction with 1 ml of 50% (v/v) aqueous iso-propanol by continuous mixing (Thermomixer comfort, Eppendorf) at 1000 rpm for 30 min at room temperature, followed by centrifugation at 2500×g for 15 min at room temperature (RT). The procedure was repeated one time. To obtain Glu proteins, the procedure above was repeated and the pellet was further extracted with 1 ml of 50% (v/v) aqueous iso-propanol, 50 mM Tris-HCl (pH 7.5) containing 1% (w/v) DTT for 30 min at 60 °C with mixing every 5 to 10 min, followed by centrifugation at 10,000×g for 10 min at RT (Fig. 1).

2.3. Protein “in-solution” digestion in alcohol

Because alcohol interferes with total protein determination according to the Bradford method [23], wheat grain extracts produced by the method described in Section 2.2 were divided into two equal aliquots. The first aliquot was then used to determine the protein concentration and second for further analysis by mass spectrometry. Protein from the first aliquot were precipitated with 5 volumes of 0.1 M ammonium acetate in 100% methanol and washed with 0.1 M ice cold ammonium acetate in 100% methanol twice. After centrifugation at 5000 g at 4 °C for 10 min, the liquid phase was removed. Precipitated proteins were reconstituted in 100 µl of solubilization solution (8 M urea, 2 M thiourea, 2%CHAPS, 50 mM DTT). Total protein concentration was then measured according to Bradford [23].

One hundred micrograms of the total protein preparation (second aliquot) was reduced with DTT (final concentration 10 mM) and alkylated with IAA (final concentration 40 mM), gently mixed, and incubated at RT for 45 min in the dark. Trypsin (Promega, Madison, WI, USA) in 1 mM NH₄HCO₃ was added to protein solution (1:25), gently mixed, and incubated overnight at 37 °C (16–20 h). The reaction was stopped with formic acid (final concentration 5%). The resulting tryptic digest was concentrated and desalted using Peptide Cleanup C18 spin columns (Agilent Technologies) according to the manufacturer’s instructions and stored at –80 °C until mass spectrometry (MS) analyses.

2.4. Mass spectrometry

Prior to MS analysis, the trypsin digested extracts were concentrated under vacuum to 20 µl using Concentrator plus

Fig. 1 – The workflow. Gliadins (Glia) and glutenins (Glu) were extracted from milled wheat grain. Gliadins and glutenins were then digested with trypsin, peptides were spiked with internal standard (predigested bovine hemoglobin, 186002327), separated by nanoAcquity ultraperformance liquid chromatography (UPLC), and analyzed by MS^E. For protein identification alternate scans between low and elevated energies were used to generate fragmentation data without precursor selection. For protein quantification, MS signals for the three most intense tryptic peptides were calculated for internal standard, bovine hemoglobin. Then, the average MS signal response from the standard was used to determine a universal signal response factor (USRF, counts/mol of protein). The USRF was then used to determine the concentration of the target proteins. This approach resulted into quantification of 34 Glia and 22 Glu proteins in at least biological triplicate.

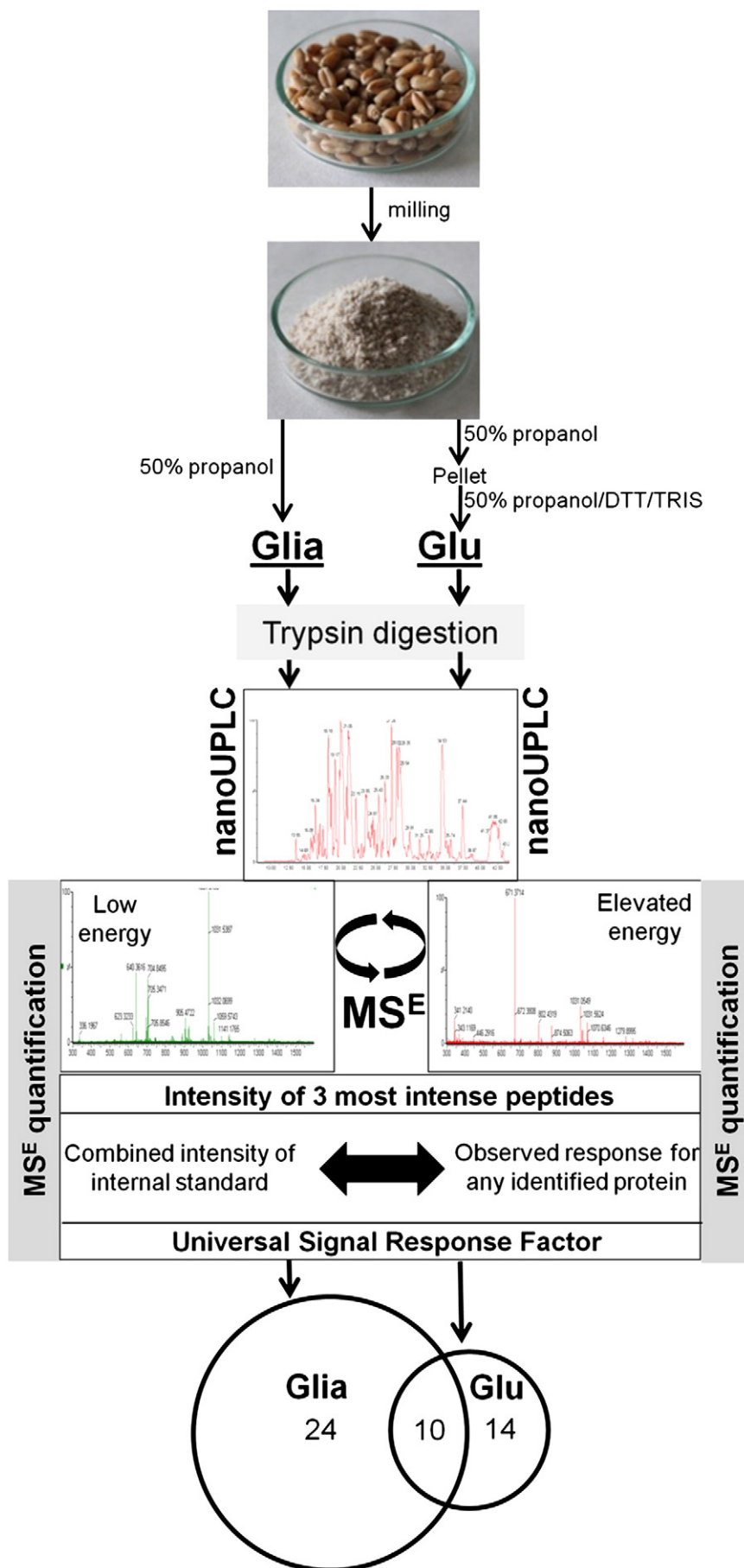


Table 1 – List of gliadin (Glia) and glutenin (Glu) proteins identified in wheat grains. The table contain UniProt accession number; protein name; MW/pI, molecular weight (kDa)/pH; organism; Glia amount (ng/μg), average concentration in ng of protein per μg of sample, STDEV, standard deviation.

Accession	Protein name	MW/pI	Genus/species	Glia		Glu	
				Amount (ng/μg)	STDEV	Amount (ng/μg)	STDEV
Sugar metabolism							
P30271	Beta amylase	24.3/4.9	Secale cereale	0.09	0.05		
Q4VM11	Beta amylase	57.1/5.2	Hordeum vulgare var distichum	3.07	1.78		
Q7X9M2	Beta amylase	30.9/8.5	Triticum aestivum	6.45	6.64		
Proteolysis							
B9VRI3	Alpha amylase inhibitor CM16 subunit	15.6/5.1	T. aestivum	0.93	0.52		
C0LF30	Serpin 1	43.1/5.3	T. aestivum	7.51	10.08		
C4P615	Monomeric alpha amylase inhibitor	16.6/7.4	Triticum turgidum subsp dicoccoides	0.62	0.20		
P16347	Endogenous alpha amylase subtilisin inhibitor	196/6.8	T. aestivum	0.99	1.23		
P17314	Alpha amylase trypsin inhibitor CM3	18.2/7.3	T. aestivum	21.05	18.81	11.31	3.05
Q41540	CM 17 protein	16.0/4.9	T. aestivum			8.44	6.38
Seed storage							
A7LHB3	Alpha gliadin	35.4/8.0	T. aestivum	1.56	0.26		
A7LHC8	Alpha gliadin	35.4/8.0	Thinopyrum ponticum × T. aestivum	6.60	4.62		
D2X6D5	Alpha gliadin	34.7/8.5	T. turgidum subsp durum			15.89	10.51
D2X6D8	Alpha gliadin	33.0/8.8	T. turgidum subsp durum	3.70	2.73		
Q2A780	Putative avenin like “a”	18.8/7.6	A. markgrafii	1.06	0.58	3.82	3.59
Q2A784	Putative avenin like “a”	18.4/7.9	T. aestivum	3.34	1.54	11.85	7.83
Q2A783	Putative avenin like “b”	32.7/7.6	T. aestivum	3.20	2.95	22.96	13.76
D2KFG9	Gliadin avenin like seed protein	22.4/6.2	T. aestivum			14.82	3.52
D2KFH2	Gliadin avenin like seed protein	11.3/7.6	T. aestivum			9.36	5.62
B0L966	D hordein	91.8/6.6	Hordeum chilense			223.44	62.14
B6DQD1	Gamma gliadin	32.7/7.9	T. aestivum × Lophopyrum elongatum	20.29	12.80		
B6DQD4	Gamma gliadin	31.6/8.3	T. aestivum × L. elongatum	33.33	4.07		
B6DQE1	Gamma gliadin	31.8/8.1	T. aestivum × L. elongatum	1.54	1.05		
B6UKJ9	Gamma gliadin	29.2/7.5	T. turgidum subsp dicoccoides	4.17	3.50	5.23	6.26
B6UKL5	Gamma gliadin	29.7/8.5	T. turgidum	12.02	10.31		
B6UKN7	Gamma gliadin	35.7/8.0	T. aestivum	3.87	3.30		
B6UKU0	Gamma gliadin	26.7/7.0	Aegilops searsii	21.01	23.91		
B8XU42	Gamma gliadin	32.5/7.5	A. searsii	10.14	1.59		
B8XU48	Gamma gliadin	18.1/8.5	T. monococcum	11.83	11.58	11.87	17.71
B8XU49	Gamma gliadin	21.7/7.9	T. monococcum	13.27	9.09		
Q1W676	Gamma gliadin	14.3/8.7	T. aestivum	3.86	2.45	19.87	12.60
Q8H738	Gamma gliadin	16.3/7.8	T. aestivum	13.79	4.22		
Q0Q5D4	Globulin 1	24.5/7.7	T. aestivum			9.16	5.38
Q0Q5D9	Globulin 1	24.5/8.1	T. aestivum			7.26	6.58
A9YSK3	High molecular weight glutenin subunit	69.7/8.1	T. aestivum	13.84	3.52		
Q7XAH4	High molecular weight glutenin subunit	12.4/6.9	T. turgidum subsp durum			2.53	2.55
A8IEC6	Low molecular weight glutenin subunit	33.8/8.9	Thinopyrum ponticum × T. aestivum	1.51	0.42		
D5FPF1	Low molecular weight glutenin subunit	42.5/9.1	Ampelocera longissima			41.99	18.16
Q1ZZT4	Low molecular weight glutenin subunit	31.8/8.1	T. aestivum	1.75	1.24	11.80	16.70

Table 1 (continued)

Accession	Protein name	MW/pI	Genus/species	Glia		Glu	
				Amount (ng/μg)	STDEV	Amount (ng/μg)	STDEV
B8XU59	Low molecular weight glutenin	31.2/7.9	<i>T. monococcum</i>			17.19	10.29
Q8W3V1	Low molecular weight glutenin subunit	23.9/8.7	<i>T. aestivum</i>	2.44	1.09		
B2BZC7	LMW m glutenin subunit	32.7/8.7	<i>T. aestivum</i>	2.59	0.68	18.92	14.82
B5ANT3	LMW glutenin	31.9/7.8	<i>T. aestivum</i>			7.57	3.97
Q00M56	LMW GS P 31	40.1/8.1	<i>T. aestivum</i>	1.40	1.45		
B2BZC7	LMW m glutenin subunit	32.7/8.7	<i>T. aestivum</i>	2.59	0.68	18.92	14.82
B5ANT3	LMW glutenin	31.9/7.8	<i>T. aestivum</i>			7.57	3.97
Q8H737	Lmw gs	22.9/8.4	<i>T. aestivum</i>	4.35	2.75	8.86	9.62
Q7X9M4	Putative uncharacterized protein	25.1/7.9	<i>T. aestivum</i>			37.93	14.89
Disease/defense							
Q41554	15 kDa grain softness protein	6.0/9.0	<i>T. aestivum</i>	7.67	10.68		

(Eppendorf, Germany). Final volume was then adjusted to 100 μl by adding 5% acetonitrile in 0.1% formic acid. Each sample was spiked with predigested bovine hemoglobin (prepared by digesting of HBA P01966, HBB P02081) internal standard (Waters 186002327, UK) at a level of 1 pmol per 5 μL injection (1.5 μg of wheat proteins). For sample analysis, 1.5 μg protein aliquots were analyzed by nanoAcquity ultraperformance liquid chromatography (UPLC) coupled to Premier quadrupole time-of-flight (QTOF) mass spectrometer (Waters, UK). The peptide mixture was injected onto a reverse-phase column (nanoAcquity UPLC column BEH 130 C18, 75 μm × 150 mm, 1.7 μm particle size) and an acetonitrile gradient (10–45% B in 40 min; A=water with 0.1% formic acid, B=acetonitrile containing 0.1% formic acid) at a flow rate of 350 nL/min was employed to elute the peptides into the Q-TOF. The column was connected to the PicoTip emitters (New Objective, USA) mounted into the nanospray source of the Q-TOF Premier. A nano-electrospray voltage of 3.4 kV was applied, with the source temperature set to 70 °C. The spectral acquisition scan rate was 1 s with a 0.05 s interscan delay. A novel multiplex MS^E approach, in which MS data are collected in an alternating, low energy (MS) and elevated energy (MS^E) mode, was used for protein identification [24,25]. This method generates both precursor and product ions in a single analytical run by abandoning the selection of a precursor ion for individual fragmentation and instead fragmenting everything. The quantification of proteins using MS^E technology is based on the observation that exponentially modified protein abundance index (emPAI) is directly proportional to the protein content in a protein mixture [26]. Thus, the combined intensity of the charged ions for the three most abundant tryptic peptides of internal standard was compared with the observed response for identified protein in a complex mixture. The average MS signal response from bovine hemoglobin was used to determine the universal signal response factor (USRF, counts/mol of protein). This information was then used to determine the concentrations for each of the target proteins by dividing the MS by USRF. In the low energy MS mode, data were collected at constant collision energy of 3 eV. In elevated energy MS^E mode, the collision energy was ramped

from 20 eV to 38 eV during each integration. During data acquisition the quadrupole analyzer was not mass selective but operated in the radio-frequency only mode. Thus, all ions were passed to the TOF analyzer. Glu-1-fibrinopeptide B (GFP) at a concentration of 1 pmol/mL was infused via the source at a flow rate of 500 nL/min and sampled every 30 s as the external mass calibrant.

2.5. Data processing and database searching

The MS^E data were processed using the ProteinLynx Global Server v. 2.4 (PLGS 2.4; Waters, UK) that provides background-subtraction, smoothing, centroiding and deisotoping. All data were then lock-spray calibrated against GFP using data collected from the reference line during acquisition and charge-state reduced to produce a precursor monoisotopic mass lists as well as associated product ion mass lists. Each processed file was searched against the non-redundant *Triticeae* plant protein UniProt database downloaded on March 1, 2011 (11,764 entries) with addition of internal standard HBA_BOVIN Hemoglobin subunit alpha sequence (Waters: 186002327; NCBI: P01966) using the search algorithm within the PLGS 2.4. The initial correlation of a precursor (MS) and possible fragment ions (MS/MS) was achieved by means of time alignment. The elution profile of precursor masses from the liquid chromatography component was used to deconvolute the mass of fragments from many precursor ions. Default search parameters were applied including the “automatic” setting for mass accuracy (10 ppm for precursor ions and 15 ppm for product ions), a minimum of 1 peptide match per protein, a minimum of 3 consecutive product ion matches per protein, and a minimum of 7 total product ion matches per protein. The maximum false positive rate (FPR) against the randomized forward database was set at 4%, and the protein quantification functionality was enabled using the internal standard (1 pmol of HBA_BOVIN Hemoglobin subunit alpha). Only 1 missed tryptic cleavage site was allowed during the search. A fixed carbamidomethyl-Cys modification was used, in addition to the following variable modifications: deamidation of Asn and Gln; oxidation of Met; and dehydration of Ser

and Thr. In order to determine protein quantities the combined intensity of the multiply charged ions for the three most abundant tryptic peptides of a quantitatively added internal standard was compared with the observed response for any identified protein in a complex mixture.

3. Results and discussion

3.1. MS^E proteomics provided reliable identification and quantification of Glia and Glu proteins in wheat grain

In order to isolate Glia and Glu proteins from wheat grain prior to MS^E, a well established two-step propanol extraction procedure was followed [22] with slight modification. In this study the ground wheat grain (flour) was not sieved (Fig. 1). For the purposes of surveying for the broadest inclusion of proteins, whole-wheat flour was considered preferable because sieving can remove larger particles that can include proteins from the seed coat. The original method using sieving extracted majority of Glia (1.94±0.05 mg/ml) and Glu (0.61±0.01 mg/ml) proteins as shown on Fig. 2 of Broeck [22]. Use of whole grain wheat flour in present study decreased efficiency of Glia extraction to 0.23±0.07 mg/ml and the efficiency of Glu extraction was not effected (0.45±0.18 mg/ml).

After protein “in-solution” digestion with trypsin, peptides were separated by nanoUPLC and subjected to analysis by data-independent MS^E technology for protein identification and quantification that is based on alternation of low and elevated energy scans to generate fragmentation data without precursor ion selection [19]. For the purpose of quantification, 1 pmol of trypsin digested bovine hemoglobin (P01966) was spiked to the wheat grain extract sample prior the MS^E analysis as internal standard (Fig. 1). Then, acquired MS^E data were processed by PLGS 2.4 (Waters, UK).

A qualification of the MS^E approach was performed in order to verify that concentration values for each target protein were reproducible. Proteins needed to be identified and quantified in triplicate; triplicates were defined as separate experiments of protein isolation from 200 g of milled wheat grain followed by MS^E quantification. This approach resulted in quantifying 34 proteins in the Glia extract and 22 in Glu extract preparations (Supplementary table). It was revealed that 11 of these proteins were quantified in both extraction fractions (Table 1). For instance, α -amylase inhibitor with accession number P17314 (Table 1). These data are in agreement with previously reported studies that indicate that proteins can be co-purified in different wheat grain protein fractions where, for instance, α -amylase inhibitor subunits were identified not only in Glia but also in the Glu fraction [27].

All quantified proteins were classified into functional groups according to Bevan [28] (Fig. 2). In total, 25 Glia and 20 Glu proteins were associated with seed storage, which was the most abundant functional category (Table 1). The second most abundant functional category consisted of proteins associated with proteolysis (groups of 5 Glia and 2 Glu; Table 1). Additionally, in the Glia extract, we quantified three beta amylases that are associated with sugar metabolism and a 15 kDa grain softness protein that is associated with disease and defense (Table 1).

In order to determine the reproducibility of quantifying individual proteins in replicate extracts, the coefficient of variations (CV) were calculated for each quantified protein (Supplementary table). Coefficients of variation ranged from 0.12 to 1.39 (Fig. 3, Supplementary table) indicating a good degree of reproducibility. More precise CV values (0.1–0.3) were obtained with MS^E quantitative proteomics only for mixture of four proteins [21]. Taking into account the highly complex matrix of the wheat grain proteins and the fact that two types of extractions are required to support detection of Glia and Glu proteins, the MS^E proteomics provided reliable quantification of proteins from wheat grain.

3.2. High concentrations of celiac disease proteins and low concentrations of baker's asthma proteins were detected in wheat grain

Wheat grain storage proteins mainly consist of gluten, which are glutamine and proline-rich proteins [29]. Gluten induces celiac disease in susceptible individuals [30]. The main components of wheat gluten that are responsible for the disease are gliadin proteins [27,31,32,33–35]. Taxonomically, the wheat gluten protein family members, γ -gliadins [36], are the most conserved taxonomically and structurally and they contain several celiac disease epitopes [37]. In our study, γ -gliadins were abundant in terms of the number of proteins in the Glia fraction with 12 quantified sequences (Table 1). The most abundant γ -gliadin sequence was B6DQD4 with 33 (±4) ng/μg of Glia extract. The distribution of the γ -gliadins that we could detect and quantify shows that 80% were extracted from the Glia fraction (Table 1). In the Glu protein extract, only three γ -gliadin sequences were quantified (B6UKJ9, B8XU48, and Q1W676) (Table 1) of which the most abundant was Q1W676 with 20 (±13) ng/μg of Glu extract (Table 1).

The protein family that includes the α -amylase trypsin inhibitors and chloroform-methanol-soluble proteins (CM proteins) in wheat grain are associated with baker's asthma and associated respiratory allergy [38]. In this study, we identified concentrations of 4 sequences of α -amylase trypsin inhibitors in the Glia fraction of which the most abundant was P17314 (21±19 ng/μg of Glia extract). This α -amylase trypsin inhibitor (P17314) was also quantified in the Glu fraction (11±3 ng/μg of Glu extract) (Table 1, Supplemental table). One CM 17 protein was quantified in the Glu fraction (Table 1).

Serpins and β -amylases have been reported as potential allergens [39–42]. This study identified one sequence (COLF30) of Serpin 1 and three sequences of β -amylases (Q7X9M2, Q4VM11, P30271) that were extracted exclusively from the Glia fraction (Table 1).

3.3. Proteins associated with viscoelastic properties of wheat are highly abundant in the Glu fraction

Another group of proteins identified in this study were associated with mechanical handling and processing characteristics of wheat grain. Low and high molecular weight glutenin subunits (HMW GS, LMW GS, respectively) play a major role in the determination of viscoelastic properties of wheat flour [43]. In total, our study quantified 2 sequences of HMW GS and 11 of LMW GS (Table 1). This is not surprising because LMW GS are

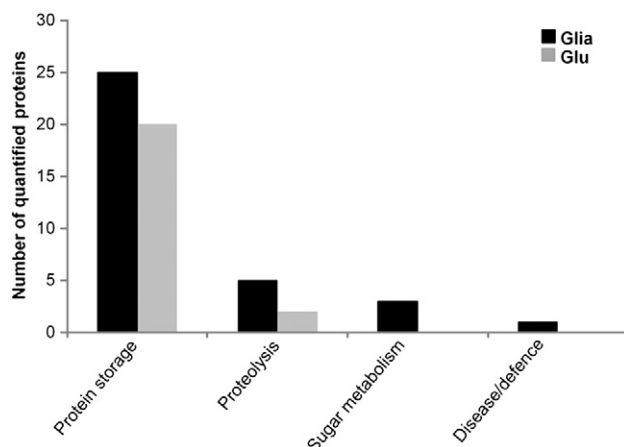


Fig. 2 – Functional classification of 34 gliadin (Glia) and 22 glutenin (Glu) proteins quantified in wheat grain.

derived from many more genes than HMW GS [43]. For the group of HMW GS proteins, sequence A9YSK3 was quantified in the Glia fraction (14 ± 3 ng/ μ g of Glia extract) and Q7XAH4 in the Glu fraction (2 ± 2 ng/ μ g of Glu extract). Out of 11 LMW GC proteins, 4 were quantified in both fractions, 3 only in the Glia fraction and 4 only in the Glu fraction (Table 1). Out of these, the most abundant was D5FPF1 which was detected in the Glu fraction at a concentration of 42 ± 18 ng/ μ g (Table 1). A majority of glutenin subunits were quantified in the Glu fraction which confirms a previous report that GS proteins are extracted with glutenins [44]. Additionally, in the Glu fraction, we identified 223 (± 62) ng of sequence B0L966 which belongs to the structural group of proteins that includes D hordein, a prolamin storage protein of barley endosperms (Table 1) that displays high degree of sequence homology to the HMW GS proteins [45].

Avenin-like proteins “a” and “b” were previously identified in wheat grain using EST libraries [10]. We identified two sequences of “a” type (Q2A784, Q2A780) and one sequence of “b” type (Q2A783) in the Glia and Glu fractions. The abundance of avenin-like proteins were higher in the Glu fraction compared with the Glia fraction (Table 1). Additionally, two sequences of Gliadin related avenin-like seed proteins (D2KFH2, D2KFG9) were identified in the Glu fraction (Table 1). Avenin-like proteins in wheat

might be associated with viscoelastic properties of wheat flour [43] because of their similarities to LMW gliadin sequences [10]. They are also likely relevant to understanding the different physical properties in various wheat varieties and how it affects their specific use in processed foods. One last protein associated with the physical properties of wheat grain is the 15 kDa grain softness protein [46]; this protein has been previously referred to as friabilin [47]. This 15 kDa protein is a mixture of puroindoline-like polypeptides and is associated with wheat grain texture [48]; a concentration of $8 (\pm 11)$ ng of sequence Q41554 was identified and measured in the Glia fraction (Table 1).

4. Concluding remarks

This study evaluated a MS^E-based proteomic approach for wheat protein identification and quantification in extracts prepared from whole wheat grain. The study design was an effort to develop reliable quantitative MS-based methods using cutting-edge MS techniques for the determination of protein content in highly complex matrices such as wheat grain and to pair this with characterizations of how best to extract and survey a wide variety of soluble proteins, particularly those proteins that can be important for wheat allergy and food processing. The primary aims of standardizing a gel-free and label-free proteomic approach for use with wheat grain were achieved with this work and provide a platform for further studies that rely on quantitative measures of wheat protein targets. The next step in this research is expected to be the use of a targeted MS^E-based proteomic approach for quantitative analysis of clinically relevant wheat grain allergens to support a comparative analysis of different wheat varieties. An understanding of the natural variability in protein expression will be critical to futures studies that attempt to understand the clinical importance of exposure to wheat proteins.

Acknowledgments

This research was supported by Syngenta Biotechnology Inc.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jprot.2012.12.011>.

REFERENCES

- [1] Johansson E, Prieto-Linde ML, Jönsson JÖ. Effects of wheat cultivar and nitrogen application on storage protein composition and breadmaking quality. *Cereal Chem* 2001;78:19–25.
- [2] Osborne TB. The vegetable proteins. London: Longmans, Green & Co.; 1924.
- [3] Ferranti P, Mamone G, Picariello G, Addeo F. Mass spectrometry analysis of gliadins in celiac disease. *J Mass Spectrom* 2007;42:1531–48.
- [4] Payne PI, Holt LM, Jarvis MG, Jackson EA. Two-dimensional fractionation of the endosperm proteins of bread wheat (*Triticum aestivum*): biochemical and genetic studies. *Cereal Chem* 1985;62:319–26.

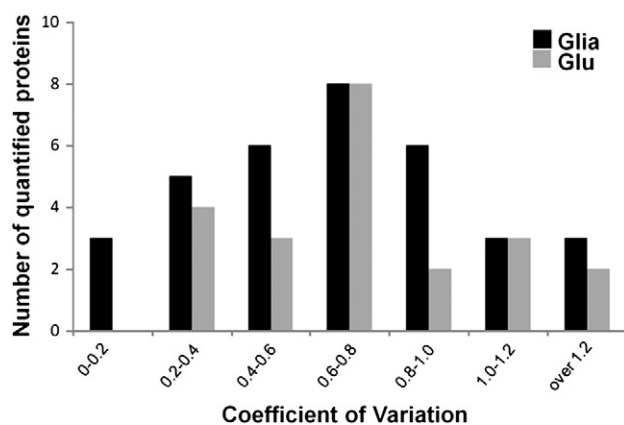


Fig. 3 – The coefficient of variation (CV), ratio of the standard deviation to the average, between protein quantifications.

- [5] Jacobsen S, Grove H, Nedenskov Jensen K, Sorensen HA, Jessen F, Hollung K, et al. Multivariate analysis of 2-DE protein patterns — practical approaches. *Electrophoresis* 2007;28:1289–99.
- [6] Wieser H. Chemistry of gluten proteins. *Food Microbiol* 2007;24:115–9.
- [7] Wieser H. Relation between gliadin structure and coeliac toxicity. *Acta Paediatr* 1996;412(85 Suppl.):3–9.
- [8] Tatham AS, Shewry PR, Belton PS. Structural studies of cereal prolamins, including wheat gluten. In: Pomeranz Y, editor. *Advances in Cereal Science and Technology*, vol. 10. St Paul: AACC; 1990.
- [9] Jackson EA, Holt LM, Payne PI. Characterisation of high molecular-weight gliadin and low-molecular-weight glutenin subunits of wheat endosperm by two-dimensional electrophoresis and the chromosomal localization of their controlling genes. *Theor Appl Genet* 1983;66:29.
- [10] Kan Y, Wan Y, Beaudoin F, Leader DJ, Edwards K, Poole R, et al. Transcriptome analysis reveals differentially expressed storage protein transcripts in seeds of *Aegilops* and wheat. *J Cereal Sci* 2006;44:75–85.
- [11] Rashtak S, Murray JA. Review article: coeliac disease, new approaches to therapy. *Aliment Pharmacol Ther* 2012;35:768–81.
- [12] Pietzak M. Celiac disease, wheat allergy, and gluten sensitivity: when gluten free is not a fad. *JPEN J Parenter Enteral Nutr* 2012;36(1 Suppl.):68S–75S.
- [13] Sapone A, Bai JC, Ciacci C, Dolinsek J, Green PH, Hadjivassiliou M, et al. Spectrum of gluten-related disorders: consensus on new nomenclature and classification. *BMC Med* 2012;10:13.
- [14] Allred LK, Ritter BW. Recognition of gliadin and glutenin fractions in four commercial gluten assays. *J AOAC Int* 2010;93:190–6.
- [15] Sharma GM. Immunoreactivity and detection of wheat proteins by commercial ELISA kits. *J AOAC Int* 2012;95:364–71.
- [16] Simonato B, Mainente F, Tolini S, Pasini G. Immunochemical and mass spectrometry detection of residual proteins in gluten fined red wine. *J Agric Food Chem* 2011;59(7):3101–10.
- [17] Silva JC, Gorenstein MV, Li GZ, Vissers JPC, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics* 2006;5:144–56.
- [18] Mbeunkui F, Goshe MB. Investigation of solubilization and digestion methods for microsomal membrane proteome analysis using data-independent LC-MSE. *Proteomics* 2011;11:898–911.
- [19] Hummel M, Cordewener JH, de Groot JC, Smeekens S, America AH, Hanson J. Dynamic protein composition of *Arabidopsis thaliana* cytosolic ribosomes in response to sucrose feeding as revealed by label free MSE proteomics. *Proteomics* 2012;12:1024–38.
- [20] Blackburn K, Cheng FY, Williamson JD, Goshe MB. Data-independent liquid chromatography/mass spectrometry (LC/MS(E)) detection and quantification of the secreted *Apium graveolens* pathogen defense protein mannitol dehydrogenase. *Rapid Commun Mass Spectrom* 2010;24:1009–16.
- [21] Levin Y, Hradetzky E, Bahn S. Quantification of proteins using data-independent analysis (MSE) in simple and complex samples: a systematic evaluation. *Proteomics* 2011;11:3273–87.
- [22] van den Broeck HC, America AH, Smulders MJ, Bosch D, Hamer RJ, Gilissen LJ, et al. A modified extraction protocol enables detection and quantification of celiac disease-related gluten proteins from wheat. *J Chromatogr B Analyt Technol Biomed Life Sci* 2009;877:975–82.
- [23] Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976;72:248–54.
- [24] Plumb RS, Johnson KA, Rainville P, Smith BW, Wilson ID, Castro-Perez JM, et al. UPLC/MS(E); a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Commun Mass Spectrom* 2006;20:1989–94.
- [25] Li GZ, Vissers JP, Silva JC, Golick D, Gorenstein MV, Geromanos SJ. Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. *Proteomics* 2009;9:1696–719.
- [26] Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, et al. Exponentially modified protein abundance index (em-PAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 2005;4:1265–72.
- [27] Pastorello EA, Farioli L, Conti A, Pravettoni V, Bonomi S, Iametti S, et al. Wheat IgE-mediated food allergy in European patients: α -amylase inhibitors, lipid transfer proteins and low-molecular-weight glutenins. *Int Arch Allergy Immunol* 2007;144:10–22.
- [28] Bevan M, Bancroft I, Bent E, Love K, Goodman H, Dean C, et al. Analysis of 19 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature* 1998;391:485–8.
- [29] Shewry PR, Tatham AS, Lazzeri P. Biotechnology of wheat quality. *J Sci Food Agric* 1997;73:397–406.
- [30] Shan L, Molberg Ø, Parrot I, Hausch F, Filiz F, Gray GM, et al. Structural basis for gluten intolerance in celiac disease. *Science* 2002;297:2275–9.
- [31] Battais F, Richard C, Szustakowski G, et al. Wheat flour allergy: an entire diagnostic tool for complex allergy. *Allerg Immunol (Paris)* 2006;38:59–61.
- [32] Palosuo K. Update on wheat hypersensitivity. *Curr Opin Allergy Clin Immunol* 2003;3:205–9.
- [33] Takizawa T, Arakawa H, Tokuyama K, Morikawa A. Identification of allergen fractions of wheat flour responsible for anaphylactic reactions to wheat products in infants and young children. *Int Arch Allergy Immunol* 2001;125:51–6.
- [34] Morita E, Kameyoshi Y, Mihara S, Hiragun T, Yamamoto S. gamma-Gliadin: a presumptive allergen causing wheat-dependent exercise-induced anaphylaxis. *Br J Dermatol* 2001;145:182–4.
- [35] Armentia A, Rodriguez R, Callejo A, et al. Allergy after ingestion or inhalation of cereals involves similar allergens in different ages. *Clin Exp Allergy* 2002;32:1216–22.
- [36] Shewry PR, Tatham AS. The prolamins storage proteins of cereal seeds: structure and evolution. *Biochem J* 1990;267:1–12.
- [37] Gänzle MG, Lõponen J, Gobbetti M. Proteolysis in sourdough fermentations: mechanisms and potential for improved bread quality. *Trends Food Sci Technol* 2008;19:513–21.
- [38] Salcedo G, Quirce S, Diaz-Perales A. Wheat allergens associated with baker's asthma. *J Investig Allergol Clin Immunol* 2011;21:81–92.
- [39] Weiss W, Huber G, Engel KH, Pethran A, Dunn MJ, Gooley AA, et al. Identification and characterization of wheat grain albumin/globulin allergens. *Electrophoresis* 1997;18:826–33.
- [40] Posch A, Weiss W, Wheeler C, Dunn MJ, Görg A. Sequence analysis of wheat grain allergens separated by two dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 1995;16:1115–9.
- [41] Sander I, Flagge A, Merget R, Halder TM, Meyer HE, Baur X. Identification of wheat flour allergens by means of 2-dimensional immunoblotting. *J Allergy Clin Immunol* 2001;107:907–13.
- [42] Sandiford CP, Tee RD, Newman Taylor AJ. The role of cereal and fungal amylases in cereal flour hypersensitivity. *Clin Exp Allergy* 1994;24:549–57.

- [43] Masci S, D'Ovidio R, Lafiandra D, Kasarda DD. Characterization of a low-molecular-weight glutenin subunit gene from bread wheat and the corresponding protein that represents a major subunit of the glutenin polymer. *Plant Physiol* 1998;118:1147–58.
- [44] Kasarda DD. Glutenin structure in relation to wheat quality. *Wheat is unique*. St. Paul, MN: American Association of Cereal Chemistry; 1989. p. 277–302.
- [45] Gu YQ, Anderson OD, Londeoré CF, Kong X, Chibbar RN, Lazo GR. Structural organization of the barley D-hordein locus in comparison with its orthologous regions of wheat genomes. *Genome* 2003;46:1084–97.
- [46] Jolly CJ, Rahman S, Kortt AA, Higgins TJV. Characterisation of the wheat Mr 15000 “grain-softness protein” and analysis of the relationship between its accumulation in the whole seed and grain softness. *Theor Appl Genet* 1993;86:589–97.
- [47] Morrison WR, Greenwell P, Law CN, Sulaiman BD. Occurrence of friabilin, a low molecular weight protein associated with grain softness, on starch granules isolated from some wheats and related species. *J Cereal Sci* 1992;15:143–9.
- [48] Rahman S, Jolly CJ, Skeritt JH, Walloscheck A. Cloning of a wheat 15-kDa grain softness protein (GSP). GSP is a mixture of puroindoline-like polypeptides. *Eur J Biochem* 1994;223: 917–25.